Draft Document for Review January 5, 2017 7:42 am SG24-7938-05

Redbooks
ibm.com/redbooks

# Implementing the IBM Storwize V7000 and IBM Spectrum Virtualize V7.8

Jon Tate

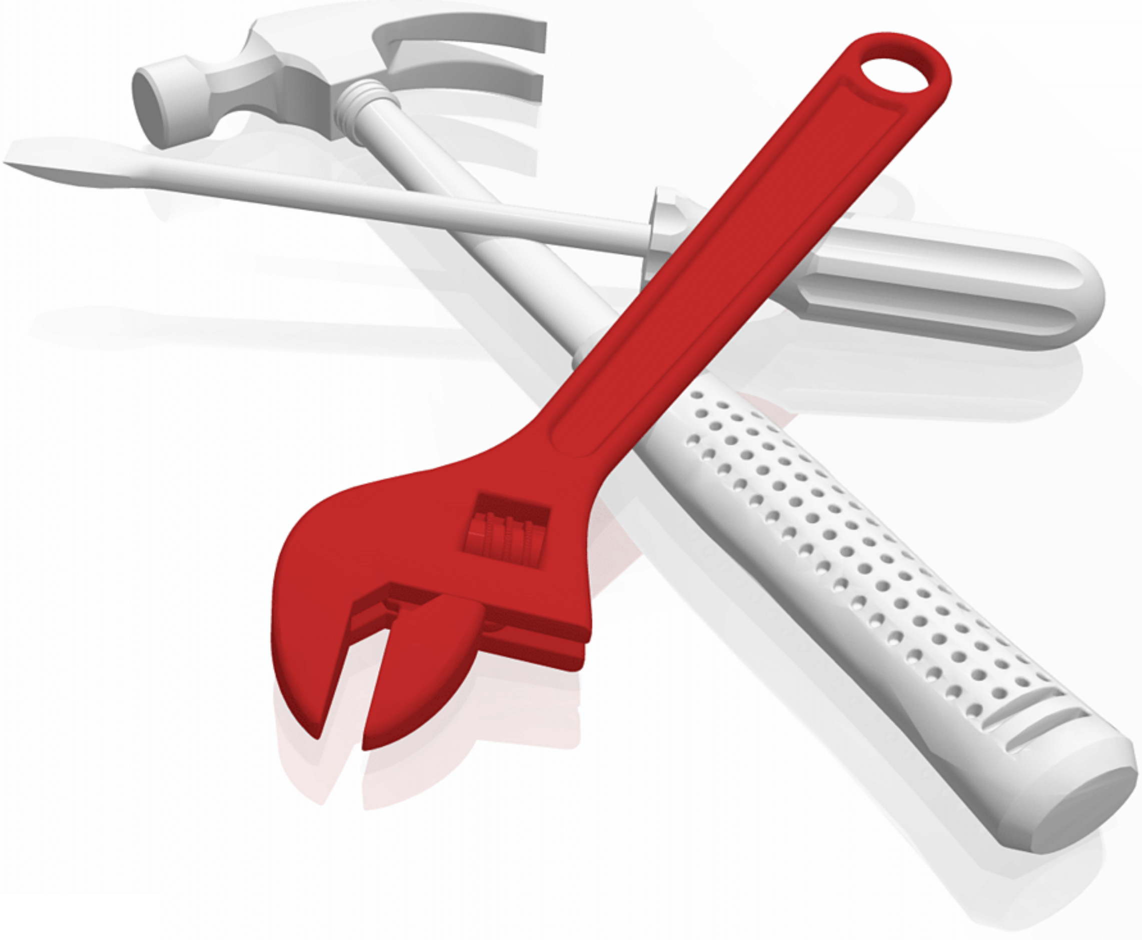
Catarina Castro

Frank Enders

Giulio Fiscella

Dharmesh Kamdar

Paulo Tomiyoshi Takeda



IBM

Redbooks

International Technical Support Organization

# Implementing the IBM Storwize V7000 and IBM Spectrum Virtualize V7.8

December 2016

SG24-7938-05

**Note:** Before using this information and the product it supports, read the information in “Notices” on page xiii.

**Sixth Edition (December 2016)**

This edition applies to IBM Spectrum Virtualize V7.8, and the associated hardware and software detailed within. Note that screenshots may differ from the generally available (GA) version as parts of this book were written with pre-GA code.

This document was created or updated on January 5, 2017.

# Contents

**Notices** . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . xiii
Trademarks . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . xiv

**Preface** . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . xv
Authors. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . xv
Now you can become a published author, too! . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . xvii
Comments welcome. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . xviii
Stay connected to IBM Redbooks . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . xviii

**Summary of changes**. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . xix
December 2016, Sixth Edition . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . xix

**Chapter 1. Introduction to storage virtualization** . . . . . . . . . . . . . . . . . . . . . . . . . . . 1
1.1 Storage virtualization terminology . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 2
1.2 Requirements driving storage virtualization . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 5
1.2.1 Benefits of using IBM Spectrum Virtualize . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 5
1.3 Latest changes and enhancements . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 6
1.3.1 IBM Storwize V7000 Gen2+ . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 6
1.3.2 IBM Spectrum Virtualize Software changes and enhancements . . . . . . . . . . . . . . 7
1.4 Summary. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 9

**Chapter 2. System Overview** . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 11
2.1 IBM Spectrum Storage and IBM Spectrum Virtualize . . . . . . . . . . . . . . . . . . . . . . . 12
2.1.1 IBM Spectrum Virtualize . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 12
2.2 Storage virtualization. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 13
2.3 IBM Storwize V7000 overview . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 14
2.3.1 IBM Storwize V7000 models. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 16
2.3.2 IBM Storwize V7000 functions . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 17
2.3.3 IBM Storwize V7000 licensing. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 19
2.4 IBM Storwize V7000 hardware . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 20
2.5 IBM Storwize V7000 components . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 20
2.5.1 Hosts. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 21
2.5.2 Host Cluster . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 21
2.5.3 Nodes . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 21
2.5.4 I/O groups . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 21
2.5.5 Cache . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 22
2.5.6 Clustered system . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 24
2.5.7 HyperSwap . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 24
2.5.8 Dense Expansion Drawers . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 25
2.5.9 Expansion Enclosures . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 26
2.5.10 RAID . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 27
2.5.11 Read Intensive Flash Drives . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 28
2.5.12 Managed disks . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 28
2.5.13 Quorum disks . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 29
2.5.14 IP Quorum. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 29
2.5.15 Storage pool . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 30
2.5.16 Volumes . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 31
2.5.17 Easy Tier. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 33
2.5.18 Encryption . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 34

2.5.19 iSCSI . . . 34
2.5.20 Real-time Compression. . . . 35
2.5.21 IP replication . . . 35
2.5.22 IBM Storwize V7000 copy services. . . . 36
2.5.23 Synchronous or asynchronous remote copy. . . . 36
2.5.24 FlashCopy and Transparent Cloud Tiering . . . 37
2.6 Business Continuity. . . . 38
2.6.1 Business Continuity with HyperSwap . . . 38
2.7 Management and support tools. . . . 39
2.7.1 IBM Assist On-site and remote service . . . 39
2.7.2 Event notifications. . . . 39
2.7.3 Simple Network Management Protocol traps . . . 40
2.7.4 Syslog messages . . . 40
2.7.5 Call Home email . . . 40
2.8 Useful IBM Storwize V7000 websites . . . 40

**Chapter 3. Planning** . . . 43
3.1 General planning rules . . . 44
3.2 Physical planning . . . 45
3.2.1 Cable connections . . . 45
3.3 Logical planning . . . 46
3.3.1 Management IP addressing plan . . . 46
3.3.2 SAN zoning and SAN connections . . . 50
3.3.3 iSCSI IP addressing plan . . . 52
3.3.4 iSCSI external storage attachment . . . 52
3.3.5 External storage configuration planning . . . 52
3.3.6 Host failover . . . 55
3.3.7 IP mirroring . . . 56
3.3.8 Planning for external storage virtualization . . . 61
3.3.9 Storwize V7000 clustered system configuration . . . 63
3.3.10 Storage pool configuration . . . 64
3.3.11 Volume configuration . . . 66
3.3.12 Host attachment planning . . . 68
3.3.13 Host mapping (LUN masking). . . . 72
3.3.14 NPIV planning. . . . 73
3.3.15 Advanced Copy Services . . . 73
3.3.16 SAN boot support . . . 80
3.3.17 Data migration from a non-virtualized storage subsystem . . . 80
3.3.18 Storwize V7000 configuration backup procedure . . . 81
3.4 Performance considerations . . . 82
3.4.1 SAN. . . . 82
3.4.2 Disk subsystems . . . 82
3.4.3 Storwize V7000. . . . 83
3.4.4 IBM Real-time Compression . . . 84
3.4.5 Performance monitoring . . . 84

**Chapter 4. Initial configuration** . . . 85
4.1 Prerequisites . . . 86
4.2 System initialization. . . . 86
4.2.1 System initialization wizard . . . 87
4.3 System setup . . . 90
4.3.1 System setup wizard. . . . 90
4.3.2 Adding a control enclosure . . . 103

4.3.3 Adding an expansion enclosure . . . . . . . . . . 105
4.4 Configuring user authentication . . . . . . . . . . 108
4.4.1 Default superuser account . . . . . . . . . . 108
4.4.2 Local authentication . . . . . . . . . . 108
4.4.3 Remote authentication . . . . . . . . . . 109
4.4.4 User groups and roles. . . . . . . . . . 118
4.5 Configuring secure communications . . . . . . . . . . 118
4.5.1 Configuring a signed certificate. . . . . . . . . . 119
4.5.2 Generating a self-signed certificate. . . . . . . . . . 122

**Chapter 5. Graphical User Interface** . . . . . . . . . . 125
5.1 Normal operations using the GUI . . . . . . . . . . 126
5.2 Introduction to the GUI . . . . . . . . . . 126
5.2.1 Dynamic menu . . . . . . . . . . 126
5.2.2 Suggested tasks . . . . . . . . . . 128
5.2.3 Notification status area . . . . . . . . . . 128
5.3 Overview window . . . . . . . . . . 131
5.3.1 Content view organization. . . . . . . . . . 131
5.3.2 Help. . . . . . . . . . 134
5.4 Viewing Cluster settings using the GUI. . . . . . . . . . 135
5.4.1 System status information. . . . . . . . . . 135
5.4.2 View IBM Storwize V7000 system properties . . . . . . . . . . 136
5.4.3 Renaming the IBM Storwize V7000 system . . . . . . . . . . 137
5.4.4 Rename a node canister. . . . . . . . . . 139
5.5 System overview . . . . . . . . . . 141
5.6 Monitoring menu . . . . . . . . . . 143
5.6.1 System details. . . . . . . . . . 144
5.6.2 Events . . . . . . . . . . 145
5.6.3 Performance . . . . . . . . . . 146
5.7 Network menu . . . . . . . . . . 147
5.7.1 Configuring the network . . . . . . . . . . 147
5.7.2 Management IP addresses . . . . . . . . . . 147
5.7.3 Service IP Information. . . . . . . . . . 148
5.7.4 iSCSI information . . . . . . . . . . 149
5.7.5 Fibre Channel information. . . . . . . . . . 150
5.8 Notifications menu. . . . . . . . . . 151
5.8.1 Email notifications . . . . . . . . . . 151
5.8.2 SNMP notifications . . . . . . . . . . 153
5.8.3 Syslog notifications . . . . . . . . . . 154
5.9 System menu . . . . . . . . . . 155
5.9.1 Date and time . . . . . . . . . . 155
5.9.2 Licensing. . . . . . . . . . 157
5.9.3 VMware Virtual Volumes. . . . . . . . . . 157
5.9.4 IP Quorum. . . . . . . . . . 158
5.9.5 Domain Name Server . . . . . . . . . . 159
5.9.6 Transparent Cloud Tiering . . . . . . . . . . 160
5.10 Setting GUI preferences . . . . . . . . . . 161
5.10.1 Navigation . . . . . . . . . . 162
5.10.2 Login Message . . . . . . . . . . 162
5.10.3 General settings . . . . . . . . . . 163

**Chapter 6. Storage pools** . . . . . . . . . . 165
6.1 Working with storage pools . . . . . . . . . . 166

6.1.1 Creating storage pools . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 167
6.1.2 Actions on storage pools. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 170
6.1.3 Child storage pools . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 171
6.1.4 Encrypted storage pools . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 174
6.2 Working with managed disks . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 175
6.2.1 Assigning managed disks to storage pools. . . . . . . . . . . . . . . . . . . . . . . . . . . 176
6.2.2 Traditional and distributed RAID . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 183
6.2.3 Actions on arrays . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 186
6.2.4 Actions on external MDisks. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 188
6.2.5 Additional actions on MDisks . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 194
6.3 Working with internal drives . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 195
6.3.1 Actions on internal drives . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 196
6.4 Working with external storage controllers . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 202
6.4.1 Fibre Channel external storage controllers . . . . . . . . . . . . . . . . . . . . . . . . . . . 203
6.4.2 iSCSI external storage controllers. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 204
6.4.3 Actions on external storage controllers . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 209

**Chapter 7. Volumes** . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 213
7.1 Introduction to volumes . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 214
7.1.1 Image mode volumes . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 215
7.1.2 Managed mode volumes. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 216
7.1.3 Cache mode for volumes . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 217
7.1.4 Mirrored volumes . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 218
7.1.5 Thin-provisioned volumes . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 221
7.1.6 Compressed Volumes. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 223
7.1.7 Volumes for various topologies. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 223
7.2 Create Volumes menu . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 224
7.3 Creating volumes using the Quick Volume Creation. . . . . . . . . . . . . . . . . . . . . . . . . 228
7.3.1 Creating Basic volumes using Quick Volume Creation. . . . . . . . . . . . . . . . . . . . 228
7.3.2 Creating Mirrored volumes using Quick Volume Creation . . . . . . . . . . . . . . . . . 231
7.4 Mapping a volume to the host . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 234
7.5 Creating Custom volumes using the Advanced menu . . . . . . . . . . . . . . . . . . . . . . . 236
7.5.1 Creating a custom thin-provisioned volume . . . . . . . . . . . . . . . . . . . . . . . . . . . 237
7.5.2 Creating Custom Compressed volumes . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 240
7.5.3 Custom Mirrored Volumes . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 241
7.6 HyperSwap and the mkvolume command . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 244
7.6.1 Volume manipulation commands . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 247
7.7 Mapping Volumes to Host after volume creation . . . . . . . . . . . . . . . . . . . . . . . . . . . 249
7.7.1 Mapping newly created volumes to the host using the wizard . . . . . . . . . . . . . . 249
7.8 Migrating a volume to another storage pool . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 251
7.9 Migrating volumes using the volume copy feature . . . . . . . . . . . . . . . . . . . . . . . . . . 254
7.10 Volume operations using CLI . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 258
7.10.1 Creating a volume. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 258
7.10.2 Volume information . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 260
7.10.3 Creating a thin-provisioned volume. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 260
7.10.4 Creating a volume in image mode . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 261
7.10.5 Adding a mirrored volume copy . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 262
7.10.6 Adding a compressed volume copy . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 266
7.10.7 Splitting a mirrored volume . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 267
7.10.8 Modifying a volume . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 268
7.10.9 I/O governing . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 269
7.10.10 Deleting a volume . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 271
7.10.11 Using volume protection . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 271

7.10.12 Expanding a volume . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 272
7.10.13 Assigning a volume to a host . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 273
7.10.14 Showing volumes to host mapping . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 274
7.10.15 Deleting a volume to host mapping. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 275
7.10.16 Migrating a volume . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 275
7.10.17 Migrating a fully managed volume to an image mode volume . . . . . . . . . . . . 276
7.10.18 Shrinking a volume . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 276
7.10.19 Showing a volume on an MDisk . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 277
7.10.20 Showing which volumes are using a storage pool . . . . . . . . . . . . . . . . . . . . 278
7.10.21 Showing which MDisks are used by a specific volume. . . . . . . . . . . . . . . . . . 278
7.10.22 Showing from which storage pool a volume has its extents . . . . . . . . . . . . . . 278
7.10.23 Showing the host to which the volume is mapped . . . . . . . . . . . . . . . . . . . . 280
7.10.24 Showing the volume to which the host is mapped . . . . . . . . . . . . . . . . . . . . 280
7.10.25 Tracing a volume from a host back to its physical disk. . . . . . . . . . . . . . . . . . 280

**Chapter 8. Hosts** . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 283
8.1 Host attachment overview. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 284
8.2 Host Clusters. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 285
8.3 N-Port Virtualization ID (NPIV) Support . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 285
8.3.1 NPIV Pre-requisites . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 287
8.3.2 Enabling NPIV on a new system. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 288
8.3.3 Enabling NPIV on an existing system . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 291
8.4 Hosts operations using the GUI . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 293
8.4.1 Creating hosts. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 294
8.4.2 Advanced host administration. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 304
8.4.3 Adding and deleting host ports . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 325
8.4.4 Host mappings overview. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 334
8.5 Hosts operations using **CLI. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 336**
8.5.1 Create a host using CLI . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 336
8.5.2 Advanced host administration using CLI. . . . . . . . . . . . . . . . . . . . . . . . . . . . 338
8.5.3 Adding and deleting a host port using CLI . . . . . . . . . . . . . . . . . . . . . . . . . . 340
8.5.4 Host clusters . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 342
8.5.5 Host Cluster Operations . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 343

**Chapter 9. Storage migration**. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 347
9.1 Storage migration overview. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 348
9.1.1 Interoperability and compatibility. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 348
9.1.2 Prerequisites . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 349
9.2 Storage migration wizard . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 349

**Chapter 10. Advanced features for storage efficiency** . . . . . . . . . . . . . . . . . . . . 363
10.1 Introduction . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 364
10.2 Easy Tier . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 365
10.2.1 Easy Tier concepts . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 365
10.2.2 SSD arrays and flash MDisks . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 367
10.2.3 Disk tiers . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 370
10.2.4 Read Intensive flash drive and Easy Tier . . . . . . . . . . . . . . . . . . . . . . . . . . 373
10.2.5 Easy Tier process . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 374
10.2.6 Easy Tier operating modes . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 375
10.2.7 Implementation considerations . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 377
10.2.8 Modifying the Easy Tier setting. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 378
10.2.9 Monitoring tools. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 388
10.2.10 More information . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 392
10.3 Thin provisioning . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 392

10.3.1 Configuring a thin-provisioned volume . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 393
10.3.2 Performance considerations . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 396
10.3.3 Limitations of virtual capacity . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 397
10.4 Real-time Compression software . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 398
10.4.1 Common use cases . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 398
10.4.2 Real-time Compression concepts . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 399
10.4.3 Random Access Compression Engine . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 400
10.4.4 Dual RACE component. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 404
10.4.5 Random Access Compression Engine in IBM Spectrum Virtualize stack. . . . . 405
10.4.6 Data write flow . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 406
10.4.7 Data read flow. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 406
10.4.8 Compression of existing data . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 407
10.4.9 Configuring compressed volumes. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 408
10.4.10 Comprestimator . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 411

**Chapter 11. Advanced Copy Services**. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 413
11.1 FlashCopy . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 414
11.1.1 Business requirements for FlashCopy . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 414
11.1.2 Backup improvements with FlashCopy. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 414
11.1.3 Restore with FlashCopy . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 415
11.1.4 Moving and migrating data with FlashCopy . . . . . . . . . . . . . . . . . . . . . . . . . . 415
11.1.5 Application testing with FlashCopy . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 416
11.1.6 Host and application considerations to ensure FlashCopy integrity . . . . . . . . . 416
11.1.7 FlashCopy attributes . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 416
11.1.8 Reverse FlashCopy . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 417
11.1.9 IBM Spectrum Protect Snapshot. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 418
11.2 FlashCopy functional overview . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 419
11.3 Implementing FlashCopy . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 420
11.3.1 FlashCopy mappings . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 420
11.3.2 Multiple Target FlashCopy . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 421
11.3.3 Consistency Groups . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 422
11.3.4 FlashCopy indirection layer. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 424
11.3.5 Grains and the FlashCopy bitmap. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 424
11.3.6 Interaction and dependency between multiple target FlashCopy mappings. . . 426
11.3.7 Summary of the FlashCopy indirection layer algorithm. . . . . . . . . . . . . . . . . . . 428
11.3.8 Interaction with the cache . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 428
11.3.9 FlashCopy and image mode volumes. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 429
11.3.10 FlashCopy mapping events . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 432
11.3.11 FlashCopy mapping states . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 434
11.3.12 Thin provisioned FlashCopy . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 436
11.3.13 Background copy . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 436
11.3.14 Serialization of I/O by FlashCopy . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 437
11.3.15 Event handling . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 438
11.3.16 Asynchronous notifications . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 439
11.3.17 Interoperation with Metro Mirror and Global Mirror . . . . . . . . . . . . . . . . . . . . 439
11.3.18 FlashCopy presets . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 440
11.4 Managing FlashCopy using GUI . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 441
11.4.1 Creating a FlashCopy mapping. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 443
11.4.2 Single-click snapshot . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 450
11.4.3 Single-click clone . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 451
11.4.4 Single-click backup . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 452
11.4.5 Creating a FlashCopy Consistency Group . . . . . . . . . . . . . . . . . . . . . . . . . . . 453
11.4.6 Creating FlashCopy mappings in a Consistency Group . . . . . . . . . . . . . . . . . . 454

11.4.7 Showing related volumes . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 458
11.4.8 Moving a FlashCopy mapping to a Consistency Group . . . . . . . . . . . . . . . . . . 459
11.4.9 Removing a FlashCopy mapping from a Consistency Group . . . . . . . . . . . . . . 460
11.4.10 Modifying a FlashCopy mapping. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 461
11.4.11 Renaming FlashCopy mapping. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 462
11.4.12 Renaming Consistency Group . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 463
11.4.13 Deleting FlashCopy mapping . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 463
11.4.14 Deleting FlashCopy Consistency Group . . . . . . . . . . . . . . . . . . . . . . . . . . . . 465
11.4.15 Starting FlashCopy process . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 465
11.4.16 Stopping FlashCopy process . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 466
11.5 Transparent Cloud Tiering . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 467
11.5.1 Considerations for using Transparent Cloud Tiering. . . . . . . . . . . . . . . . . . . . 468
11.5.2 Transparent Cloud Tiering as a backup solution and data migration . . . . . . . . 468
11.5.3 Restore using Transparent Cloud Tiering . . . . . . . . . . . . . . . . . . . . . . . . . . . . 469
11.5.4 Transparent Cloud Tiering restrictions . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 469
11.6 Implementing Transparent Cloud Tiering . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 470
11.6.1 DNS Configuration . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 470
11.6.2 Enabling Transparent Cloud Tiering . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 470
11.6.3 Creating and Managing Cloud Snapshots . . . . . . . . . . . . . . . . . . . . . . . . . . . 473
11.6.4 Restore Cloud Snapshots . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 476
11.7 Volume mirroring and migration options . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 479
11.8 Native IP replication . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 480
11.8.1 Native IP replication technology . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 480
11.8.2 IBM Storwize System Layers . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 482
11.8.3 IP partnership limitations. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 483
11.8.4 VLAN support . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 484
11.8.5 IP partnership and terminology . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 485
11.8.6 States of IP partnership . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 486
11.8.7 Remote copy groups. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 487
11.8.8 Supported configurations . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 488
11.9 Remote Copy . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 500
11.9.1 Multiple IBM Storwize V7000 system mirroring. . . . . . . . . . . . . . . . . . . . . . . . 501
11.9.2 Importance of write ordering . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 503
11.9.3 Remote copy intercluster communication . . . . . . . . . . . . . . . . . . . . . . . . . . . 505
11.9.4 Metro Mirror overview . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 506
11.9.5 Synchronous remote copy . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 507
11.9.6 Metro Mirror features. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 508
11.9.7 Metro Mirror attributes. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 509
11.9.8 Practical use of Metro Mirror. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 509
11.9.9 Global Mirror Overview . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 510
11.9.10 Asynchronous remote copy . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 511
11.9.11 Global Mirror features . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 512
11.9.12 Using Change Volumes with Global Mirror . . . . . . . . . . . . . . . . . . . . . . . . . . 514
11.9.13 Distribution of work among nodes. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 517
11.9.14 Background copy performance . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 517
11.9.15 Thin-provisioned background copy . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 518
11.9.16 Methods of synchronization . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 518
11.9.17 Practical use of Global Mirror . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 519
11.9.18 Valid combinations of FlashCopy, Metro Mirror, and Global Mirror . . . . . . . . 519
11.9.19 Remote Copy configuration limits . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 519
11.9.20 Remote Copy states and events. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 520
11.10 Remote Copy commands . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 527
11.10.1 Remote Copy process . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 527

11.10.2 Listing available system partners . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 528
11.10.3 Changing the system parameters . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 528
11.10.4 System partnership . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 530
11.10.5 Creating a Metro Mirror/Global Mirror consistency group . . . . . . . . . . . . . . . . 531
11.10.6 Creating a Metro Mirror/Global Mirror relationship . . . . . . . . . . . . . . . . . . . . . 531
11.10.7 Changing Metro Mirror/Global Mirror relationship . . . . . . . . . . . . . . . . . . . . . 531
11.10.8 Changing Metro Mirror/Global Mirror consistency group . . . . . . . . . . . . . . . . 532
11.10.9 Starting Metro Mirror/Global Mirror relationship . . . . . . . . . . . . . . . . . . . . . . 532
11.10.10 Stopping Metro Mirror/Global Mirror relationship . . . . . . . . . . . . . . . . . . . . 532
11.10.11 Starting Metro Mirror/Global Mirror consistency group . . . . . . . . . . . . . . . . 533
11.10.12 Stopping Metro Mirror/Global Mirror consistency group . . . . . . . . . . . . . . . 533
11.10.13 Deleting Metro Mirror/Global Mirror relationship . . . . . . . . . . . . . . . . . . . . 533
11.10.14 Deleting Metro Mirror/Global Mirror consistency group. . . . . . . . . . . . . . . . 533
11.10.15 Reversing Metro Mirror/Global Mirror relationship . . . . . . . . . . . . . . . . . . . 534
11.10.16 Reversing Metro Mirror/Global Mirror consistency group . . . . . . . . . . . . . . 534
11.11 Managing Remote Copy using the GUI . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 534
11.11.1 Creating Fibre Channel partnership . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 535
11.11.2 Creating stand-alone remote copy relationships. . . . . . . . . . . . . . . . . . . . . . . 537
11.11.3 Creating Consistency Group. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 541
11.11.4 Renaming Consistency Group . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 545
11.11.5 Renaming remote copy relationship . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 546
11.11.6 Moving stand-alone remote copy relationship to Consistency Group. . . . . . . 547
11.11.7 Removing remote copy relationship from Consistency Group . . . . . . . . . . . . 548
11.11.8 Starting remote copy relationship . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 549
11.11.9 Starting remote copy Consistency Group . . . . . . . . . . . . . . . . . . . . . . . . . . . 550
11.11.10 Switching copy direction . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 551
11.11.11 Switching the copy direction for a Consistency Group . . . . . . . . . . . . . . . . . 552
11.11.12 Stopping a remote copy relationship. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 554
11.11.13 Stopping Consistency Group . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 555
11.11.14 Deleting stand-alone remote copy relationships . . . . . . . . . . . . . . . . . . . . . 556
11.11.15 Deleting Consistency Group . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 558
11.12 Troubleshooting remote copy . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 558
11.12.1 1920 error . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 558
11.12.2 1720 error . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 560

**Chapter 12. Encryption**. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 563
12.1 Introducing encryption. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 564
12.2 Defining encryption of data at rest . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 564
12.2.1 Encryption keys. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 567
12.2.2 Encryption licenses . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 567
12.3 Activating encryption . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 568
12.3.1 Obtaining an encryption license . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 568
12.3.2 Start activation process during initial system setup . . . . . . . . . . . . . . . . . . . . 569
12.3.3 Start activation process on a running system . . . . . . . . . . . . . . . . . . . . . . . . . 571
12.3.4 Activate the license automatically. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 572
12.3.5 Activate the license manually . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 575
12.4 Enabling encryption. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 577
12.4.1 Starting the encryption enablement process. . . . . . . . . . . . . . . . . . . . . . . . . . 578
12.4.2 Enabling encryption using a key server . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 580
12.4.3 Enabling encryption using USB flash drives . . . . . . . . . . . . . . . . . . . . . . . . . . 585
12.5 Using encryption . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 589
12.5.1 Encrypted pools . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 589
12.5.2 Encrypted child pools . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 591

12.5.3 Encrypted arrays . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 593
12.5.4 Encrypted MDisks . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 594
12.5.5 Encrypted volumes . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 596
12.6 Rekeying an encryption-enabled system . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 597
12.6.1 Rekeying using a key server . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 597
12.6.2 Rekeying using USB flash drives . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 599
12.7 Disabling encryption . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 601
12.8 Restrictions . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 603

**Chapter 13. RAS, monitoring, and troubleshooting** . . . . . . . . . . . . . . . . . . . . . . . . . 605
13.1 Reliability, availability, and serviceability . . . . . . . . . . . . . . . . . . . . . . . . . . . . 606
13.1.1 Node canisters . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 606
13.1.2 Expansion canisters . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 608
13.1.3 Dense Drawer Enclosures LED . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 609
13.1.4 Enclosure SAS cabling . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 610
13.1.5 Power . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 612
13.2 Shutting down IBM Storwize V7000 . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 613
13.2.1 Shutting down a node canister . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 614
13.2.2 Shutting down a system . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 616
13.2.3 Shutting down and powering on an IBM Storwize V7000 infrastructure . . . . . . 617
13.3 Configuration backup . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 617
13.3.1 Backup using CLI . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 618
13.3.2 Backup using the GUI . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 619
13.4 Software update . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 620
13.4.1 Precautions before the update . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 620
13.4.2 IBM Storwize V7000 update test utility . . . . . . . . . . . . . . . . . . . . . . . . . . . 621
13.4.3 Update procedure to V7.8 . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 621
13.4.4 Updating the IBM Storwize V7000 cluster manually . . . . . . . . . . . . . . . . . . 629
13.4.5 Manually updating the Storwize V7000 . . . . . . . . . . . . . . . . . . . . . . . . . . . 635
13.5 Critical Fix Notification feature . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 636
13.6 Troubleshooting and fix procedures . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 636
13.6.1 Managing event log . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 638
13.6.2 Running a fix procedure . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 639
13.6.3 Event log details . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 643
13.7 Monitoring . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 644
13.7.1 Email notifications and the Call Home function . . . . . . . . . . . . . . . . . . . . . . 645
13.7.2 SNMP Configuration . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 650
13.7.3 Syslog notifications . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 651
13.8 Audit log . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 652
13.9 Collecting support information using the GUI and the CLI . . . . . . . . . . . . . . . . . . 654
13.9.1 Collecting information using the GUI. . . . . . . . . . . . . . . . . . . . . . . . . . . . . 654
13.9.2 Collecting logs using the CLI . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 657
13.9.3 Uploading files to the Support Center . . . . . . . . . . . . . . . . . . . . . . . . . . . 658
13.10 Service Assistant Tool. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 660

**Appendix A. Performance data and statistics gathering** . . . . . . . . . . . . . . . . . . . . . 663
Storwize V7000 performance overview . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 664
Performance considerations . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 664
IBM Spectrum Virtualize performance perspectives . . . . . . . . . . . . . . . . . . . . . . . 665
Performance monitoring. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 665
Collecting performance statistics . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 665
Real-time performance monitoring . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 667
Performance data collection and Tivoli Storage Productivity Center for Disk . . . . . . . . 674

**Appendix B. CLI setup and SAN boot** . . . . . . . . . . 675
CLI setup . . . . . . . . . . 676
Basic setup . . . . . . . . . . 676
SAN Boot . . . . . . . . . . 687
Enabling SAN Boot for Microsoft Windows . . . . . . . . . . 688
Enabling SAN Boot for VMware . . . . . . . . . . 688

**Appendix C. Terminology** . . . . . . . . . . 689
Commonly encountered terms . . . . . . . . . . 690

**Related publications** . . . . . . . . . . 707
IBM Redbooks . . . . . . . . . . 707
Other resources . . . . . . . . . . 708
Referenced websites . . . . . . . . . . 708
Help from IBM . . . . . . . . . . 709

# Notices

This information was developed for products and services offered in the US. This material might be available from IBM in other languages. However, you may be required to own a copy of the product or product version in that language in order to access it.

IBM may not offer the products, services, or features discussed in this document in other countries. Consult your local IBM representative for information on the products and services currently available in your area. Any reference to an IBM product, program, or service is not intended to state or imply that only that IBM product, program, or service may be used. Any functionally equivalent product, program, or service that does not infringe any IBM intellectual property right may be used instead. However, it is the user's responsibility to evaluate and verify the operation of any non-IBM product, program, or service.

IBM may have patents or pending patent applications covering subject matter described in this document. The furnishing of this document does not grant you any license to these patents. You can send license inquiries, in writing, to:
*IBM Director of Licensing, IBM Corporation, North Castle Drive, MD-NC119, Armonk, NY 10504-1785, US*

INTERNATIONAL BUSINESS MACHINES CORPORATION PROVIDES THIS PUBLICATION "AS IS" WITHOUT WARRANTY OF ANY KIND, EITHER EXPRESS OR IMPLIED, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF NON-INFRINGEMENT, MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE. Some jurisdictions do not allow disclaimer of express or implied warranties in certain transactions, therefore, this statement may not apply to you.

This information could include technical inaccuracies or typographical errors. Changes are periodically made to the information herein; these changes will be incorporated in new editions of the publication. IBM may make improvements and/or changes in the product(s) and/or the program(s) described in this publication at any time without notice.

Any references in this information to non-IBM websites are provided for convenience only and do not in any manner serve as an endorsement of those websites. The materials at those websites are not part of the materials for this IBM product and use of those websites is at your own risk.

IBM may use or distribute any of the information you provide in any way it believes appropriate without incurring any obligation to you.

The performance data and client examples cited are presented for illustrative purposes only. Actual performance results may vary depending on specific configurations and operating conditions.

Information concerning non-IBM products was obtained from the suppliers of those products, their published announcements or other publicly available sources. IBM has not tested those products and cannot confirm the accuracy of performance, compatibility or any other claims related to non-IBM products. Questions on the capabilities of non-IBM products should be addressed to the suppliers of those products.

Statements regarding IBM's future direction or intent are subject to change or withdrawal without notice, and represent goals and objectives only.

This information contains examples of data and reports used in daily business operations. To illustrate them as completely as possible, the examples include the names of individuals, companies, brands, and products. All of these names are fictitious and any similarity to actual people or business enterprises is entirely coincidental.

COPYRIGHT LICENSE:

This information contains sample application programs in source language, which illustrate programming techniques on various operating platforms. You may copy, modify, and distribute these sample programs in any form without payment to IBM, for the purposes of developing, using, marketing or distributing application programs conforming to the application programming interface for the operating platform for which the sample programs are written. These examples have not been thoroughly tested under all conditions. IBM, therefore, cannot guarantee or imply reliability, serviceability, or function of these programs. The sample programs are provided "AS IS", without warranty of any kind. IBM shall not be liable for any damages arising out of your use of the sample programs.

# Trademarks

IBM, the IBM logo, and ibm.com are trademarks or registered trademarks of International Business Machines Corporation, registered in many jurisdictions worldwide. Other product and service names might be trademarks of IBM or other companies. A current list of IBM trademarks is available on the web at "Copyright and trademark information" at http://www.ibm.com/legal/copytrade.shtml

The following terms are trademarks or registered trademarks of International Business Machines Corporation, and might also be trademarks or registered trademarks in other countries.

AIX®
DB2®
developerWorks®
DS4000®
DS5000™
DS8000®
Easy Tier®
FlashCopy®
GPFS™
HyperSwap®
IBM®
IBM FlashSystem®
IBM Spectrum™
IBM Spectrum Accelerate™
IBM Spectrum Control™
IBM Spectrum Protect™
IBM Spectrum Scale™
IBM Spectrum Storage™
IBM Spectrum Virtualize™
Jazz™
PowerHA®
Rational®
Real-time Compression™
Redbooks®
Redbooks (logo) ®
Storwize®
System Storage®
Tivoli®
XIV®

The following terms are trademarks of other companies:

Intel, Intel Xeon, Intel logo, Intel Inside logo, and Intel Centrino logo are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries.

Linux is a trademark of Linus Torvalds in the United States, other countries, or both.

Microsoft, Windows, and the Windows logo are trademarks of Microsoft Corporation in the United States, other countries, or both.

Java, and all Java-based trademarks and logos are trademarks or registered trademarks of Oracle and/or its affiliates.

UNIX is a registered trademark of The Open Group in the United States and other countries.

Other company, product, or service names may be trademarks or service marks of others.

# Preface

Continuing its commitment to developing and delivering industry-leading storage technologies, IBM® introduces the IBM Storwize® V7000 solution, powered by IBM Spectrum™ Virtualize, an innovative storage offering that delivers essential storage efficiency technologies and exceptional ease of use and performance, all integrated into a compact, modular design that is offered at a competitive, midrange price.

The IBM Storwize V7000 solution incorporates some of the top IBM technologies typically found only in enterprise-class storage systems, raising the standard for storage efficiency in midrange disk systems. This cutting-edge storage system extends the comprehensive storage portfolio from IBM and can help change the way organizations address the ongoing information explosion.

This IBM Redbooks® publication introduces the features and functions of the IBM Storwize V7000 and IBM Spectrum Virtualize™ V7.8 system through several examples. This book is aimed at pre-sales and post-sales technical support and marketing, storage administrators, and will help you understand the architecture of the Storwize V7000, how to implement it, and take advantage of the industry-leading functions and features.

## Authors

This book was produced by a team of specialists from around the world working at the International Technical Support Organization, San Jose Center.

**Jon Tate** is a Project Manager for IBM System Storage® SAN Solutions at the ITSO, San Jose Center. Before joining the ITSO in 1999, he worked in the IBM Technical Support Center, providing Level 2/3 support for IBM mainframe storage products. Jon has 31 years of experience in storage software and management, services, and support. He is an IBM Certified IT Specialist, an IBM SAN Certified Specialist, and is Project Management Professional (PMP) certified. He is also the UK Chairman of the Storage Networking Industry Association (SNIA).

**Catarina Castro** is a Software Engineer for IBM Systems at the Manchester Lab. She has worked with IBM SAN Volume Controller and the Storwize family for the past two years in a variety of test and development roles. She is currently designing and developing an upcoming feature for IBM Spectrum Virtualize.

**Frank Enders** has worked for the last ten years for EMEA Storwize Level 2 support in Mainz, Germany, and his duties include pre- and post-sales support. He has worked for IBM Germany for more than 20 years and started as a technician in disk production for IBM Mainz, and changed to magnetic head production four years later. When IBM closed disk production in Mainz in 2001 he changed his role and continued working for IBM within the ESCC Mainz as a team member of the Installation Readiness team for products such as the DS8000®, DS6000™, and the IBM System Storage SAN Volume Controller. During that time he studied for four years to gain a diploma in Electrical Engineering.

**Giulio Fiscella** is a Software Engineer for IBM Systems at the IBM Manchester Lab. He works on the whole IBM Spectrum Virtualize portfolio developing and testing new functionality. He has worked on several projects since the V7.2.0 release and he was fully responsible for different features released in V7.7.1 and V7.8.0.

**Dharmesh Kamdar** has been working in IBM Systems group for over 14 years as a Senior Software Engineer. He is working in the Open Systems Lab (OSL), where he focuses on interoperability testing of a range of IBM storage products with various vendor products, including operating systems, clustering solutions, virtualization platforms, volume managers, and file systems.

**Paulo Tomiyoshi Takeda** is a SAN and Storage Disk specialist at IBM Brazil. He has over nine years of experience in the IT arena and is an IBM Certified IT Specialist. He holds a bachelors degree in Information Systems from UNIFEB (Universidade da Fundação Educacional de Barretos) and is IBM Certified for IBM DS8000® and IBM Storwize V7000. His areas of expertise include planning, configuring, and troubleshooting DS8000 SAN Volume Controller and IBM Storwize V7000. He is involved in storage-related projects such as capacity growth planning, SAN consolidation, storage microcode upgrades, and copy services in the Open Systems environment.

Thanks to the authors of the previous edition:

Maximilian Hart, Hartmut Lonzer, Tarik Jose Maluf, Libor Miklas, Jon Parkes, Anthony Saine, Lev Sturmer, Marcin Tabinowski

Thanks to the following people for their contributions to this project:

Christopher Bulmer, Paul Cashman, Carlos Fuente, Katja Gebuhr, Warren Hawkins, Gareth Jones, Evelyn Perez, Mark Visser, Stephen Wright
**IBM Hursley, UK**

Nick Clayton
**IBM Systems, UK**

Navin Manohar
Terry Niemeyer
**IBM Systems, US**

Chris Saul
**IBM Systems, US**

Barry Whyte
**IBM Systems, New Zealand**

Dhiraj K Verma, Akshat Mithal
**IBM Systems, India**

Ramapriya Krishnamurthy
Troy Lee
**IBM Systems, US**

Bill Wiegand
**IBM Systems, US**

Da Lu
**IBM Systems, China**

Angelo Bernasconi
**IBM Systems, Italy**

Antonio Rainero
**IBM GTS, Italy**

Special thanks to the Brocade Communications Systems staff in San Jose, California for their support of this residency in terms of equipment and support in many areas:

Silviano Gaona, Sangam Racherla, Brian Steffler, Marcus Thordal
**Brocade Communications Systems**

## Now you can become a published author, too!

Here's an opportunity to spotlight your skills, grow your career, and become a published author—all at the same time! Join an ITSO residency project and help write a book in your area of expertise, while honing your experience using leading-edge technologies. Your efforts will help to increase product acceptance and customer satisfaction, as you expand your network of technical contacts and relationships. Residencies run from two to six weeks in length, and you can participate either in person or as a remote resident working from your home base.

Find out more about the residency program, browse the residency index, and apply online at:

**ibm.com**/redbooks/residencies.html

# Comments welcome

Your comments are important to us!

We want our books to be as helpful as possible. Send us your comments about this book or other IBM Redbooks publications in one of the following ways:

- Use the online **Contact us** review Redbooks form found at:

  **ibm.com**/redbooks

- Send your comments in an email to:

  redbooks@us.ibm.com

- Mail your comments to:

  IBM Corporation, International Technical Support Organization
  Dept. HYTD Mail Station P099
  2455 South Road
  Poughkeepsie, NY 12601-5400

# Stay connected to IBM Redbooks

- Find us on Facebook:

  http://www.facebook.com/IBMRedbooks

- Follow us on Twitter:

  http://twitter.com/ibmredbooks

- Look for us on LinkedIn:

  http://www.linkedin.com/groups?home=&gid=2130806

- Explore new Redbooks publications, residencies, and workshops with the IBM Redbooks weekly newsletter:

  https://www.redbooks.ibm.com/Redbooks.nsf/subscribe?OpenForm

- Stay current on recent Redbooks publications with RSS Feeds:

  http://www.redbooks.ibm.com/rss.html

# Summary of changes

This section describes the technical changes made in this edition of the book and in previous editions. This edition might also include minor corrections and editorial changes that are not identified.

Summary of Changes
for SG24-7938-05
for Implementing the IBM Storwize V7000 and IBM Spectrum Virtualize V7.8
as created or updated on January 5, 2017.

## December 2016, Sixth Edition

This revision includes the changes to the GUI and CLI that support the GA release of Spectrum Virtualize V7.8 and the hardware described and supported at the time of writing.

# 1

# Introduction to storage virtualization

This chapter will defines the concept of *storage virtualization* and an overview on how to apply virtualization to address the challenges of storage requirements.

This chapter includes the following topics:

- Storage virtualization terminology
- Requirements driving storage virtualization
- Latest changes and enhancements
- Summary

## 1.1 Storage virtualization terminology

*Storage virtualization* is a term that is used extensively throughout the storage industry, it can be applied to various technologies and underlying capabilities. In reality, most storage devices technically can claim to be virtualized in one form or another. Therefore, this chapter starts by defining the concept of storage virtualization as it is used in this book.

IBM defines storage virtualization in the following manner:

- Storage virtualization is a technology that makes one set of resources resemble another set of resources, preferably with more desirable characteristics.
- It is a logical representation of resources that is not constrained by physical limitations and hides part of the complexity. It also adds or integrates new function with existing services and can be nested or applied to multiple layers of a system.

When the term storage virtualization is mentioned, it is important to understand that virtualization can be implemented at various layers within the I/O stack. There must be clearly distinguish between virtualization at the disk layer (block-based) and virtualization at the file system layer (file-based).

The focus of this publication is virtualization at the disk layer, which is referred to as *block-level virtualization*, or the *block aggregation layer*. A description of file system virtualization is beyond the scope of this book.

For more information about file system virtualization, see the following resource:

- IBM Spectrum Scale™ (based on IBM General Parallel File System, IBM GPFS™):

  https://ibm.biz/Bdr7ME

The Storage Networking Industry Association’s (SNIA) block aggregation model provides a useful overview of the storage domain and the layers, as shown in Figure 1-1 on page 3. It illustrates several layers of a storage domain:

- File
- Block aggregation
- Block subsystem layers.

The model splits the block aggregation layer into three sublayers. Block aggregation can be realized within hosts (servers), in the storage network (storage routers and storage controllers), or in storage devices (intelligent disk arrays).

The IBM implementation of a block aggregation solution is IBM Spectrum Virtualize software, running on IBM SAN Volume Controller and IBM Storwize family.

The IBM SAN Volume Controller is implemented as a clustered appliance in the storage network layer. The IBM Storwize family is deployed as modular storage that provides capabilities to virtualize its own internal storage and external storage.

The IBM Spectrum Virtualize software can also be deployed as *software-defined* storage solution on third-party x86-based platforms, but that is beyond the intended scope of this publication.

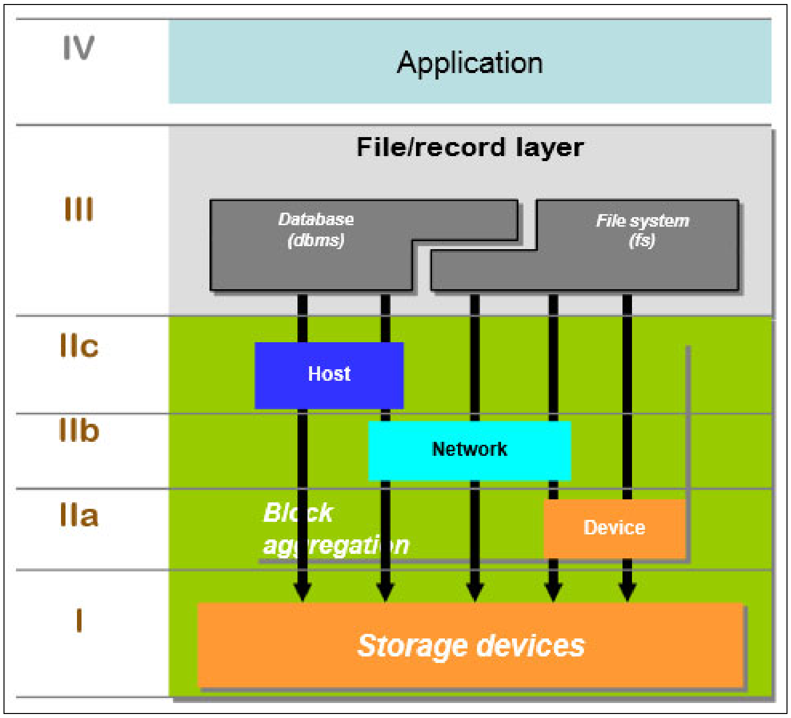


*Figure 1-1 SNIA block aggregation model*[1]

The key concept of virtualization is to decouple the storage from the storage functions that are required in the storage area network (SAN) environment.

*Decoupling* means abstracting the physical location of data from the logical representation of the data. The virtualization engine presents logical entities to the user and internally manages the process of mapping these entities to the actual location of the physical storage.

The actual mapping that is performed depends on the specific implementation, such as the granularity of the mapping, which can range from a small fraction of a physical disk up to the full capacity of a physical disk. A single block of information in this environment is identified by its *logical unit number* (LUN), which is the physical disk, and an offset within that LUN, which is known as a *logical block address* (LBA).

The term *physical disk* is used in this context to describe a piece of storage that might be carved out of a Redundant Array of Independent Disks (RAID) array in the underlying disk subsystem.

Specific to the IBM Spectrum Virtualize implementation, the address space that is mapped between the logical entity is referred to as a *volume*. The array of physical disks is referred to as *managed disks* (MDisks). The combination of several *managed disks* is referred to as a *Storage Pool.*

Figure 1-2 shows an overview of block-level virtualization.

[1] Source: Storage Networking Industry Association.

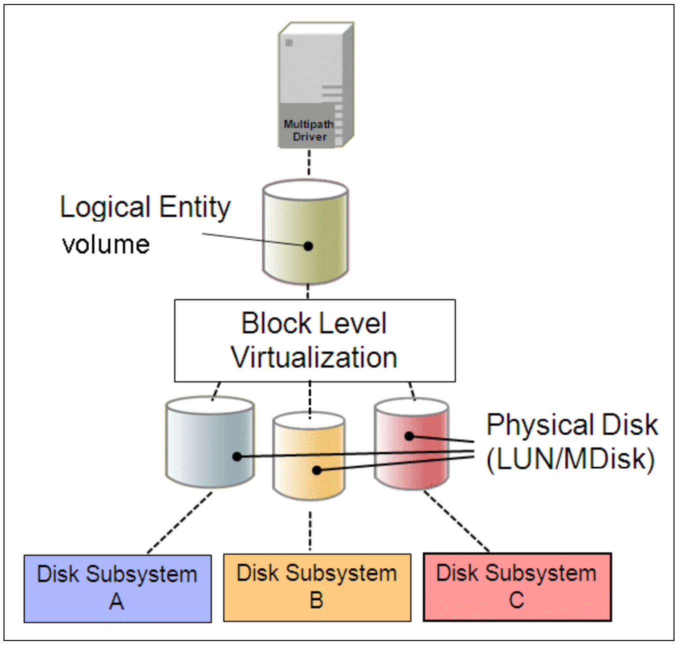


*Figure 1-2 Block-level virtualization overview*

The server and application are aware of the logical entities only, and they access these entities by using a consistent interface that is provided by the virtualization layer.

The functionality of a volume that is presented to a server, such as expanding or reducing the size of a volume, mirroring a volume, creating an IBM FlashCopy®, and thin provisioning, is implemented in the virtualization layer. It does not rely in any way on the functionality that is provided by the underlying disk subsystem. Data that is stored in a virtualized environment is stored in a location-independent way, which enables a user to move or migrate data between physical locations, which are referred to as *storage pools*.

The block-level storage virtualization can be referred as the *cornerstones of virtualization*. These cornerstones of virtualization are the core benefits that a product, such as IBM Spectrum Virtualize, can provide over the traditional directly attached or SAN storage.

IBM Spectrum Virtualize software provides the following benefits:

- Online volume migration while applications are running, which is possibly the greatest single benefit for storage virtualization. This capability enables data to be migrated on and between the underlying storage subsystems without any effect on the servers and applications. In fact, this migration is performed without the knowledge of the servers and applications that it even occurred.
- Simplified storage management by providing a single image for multiple controllers, and a consistent user interface for provisioning heterogeneous storage.
- Enterprise-level Copy Services functions. Performing Copy Services functions without the dependencies on the storage subsystems. Therefore, it enables the source and target copies to be on other storage subsystem types.
- Storage usage can be increased by pooling storage across the SAN.
- System performance is often improved with IBM Spectrum Virtualize and IBM Storwize V7000 as a result of volume striping across multiple arrays or controllers and the other cache that it provides.
- Software-based encryption capabilities to provide improved data security among storage virtualization solutions.

- Data replication to cloud storage using advanced copy services for data migration and backup solutions.

IBM Spectrum Virtualize software delivers all these functions in a homogeneous way on a scalable and high availability software platform over any attached storage and to any attached server.

## 1.2 Requirements driving storage virtualization

Today, many organizations are searching affordable and efficient ways to store, use, protect and manage the data. An emphasis is put on the IBM Cognitive era of clients' businesses and their dynamic infrastructure. Therefore, a storage environment requires an easy to manage interface and flexibility to support many applications, servers and mobility requirements.

Business demands change quickly.

The following key client concerns drive storage virtualization:

- Growth in data center costs
- Inability of information technology (IT) organizations to respond quickly to business demands
- Poor asset usage
- Poor availability or service levels
- Lack of skilled staff for storage administration

You can see the importance of addressing the complexity of managing storage networks by applying the total cost of ownership (TCO) metric to storage networks. Industry analyses show that storage acquisition costs are only about 20% of the TCO. Most of the remaining costs relate to managing the storage system.

But how much of the management of multiple systems, with separate interfaces, can be handled as a single entity? In a non-virtualized storage environment, every system is an "island" that must be managed separately.

### 1.2.1 Benefits of using IBM Spectrum Virtualize

IBM Storwize V7000 running IBM Spectrum Virtualize software, reduces the number of separate environments that must be managed down to a single environment. It also provides a single interface for storage management and a variety of functions. After the initial configuration of the storage subsystems, all of the day-to-day storage management operations are performed using the graphical user interface of the IBM Spectrum Virtualize.

Because IBM Spectrum Virtualize provides many functions, such as mirroring and ibm FlashCopy, there is no need to acquire additional subsets of applications for each attached disk subsystem that is virtualized by IBM Spectrum Virtualize.

Today, it is typical that open systems run at less than 50% of the usable capacity that is provided by the RAID disk subsystems. The use of the installed raw capacity in the disk subsystems shows usage numbers of less than 35%, depending on the RAID level that is used. A block-level virtualization solution, such as IBM Spectrum Virtualize, can allow a significant increase to approximately 75 - 80%.

With IBM Spectrum Virtualize, free space does not need to be maintained and managed in each storage subsystem, which further increases capacity usage.

## 1.3 Latest changes and enhancements

The IBM Spectrum V7.3 and its related hardware upgrade represented an important milestone in the product line development, with further enhancements up to and including V7.8. The internal architecture of IBM Spectrum Virtualize is significantly rebuilt, enabling the system to break the previous limitations in terms of scalability, flexibility and functionality.

The intent of this book is to cover the major software changes and provide a brief summary of supported hardware.

### 1.3.1 IBM Storwize V7000 Gen2+

The IBM Storwize V7000 Gen2 model is a modular, virtualized, enterprise-class storage solution technology. Each IBM Storwize V7000 module is delivered in 2U 19-inch rack-mounted enclosure. The IBM Storwize V7000 can be easily managed using a web-based graphical user interface and it provides a number of functions and features.

The control enclosure features two redundant controllers combined into a cluster with single management. Each controller contains one 8-core Intel Xeon processor with 32 GB or 64 GB of cache; one hardware compression accelerator card and an additional slot for the second accelerator.

A front view of the IBM Storwize V7000 Gen2 is shown in Figure 1-3.

*Figure 1-3 Front view of IBM Storwize V7000 Gen2*

Similar to the previous version, the IBM Storwize V7000 Gen2+ base model comes as 2U-19-inch rack-mounted enclosure and each controller features one 10-core Intel Xeon processor with 32 GB memory cache and optional cache upgrade to 64 GB. The IBM Storwize V7000 Gen 2 Plus, comes standard with integrated hardware compression card to support Real-time Compression workloads.

The IBM Storwize V7000 Gen 2 and Gen2 Plus support 1-Gb iSCSI connectivity as standard with options for 16-Gb Fibre Channel and 10-Gb for iSCSI and FCoE connectivity.

The IBM Storwize V7000 Gen2 is referred as model 524 and Gen2+ as 624. In this publication, both models of the IBM Storwize V7000 will be referred as only ‘IBM Storwize V7000’.

> **Note:** For comprehensive list of supported configurations refer to IBM Storwize V7000 configuration limits and restrictions in the following URL:
>
> http://www-01.ibm.com/support/docview.wss?uid=ssg1S1009561

### 1.3.2 IBM Spectrum Virtualize Software changes and enhancements

This section provides an overview of the relevant software enhancements incorporated to the IBM Storwize V7000 code.

At the V7.3 announcement, IBM included the following changes:

- IBM Easy Tier version 3 including storage pool balancing within the same tier and extent migration across any two adjacent tier in a three tiers storage pool. This function is enabled automatically in IBM Spectrum Virtualize software, and does not need any licenses.
- New cache architecture. Advanced cache algorithm splits the original single cache into upper and lower caches. The upper cache provides cache partitioning and fast write response times as it is the highest I/O stack layer. The lower cache provides a number of cache functions to support destage, read caching and prefetching.

At V7.4 IBM announced the following changes:

- The most noticeable change in V7.4 after the first login is the modified graphical user interface (GUI) with the new layout of the system panel, including enhanced functions that are available directly from the welcome window.
- The concept of the GUI design conforms to the well-known approach from IBM System Storage XIV Gen3 and IBM FlashSystem® 840. It provides common, unified procedures to manage all these systems in a similar way, enabling administrators to simplify their operational procedures across all systems.
- *Child pools* are new objects, which are created from the physical storage pool and provide most of the functions of managed disk groups (MDiskgrps), for example, volume creation. However, the user can specify the capacity of the child pool at creation.
- A new level of volume protection prevents users from removing mappings of volumes that are considered active. *Active* means that the system has detected recent I/O activity to the volume from any host within a protection period that is defined by the user. This behavior is enabled by system-wide policy settings. The detailed volume view contains the new field that indicates when the volume was last accessed.
- A user can replace a failed flash drive by removing it from the expansion enclosures unit and installing a new replacement drive, without requiring a Directed Maintenance Procedure (DMP) to supervise the action.
- The user determines that the fault light-emitting diode (LED) is illuminated for a drive, so they can expect to be able to reseat or replace the drive in that slot. The system automatically performs the drive hardware validation tests and promotes the unit into the configuration if these checks pass.
- Additional enhancements to T10 Data Integrity Field for the data stored on drives for IBM Storwize Gen2.
- Improved performance of Real-time Compression by double I/O operations per second (IOPS) on the model 2145-DH8 and 2076-524 (when it is equipped with both Compression Accelerator cards). It introduces two separate software compression engines (RACE), taking advantage of multi-core controller architecture. Hardware resources are shared between both RACE engines.

At V7.5 announcement, IBM included the following changes:

- Direct host attachment via 16 Gbps FC adapters with all operating systems *except* IBM AIX®.
- Support for Microsoft Offloaded Data Transfer (ODX).
- Introduction of IBM HyperSwap® topology. It enables each volume to be presented by two I/O groups. The configuration tolerates combinations of node and site failures, using a flexible choice of host multipathing driver interoperability.
- Support of VMware vSphere v6.0 Virtual Volumes (VVol). Each virtual machine (VM) keeps different types of data each in a VVol, each of which is presented as a volume (logical unit is SCSI) by the IBM Spectrum Virtualize system. Therefore, each VM owns a small number of volumes.

At V7.6, IBM announced the following changes:

- Visual and functional enhancements in the GUI, with changed menu layout and an integrated performance meter on main page.
- Implementation of Distributed RAID, which differs from traditional RAID arrays by eliminating dedicated spare drives. Spare capacity is rather spread across disks, making the reconstruction of failed disk faster.
- Introduced software encryption enabled by IBM Spectrum Virtualize and using AES256-XTS algorithm. Encryption is enabled on the storage pool level. All newly created volumes in such pool are automatically encrypted. An encryption license with Universal Serial Bus (USB) flash drives is required.
- Developed the *Comprestimator* tool, which is included in IBM Spectrum Virtualize software. It provides statistics to estimate potential storage savings. Available from the CLI, it does not need compression licenses and does not trigger any compression process. It uses the same estimation algorithm as an external host-based application, so results are similar.
- Enhanced GUI wizard for initial configuration of HyperSwap topology. IBM Spectrum Virtualize now allows IP-attached quorum disks in HyperSwap system configuration.
- Increased the maximum number of iSCSI hosts attached to the system to 2048 (512 host iSCSI qualified names (IQNs) per I/O group) with a maximum of four iSCSI sessions per SVC node (8 per I/O group).
- Improved and optimized read I/O performance in HyperSwap system configuration by parallel read from primary and secondary local volume copies. Both copies must be in a synchronized state.
- Extends the support of VVols. Using IBM Spectrum Virtualize you can manage one-to-one partnership of VM drives to IBM Storwize V7000 volumes. It eliminates single, shared volume (datastore) I/O contention.
- Customizable login banner. Using CLI commands, you can now define and show a welcome message or important disclaimer on the login window to users. This banner is shown in GUI or CLI login window.

At V7.7, IBM announced the following software changes:

- Introduction of new software capabilities to support encryption protection of data stored in D-RAID volumes.
- Enhanced external virtualization flexibility to support iSCSI-based external storage virtualization.
- IP Link compression algorithm to improve usage of IP networks. This will reduce the volume of data that must be transmitted during remote copy operations.
- GUI support for IP Quorum to help administrators to configure the IBM Storwize V7000 in a HyperSwap solution using IP Quorum solution.
- GUI support for Comprestimator to assist administrators to use the compression functions to display thin provisioning and compression estimates analysis for single or multiple volumes.
- Each clustered system can support up to 10.000 volumes.

With V7.8, IBM incorporated the following innovative changes as follows:

- GUI support for Cloud Storage. In the V7.8, IBM introduced software capabilities to interface the IBM Spectrum Virtualize to an external cloud storage service provider. This functionality is priced as additional per software code.
- Support for Dense Drawer which provides rack-mounted high density disk expansion enclosures that connects via SAS to Storwize V7000 enclosures. Each Dense Drawer can support up to 92 drives.
- Extended support for additional Tier 1 Solid State Drives (SSD). The IBM Spectrum Virtualize software features new attributes to manage Read Intensive SSDs; an alert is logged in the system event log when the drive endurance reaches 95%.
- Improved security by extending the encryption capabilities to software that complies with the Key Management Interoperability Protocol (KMIP), such as IBM Security Key Lifecycle manager (SKLM).
- Increased flexibility to support IBM Storwize model conversion to enable clients to convert existing systems to more powerful models.

## 1.4 Summary

The use of storage virtualization is the foundation for a flexible and reliable storage solution helps enterprises to better align business and IT by optimizing the storage infrastructure and storage management to meet business demands.

The IBM Spectrum Virtualize running on IBM Storwize V7000 Gen2 is a mature, eighth-generation virtualization solution that uses open standards and complies with the Storage Networking Industry Association (SNIA) storage model. The IBM Storwize V7000 is a powerful modular storage, in-band block virtualization process in which intelligence (including advanced storage functions) is ported from individual storage devices to the storage network.

IBM Spectrum Virtualize can improve the usage of the storage resources, simplify the storage management, and improve the availability of business applications.

2

# System Overview

This chapter provides an overview of IBM Spectrum Virtualize software and IBM Storwize V7000 architecture and components.

The initial section describes the IBM Spectrum family and the IBM virtualization portfolio. Then the next part covers the IBM Storwize V7000 hardware, components and the software elements that involves the IBM Storwize V7000 platform.

The last part of this chapter provides an overview of the useful management and support tools that helps to assist the management of the IBM Storwize V7000.

This chapter covers following topics:

- IBM Spectrum Storage and IBM Spectrum Virtualize
- Storage virtualization
- IBM Storwize V7000 overview
- IBM Storwize V7000 hardware
- IBM Storwize V7000 components
- Management and support tools
- Useful IBM Storwize V7000 websites

All of the topics we cover in this chapter are described in more detail later in this book.

## 2.1 IBM Spectrum Storage and IBM Spectrum Virtualize

IBM Spectrum Storage™ is an industry-leading, comprehensive, software-defined storage portfolio designed to address data storage inefficiencies by changing the economics of storage with a layer of intelligent software. The IBM Spectrum Storage family includes six members, creating an efficient data footprint that dynamically stores every bit of data at the optimal cost, helping maximize performance and ensuring security:

- *IBM Spectrum Accelerate™* is an agile, software-defined storage solution for enterprises and cloud that builds on the client-proven and mature IBM XIV storage software. The key characteristic of IBM Spectrum Accelerate is that it can be easily deployed and run on purpose-built or existing hardware (H/W) chosen by the customer.
- *IBM Spectrum Scale™* is a proven, scalable, high-performance data and file management solution intended to be used by diverse workloads where performance, reliability, and availability of data are essential to the business. It is based on the former IBM General Parallel File System (IBM GPFS™).
- *IBM Spectrum Virtualize* is an industry-leading storage virtualization solution that enhances existing storage to improve resource usage and productivity to achieve a simpler, more scalable, and cost-efficient information technology (IT) infrastructure. It is a software core composed of the IBM SAN Volume Controller and IBM Storwize family of products.
- *IBM Spectrum Control™* is an on-premise software solution that provides visibility into how your storage is performing and whether it is tiered correctly to reduce costs. Based on IBM Tivoli® Storage Productivity Center, it provides consumer-oriented features and enhancements that offer the ability to monitor and manage your storage infrastructure efficiently.
- *IBM Spectrum Protect* is an industry-leading data protection solution that enables advanced data backup and data recovery features for virtual, physical, cloud, and software-defined environments – as well as core applications and remote facilities. It was formerly known as IBM Tivoli Storage Manager.
- *IBM Spectrum Archive™* together with IBM Linear Tape File System™ (LTFS) provides intuitive drag-and-drop based archiving functions to users without any specific need for device-specific software. It conforms to data retention compliance requirements with user-defined retention policies. It was formerly known as IBM System Storage Archive Manager.

This publication focus on the IBM Storwize V7000 products and describes how IBM Spectrum Virtualize V7.8 powering the IBM Storwize V7000 improves storage efficiency and enables data transparency in strategic data centers or small to medium computer rooms.

### 2.1.1 IBM Spectrum Virtualize

IBM Spectrum Virtualize is a software-enabled storage virtualization engine that provides a single point of control for storage resources within the data centers. IBM Spectrum Virtualize is a core software engine of well-established and industry-proven IBM storage virtualization solutions, such as IBM SAN Volume Controller and all versions of IBM Storwize family of products (IBM Storwize V3700, IBM Storwize V5000, IBM Storwize V7000, and IBM FlashSystem™ V9000).

> **Naming:** With the introduction of the IBM Spectrum Storage family, the *software* that runs on IBM SAN Volume Controller and IBM Storwize family products is called IBM Spectrum Virtualize. The name of the underlying *hardware* platform remains intact.

The objectives of IBM Spectrum Virtualize are to manage storage resources in your IT infrastructure, and to ensure that they are used to the advantage of your business. These processes take place quickly, efficiently, and in real time, while avoiding increases in administrative costs.

Although IBM Spectrum Virtualize is a core software engine of the whole family of IBM Storwize products (see Figure 2-1). The contents of this book is intentionally related to the deployment considerations of IBM Storwize V7000 Gen2 and Gen2+.

Throughout this book, the term "IBM Storwize V7000" is used to refer to both models of the IBM Storwize V7000 when the text applies similarly to both.

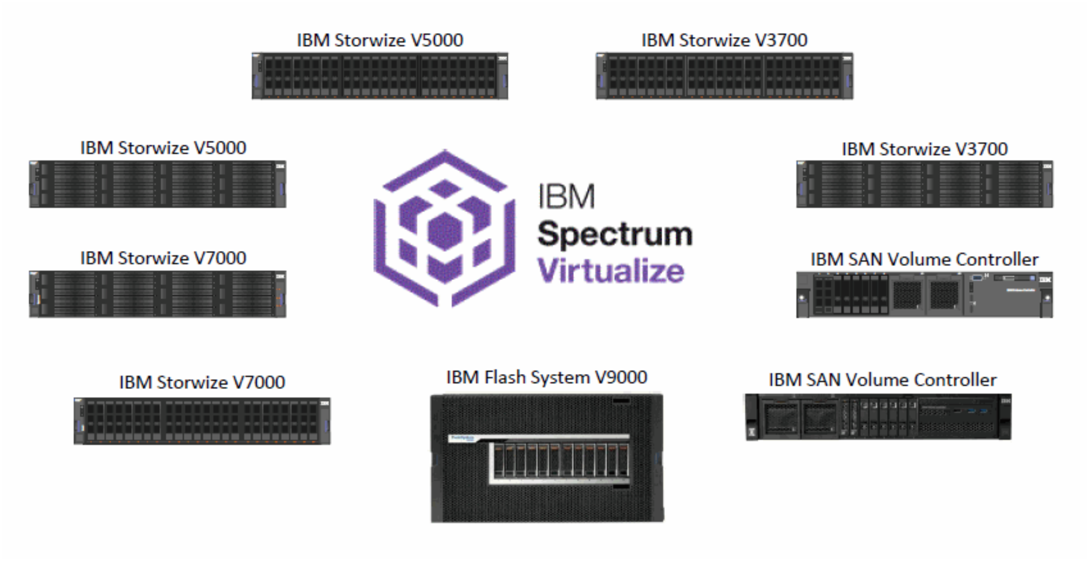


*Figure 2-1 IBM Spectrum Virtualize software*

## 2.2 Storage virtualization

Storage virtualization, like server virtualization, is one of the foundations of building a flexible and reliable infrastructure solution that enables companies to better align their IT needs. Storage virtualization enables an organization, in sense of affordability and manageability, to implement storage pools across several physically separate disk systems (which might be from different vendors).

Storage can then be deployed from these pools, and can be migrated transparently between pools without interruption to the attached host systems and their applications. Storage virtualization provides a single set of tools for advanced functions, such as instant copy and remote mirroring solutions, which enables faster and seamless deployment of storage regardless of the underlying hardware.

Because the storage virtualization represented by IBM Spectrum Virtualize is a software-enabled functionality, it offers additional features that are typically not available on a regular pure storage subsystem. These include but are not limited to the following features:

- Data compression
- Software and hardware encryption
- IBM Easy Tier®
- Thin provisioning
- Mirroring and copy services

- Interface to Cloud Service Providers

Figure 2-2 shows these features at a glance.

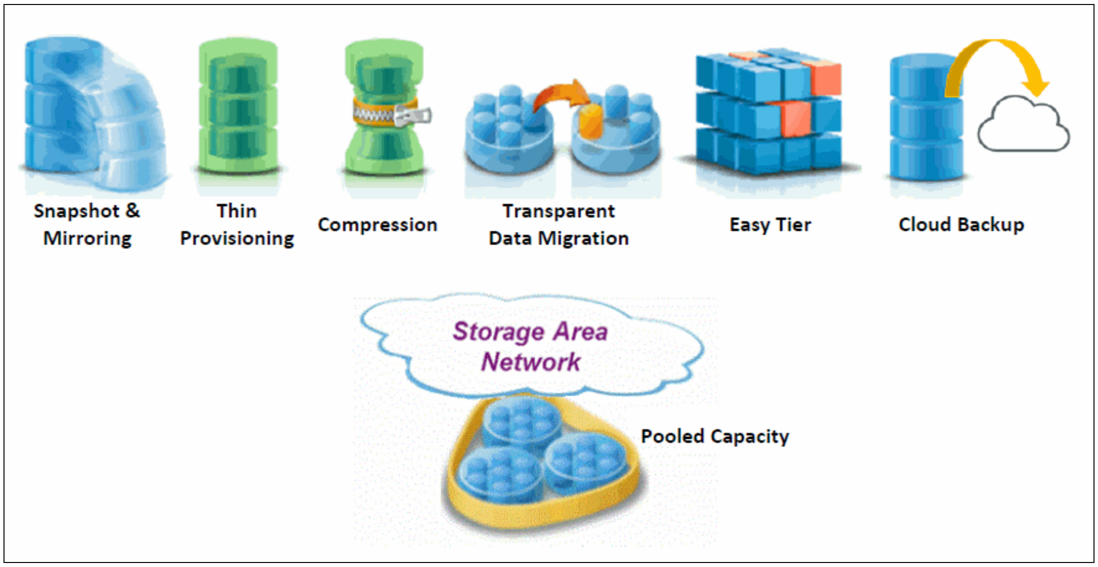


*Figure 2-2 IBM Spectrum Storage virtualization*

There are five top reasons to choose software-defined storage virtualization infrastructure and benefit from IBM Spectrum Virtualize:

- Lower license cost. Avoid purchasing licenses from multiple storage vendors for advanced features (replication, tiering, snapshots, compression, and so on) of each external storage subsystem. Manage them centrally from IBM Spectrum Virtualize.
- Feed more data on existing physical disks. External storage area network (SAN) disk arrays have physical boundaries. Although one subsystem might run out of space, another has free capacity. virtualization removes these boundaries.
- Choose lower-cost disk arrays. Less-demanding applications can easily run on cheaper disks with lower performance. IBM Spectrum Virtualize automatically and transparently moves data up and down between low-performance and high-performance disk arrays (tiering).
- Could FlashCopy. IBM Spectrum Virtualize has interface to external cloud service providers. IBM Spectrum Cloud FlashCopy supports full and incremental backup and restore from cloud snapshots.
- Quick adoption of new technologies. IBM Spectrum Virtualize seamlessly integrates invention in storage technologies, such as new array types, new disk vendors, and so on.
- Extended high availability (HA). Cross-site virtualization, workload migration, and copy services enhance options for deployment of high availability scenarios or disaster recovery (DR) solutions (IBM SAN Volume Controller Enhanced Stretched Cluster, IBM HyperSwap®).

## 2.3 IBM Storwize V7000 overview

IBM Storwize V7000 solution incorporates IBM Spectrum Virtualize software and provides a modular storage system that includes the capability to virtualize its internal and external SAN-attached storage. IBM Storwize V7000 solution is built upon IBM Spectrum Virualize.

IBM Storwize V7000 system provides several configuration options that are aimed at simplifying the implementation process. These configuration options conform to different implementation scenarios regarding the size of your data center and SAN and local area network (LAN) topology. IBM Storwize V7000 system is a clustered, scalable, midrange storage system, easy to deploy and easy to use.

Figure 2-3 shows a high-level overview of IBM Storwize V7000.

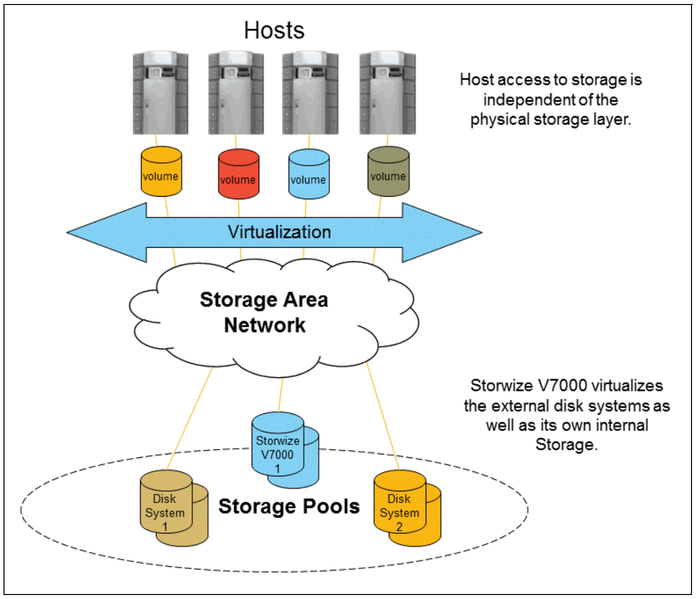

Hosts
Host access to storage is independent of the physical storage layer.
volume
volume
volume
volume
Virtualization
Storage Area Network
Storwize V7000 virtualizes the external disk systems as well as its own internal Storage.
Storwize V7000 1
Disk System 1
Storage Pools
Disk System 2

*Figure 2-3 IBM Storwize V7000 overview*

The IBM Spectrum Virtualize software that runs on IBM Storwize V7000 provides a graphical user interface (GUI) that enables storage to be deployed quickly and efficiently. The GUI is provisioned by IBM Spectrum Virtualize code and there is no need for a separate console.

The management GUI contains a series of preestablished configuration options that are called *presets*, and that use common settings to quickly configure objects on the system. Presets are available for creating volumes and IBM FlashCopy® mappings, and for setting up a Redundant Array of Independent Disks (RAID) configuration, including traditional RAIDs and the new feature of distributed RAID.

An IBM Storwize V7000 solution provides a choice of up to 480 x 3.5 inch or 1056 x 2.5 inch serial-attached SCSI (SAS) drives for the internal storage in a clustered system. The solution uses SAS cables and connectors to attach to the optional expansion enclosures. In a clustered system, the IBM Storwize V7000 can provide up to 8 pebibytes (PiB) of raw capacity (with 8-terabyte (TB) nearline SAS disks).

When virtualizing external storage arrays, the IBM Storwize V7000 system can provide up to 32 PiB of usable capacity. An IBM Storwize V7000 system supports a range of external disk systems similar to what IBM SAN Volume Controller supports today. See Figure 2-4 for a view of an IBM Storwize V7000 control enclosure.

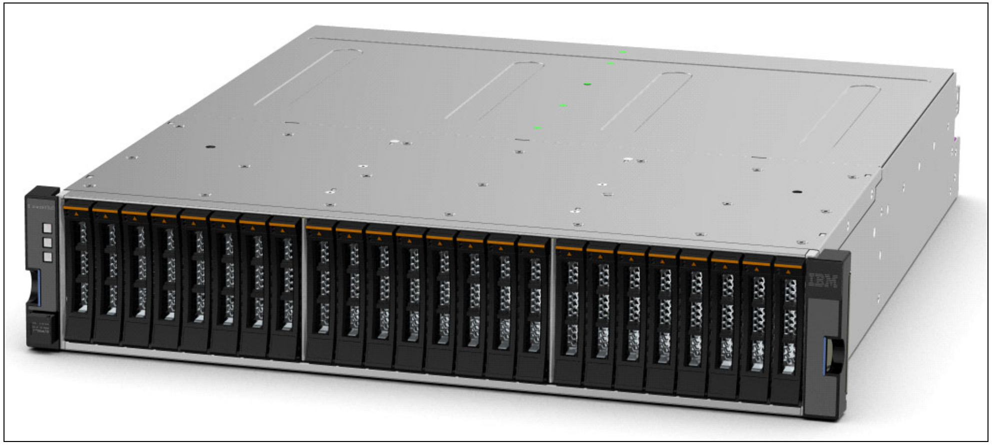


*Figure 2-4 Top-front view of a Storwize V7000 control enclosure*

The IBM Storwize V7000 solution consists of 1 - 4 control enclosures and optionally, up to 36 expansion enclosures (and supports the intermixing of the different expansion enclosures). Within each enclosure are two canisters. Control enclosures contain two node canisters, and expansion enclosures contain two expansion canisters.

## 2.3.1 IBM Storwize V7000 models

The IBM Storwize V7000 consists of enclosures and drives. An enclosure contains two canisters that are seen as part of the enclosure, although they can be replaced independently.

**Additional information:** For the most up-to-date information about features, benefits, and specifications of IBM Storwize V7000 models, see the following address:

https://ibm.biz/BdscVY

The information in this book is valid at the time of writing and covers IBM Spectrum Virtualize V7.8, but as IBM Storwize V7000 matures, expect to see new features and enhanced specifications.

The IBM Storwize V7000 models are described in Table 2-1.

*Table 2-1 IBM Storwize V7000 models*

| Model | Cache | Fibre Channel (FC) / iSCSI / SAS ports | Drive slots | Power supply |
|---|---|---|---|---|
| 2076-AF1 (with two node canisters Gen2+) | 64 or 128 gigabytes (GB) | 16 x 16 gigabit (Gb) / 6 x 1 Gb + 8x 10 Gb / 4 x 12 Gb | 24 x 2.5 inch (All Flash) | Integrated dual power supplies with battery |
| 2076-624 (with two node canisters Gen2+) | 64 or 128 gigabytes (GB) | 16 x 16 gigabit (Gb) / 6 x 1 Gb + 8x 10 Gb / 4 x 12 Gb | 24 x 2.5 inch | Integrated dual power supplies with battery |
| 2076-524 (with two node canisters Gen2) | 32 or 64 gigabytes (GB) | 4 x 16 gigabit (Gb) / 4 x 1 Gb + 4 x 10 Gb / 4 x 12 Gb | 24 x 2.5 inch | Integrated dual power supplies with battery |

| Model | Cache | Fibre Channel (FC) / iSCSI / SAS ports | Drive slots | Power supply |
|---|---|---|---|---|
| 2076-212 (with two expansion canisters) | Not applicable (N/A) | -- / -- / 4 x 12 Gb | 12 x 3.5 inch | Integrated dual power supplies |
| 2076-224 (with two expansion canisters) | N/A | -- / -- / 4 x 12 Gb | 24 x 2.5 inch | Integrated dual power supplies |
| 2076-12F (with two expansion canisters Gen2) | N/A | -- / -- / 4 x 12 Gb | 12 x 3.5 inch | Integrated dual power supplies (attaches to 2076-524 and 2076-624 only) |
| 2076-24F (with two expansion canisters Gen2) | N/A | -- / -- / 4 x 12 Gb | 24 x 2.5 inch | Integrated dual power supplies (attaches to 2076-524 and 2076-624 only) |

**Note:** The first generation of control enclosures (2076 - models 112, 124, 312, and 324) has been withdrawn from marketing. However, expansion enclosures 2076-212 and 2076-224 can still be ordered (see Table 2-1 on page 16) as they only attach to those control enclosures. Intermix of control enclosures with expansion enclosures of different generations is not a supported combination, and is refused by IBM Spectrum Virtualize software.

## 2.3.2 IBM Storwize V7000 functions

The following functions are available with the current release of IBM Spectrum Virtualize:

- Thin provisioning

  Traditional fully allocated volumes allocate real physical disk capacity for an entire volume even if that capacity is never used. Thin-provisioned volumes allocate real physical disk capacity only when data is written to the logical volume.

- Volume mirroring

  Provides a single volume image to the attached host systems while maintaining pointers to two copies of data in separate storage pools. Copies can be on separate disk storage systems that are being virtualized. If one copy is failing, IBM Storwize V7000 provides continuous data access by redirecting input/output (I/O) to the remaining copy. When the copy becomes available, automatic resynchronization occurs.

- FlashCopy

  Provides a volume level point-in-time copy function for any storage being virtualized by IBM Spectrum Virtualize. This function creates copies for backup, parallel processing, testing, and development, and has the copies available almost immediately.

  IBM Storwize V7000 includes the following IBM FlashCopy functions enabled by IBM Spectrum Virtualize V7.8:

  – Full or incremental copy

  This function copies only the changes from either the source or target data since the last FlashCopy operation, and enables completion of point-in-time online backups much more quickly than using traditional FlashCopy.

  – Multitarget FlashCopy

    IBM Storwize V7000 supports copying of up to 256 target volumes from a single source volume. Each copy is managed by a unique mapping and in general, each mapping acts independently and is not affected by other mappings sharing the source volume.

  – Cascaded FlashCopy

    This function is used to create copies of copies, and supports full, incremental, or nocopy operations.

  – Reverse FlashCopy

    This function enables data from an earlier point-in-time copy to be restored with minimal disruption to the host.

  – FlashCopy nocopy with thin provisioning

    This function provides a combination of using thin-provisioned volumes and FlashCopy together to help reduce disk space requirements when making copies.

    This option has two variations:

    - Space-efficient source and target with background copy

      Copies only the allocated space.

    - Space-efficient target with no background copy

      Copies only the space used for changes between the source and target and is generally referred to as a *snapshot*.

    This function can be used with multitarget, cascaded, and incremental FlashCopy.

  – Consistency groups

    Consistency groups address the issue where application data is on multiple volumes. By placing the FlashCopy relationships into a consistency group, commands can be issued against all of the volumes in the group. This action provides a consistent point-in-time copy of all of the data, even if it is on a physically separate volume.

    FlashCopy mappings can be members of a consistency group, or they can be operated in a stand-alone manner, not as part of a consistency group. FlashCopy commands can be issued to a FlashCopy consistency group, which affects all FlashCopy mappings in the consistency group, or to a single FlashCopy mapping if it is not part of a defined FlashCopy consistency group.

- Metro Mirror

  Provides a *synchronous* remote mirroring function up to approximately 300 kilometers (km) between sites. Because the host I/O only completes after the data is cached at both locations, performance requirements might limit the practical distance. Metro Mirror provides fully synchronized copies at both sites with zero data loss after the initial copy is completed. Metro Mirror can operate between multiple IBM Storwize systems and IBM SAN Volume Controllers.

- Global Mirror

  Provides a long-distance *asynchronous* remote mirroring function up to approximately 8,000 km between sites. With Global Mirror, the host I/O completes locally and the changed data is sent to the remote site later. This maintains a consistent recoverable copy of data at the remote site, which lags behind the local site. Global Mirror can operate between multiple IBM Storwize systems and IBM SAN Volume Controllers.

- External virtualization

  IBM Storwize V7000 provides a data migration function that can be used to import external storage systems into the IBM Storwize V7000 system. The following tasks can be accomplished:

  – Move volumes nondisruptively onto a newly installed storage system
  – Move volumes to rebalance a changed workload
  – Migrate data from other back-end storage to IBM Storwize V7000-managed storage

- Software Encryption

  IBM Storwize V7000 Gen2 and IBM Storwize Gen2+ provide optional encryption of data-at-rest functionality, which protects against the potential exposure of sensitive user data and user metadata that is stored on discarded, lost, or stolen storage devices. Encryption can only be enabled and configured on enclosures that support encryption.

  Starting with IBM Spectrum Virtualize V7.6, IBM Spectrum Virtualize offers software-enabled encryption, which includes encryption of internal or external storage.

- IBM Easy Tier

  Provides a mechanism to seamlessly migrate hot spots to the most appropriate tier within the IBM Storwize V7000 solution. This migration could be to internal drives within the IBM Storwize V7000, or to external storage systems that are virtualized by IBM Storwize V7000. Independently on Easy Tier, IBM Storwize V7000 provides an automatic storage pool balancing that is enabled by default and requires no license.

- IBM Real-time Compression

  IBM Real-time Compression technology is based on the Random Access Compression Engine (RACE). RACE is an integral part of the software stack of IBM Spectrum Virtualize V6.4 and later. This integration does not alter the behavior of the system, so previously existing features are supported for compressed volumes. Starting with IBM Spectrum Virtualize V7.6, two software-enabled RACE engines sharing HW resources for compression are required.

## 2.3.3 IBM Storwize V7000 licensing

With the broad range of technical features and capabilities of the IBM Storwize V7000 Gen2 and Gen2+, including Copy Services, External Virtualization, Easy Tier, and Real-time Compression, IBM simplified the licensing model to include these new features. The IBM Storwize V7000 offers two ways of license procurement:

- Fully flexible
- Bundled (license packages)

The license model is based on a *license-per-enclosure* concept familiar from the first generation of IBM Storwize V7000. However, the second generation offers more flexibility to match your specific needs.

> **Upgrade:** Installing or upgrading the code on the *first* generation of IBM Storwize V7000 to the V7.8 does not change your existing license model or license needs.

The conceptual model of the licensing in IBM Storwize V7000 Gen2 and Gen2+ is depicted in Figure 2-5.

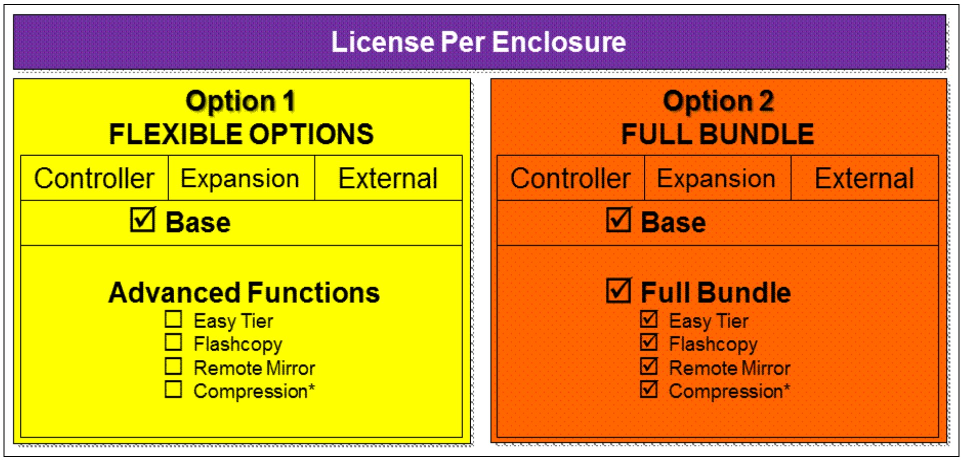


*Figure 2-5 Licensing model on IBM Storwize V7000 Gen2 and Gen2+*

The base module is represented by *IBM Spectrum Virtualize family* and is mandatory for every controller, enclosure, or externally managed controller unit. Additional licensed features can be purchased on-demand, either as a full software bundle or each feature separately. Any additional licenses need to be procured per every existing enclosure where they are planned to be used.

## 2.4 IBM Storwize V7000 hardware

In conjunction with the previous release of the IBM Spectrum Virtualize V7.3, IBM introduced a hardware refresh for the IBM Storwize V7000 platform. These improvements are further enhanced with the current code of IBM Spectrum Virtualize software. In this section, we introduce and partially repeat these hardware changes and software improvements associated with the most recent model. These include the following changes and improvements:

- New internal component layout, such as canister and ports
- Integrated battery pack within node canisters
- Enhanced scalability and flexibility with 16 gigabits per second (Gbps) I/O adapters
- Improved Real-time Compression engine with hardware assistance
- Extended disk drive support

To meet these objectives, the base hardware configuration of the IBM Storwize V7000 Gen2 and Gen2+ were substantially improved to support more advanced processors, more memory, and faster interconnects.

## 2.5 IBM Storwize V7000 components

The IBM Storwize V7000 is a modular, midrange virtualization RAID storage subsystem that employs the IBM Spectrum Virtualize software engine. It has the following benefits:

- Brings enterprise technology to midrange storage.
- Specialty administrators are not required.
- Client setup and service are simplified.
- The system can grow incrementally as storage capacity and performance needs change.
- Multiple storage tiers are in a single system with nondisruptive migration between them.
- Simple integration can be done into the server environment.

The IBM Storwize V7000 consists of a set of drive enclosures. Control enclosures contain disk drives and two nodes (an I/O group), which are attached to the SAN fabric or 10 gigabit Ethernet (GbE) fast Ethernet. Expansion enclosures contain drives and are attached to control enclosures.

The simplest use of the IBM Storwize V7000 is as a traditional RAID subsystem. The internal drives are configured into RAID arrays, and virtual disks are created from those arrays. With IBM Spectrum Virtualize software, this usage is extended by the deployment of distributed RAID arrays, which shrink the reconstruction time of failed drives.

IBM Storwize V7000 supports spinning disks and flash drives. When different tiers are employed, the IBM Storwize V7000 uses IBM Easy Tier to automatically place volume hot spots on better-performing storage. Even without the Easy Tier enabled, the Storage Pool balancing is available and enabled by default to equally distribute workload equally across all MDisks in that tier.

## 2.5.1 Hosts

A host system is a server that is connected to IBM Storwize V7000 through a Fibre Channel connection, Fibre Channel over Ethernet (FCoE), or through an Internet SCSI (iSCSI) connection.

Hosts are defined to IBM Spectrum Virtualize by identifying their worldwide port names (WWPNs) for Fibre Channel hosts. For iSCSI hosts, they are identified by using their iSCSI names. The iSCSI names can either be iSCSI qualified names (IQNs) or extended unique identifiers (EUIs).

## 2.5.2 Host Cluster

Host cluster is an host object in the IBM Storwize V7000. A host cluster is a combination of two or more servers that is connected to IBM Storwize V7000 through a Fibre Channel, Fibre Channel over Ethernet (FCoE), or through an Internet SCSI (iSCSI) connection. Host cluster object can see the same set of volumes.

## 2.5.3 Nodes

An IBM Storwize V7000 can have 2 - 8 hardware components, called *nodes* or *node canisters*, which provide the virtualization of internal and external volumes, and the cache and copy services (remote copy) functions. A clustered system consists of 1 - 4 node pairs.

One of the nodes within the system is known as the *configuration node*. It is the node that manages configuration activity for the clustered system. If this node fails, the system nominates automatically another node to become the configuration node.

## 2.5.4 I/O groups

Within IBM Storwize V7000, there are 1 - 4 pairs of node canisters known as *I/O groups*. The IBM Storwize V7000 with installed IBM Spectrum Virtualize supports eight node canisters in the clustered system, which provides four I/O groups. When a host server performs I/O to one of its volumes, all the I/Os for a specific volume are directed to the I/O group. Also, under normal conditions, the I/Os for that specific volume are always processed by the same node within the I/O group.

One node of the I/O group acts as a preferred node for their own specific subset of the total number of volumes that the I/O group presents to the host servers (a maximum of 2048 volumes per I/O group). However, each node also acts as a failover node for its partner node within the I/O group, so a node takes over the I/O workload from its partner node, if required, with no effect to the server's application.

In an IBM Storwize V7000 environment, using active/active architecture, the I/O handling for a volume can be managed by both nodes of the I/O group. Therefore, it is mandatory for servers that are connected through Fibre Channel connectors to use multipath device drivers to be able to handle this capability.

The IBM Storwize V7000 I/O groups are connected to the SAN so that all application servers accessing volumes from the I/O group have access to them. Up to 2048 host server objects can be defined in four I/O groups.

**Important:** The active/active architecture provides availability to process I/Os for both controller nodes. It enables the application to continue running smoothly, even if the server has only one access route or path to the storage controller. This type of architecture eliminates the path and logical unit number (LUN) thrashing typical of an active/passive architecture.

### 2.5.5 Cache

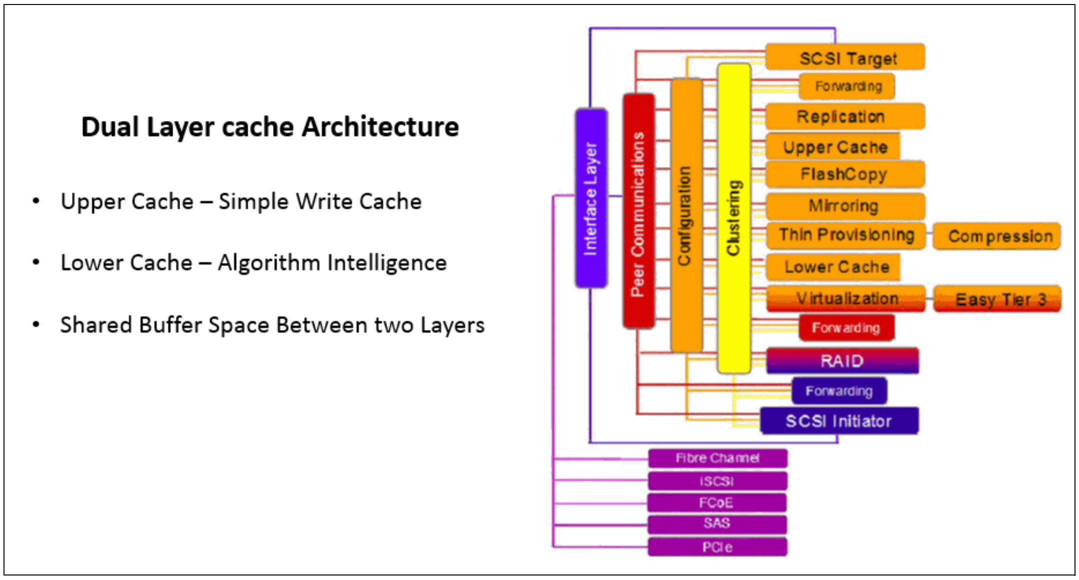
The primary benefit of storage cache is to improve I/O response time. Reads and writes to a magnetic disk drive experience seek time and latency time at the drive level, which can result in 1 ms - 10 ms of response time (for an enterprise-class disk).

Cache is allocated in 4 KiB segments. A *segment* holds part of one track. A *track* is the unit of locking and destaging granularity in the cache. The cache virtual track size is 32 KiB (eight segments). A track might be only partially populated with valid pages. The IBM Storwize V7000 combines writes up to a 256 KiB track size if the writes are in the same tracks before destage. For example, if 4 KiB is written into a track, another 4 KiB is written to another location in the same track.

Therefore, the blocks that are written from the IBM Storwize V7000 to the disk subsystem can be any size between 512 bytes up to 256 KiB. The large cache and advanced cache management algorithms allow it to improve on the performance of many types of underlying disk technologies. The system's capability to manage, in the background, the destaging operations that are incurred by writes (in addition to still supporting full data integrity) assists with system's capability in achieving good database performance.

The cache is separated into two layers: an upper cache and a lower cache.

Figure 2-6 on page 23 shows the separation of the upper and lower cache.



Dual Layer cache Architecture

- Upper Cache – Simple Write Cache
- Lower Cache – Algorithm Intelligence
- Shared Buffer Space Between two Layers

*Figure 2-6 Separation of upper and lower cache*

The upper cache delivers the following functionality, which enables the system to streamline data write performance:

- Provides fast write response times to the host by being as high up in the I/O stack as possible
- Provides partitioning

The lower cache delivers the following additional functionality:

- Ensures that the write cache between two nodes is in sync
- Caches partitioning to ensure that a slow back end cannot use the entire cache
- Uses a destage algorithm that adapts to the amount of data and the back-end performance
- Provides read caching and prefetching

Combined, the two levels of cache also deliver the following functionality:

- Pins data when the LUN goes offline
- Provides enhanced statistics for IBM Tivoli® Storage Productivity Center, and maintains compatibility with an earlier version
- Provides trace for debugging
- Reports medium errors
- Resynchronizes cache correctly and provides the atomic write functionality
- Ensures that other partitions continue operation when one partition becomes 100% full of pinned data
- Supports fast-write (two-way and one-way), flush-through, and write-through
- Integrates with T3 recovery procedures
- Supports two-way operation
- Supports none, read-only, and read/write as user-exposed caching policies
- Supports flush-when-idle
- Supports expanding cache as more memory becomes available to the platform
- Supports credit throttling to avoid I/O skew and offer fairness/balanced I/O between the two nodes of the I/O Group
- Enables switching of the preferred node without needing to move volumes between I/O Groups

Depending on the size, age, and technology level of the disk storage system, the total available cache in the system can be larger, smaller, or about the same as the cache that is associated with the disk storage.

Because hits to the cache can occur in either the IBM Storwize V7000 or the disk controller level of the overall system, the system as a whole can take advantage of the larger amount of cache wherever the cache is located. Therefore, if the storage controller level of the cache has the greater capacity, expect hits to this cache to occur, in addition to hits in the IBM Storwize V7000's cache.

Also, regardless of their relative capacities, both levels of cache tend to play an important role in enabling sequentially organized data to flow smoothly through the system. The IBM Storwize V7000 cannot increase the throughput of the underlying disks in all cases, because this increase depends on both the underlying storage technology and the degree to which the workload exhibits *hotspots* or sensitivity to cache size or cache algorithms.

### 2.5.6 Clustered system

A clustered system consists of 1 - 4 pairs of nodes. All configuration, monitoring, and service tasks are performed at the system level, and the configuration settings are replicated across all node canisters in the clustered system. To facilitate these tasks, 1 or 2 management IP addresses are set for the system.

A process is provided to back up the system configuration data to disk so that the clustered system can be restored if there is a disaster. This method does not back up application data, only the IBM Storwize V7000 system configuration information.

**System configuration backup:** After backing up the system configuration, save the backup data outside of the SAN. If you are unable to access the IBM Storwize V7000, you do not have access to the backup data if it is on the SAN.

For the purposes of remote data mirroring, two or more clustered systems (IBM Spectrum Virtualize systems starting from software V7.1) must form a partnership before creating relationships between mirrored volumes.

**Important:** IBM Storwize V7000 V6.3 introduced the **`layer`** parameter. It can be changed by running **`chsystem`** using only the command-line interface (CLI). The default is the `storage` layer, and you must change it to `replication` if you need to set up a copy services relationship between the IBM Storwize family and IBM SAN Volume Controller.

One node is designated as the configuration node canister. It is the only node that activates the system IP address. If the configuration node canister fails, the system chooses a new configuration node and the new configuration node takes over the system IP addresses.

The system can be configured using either the IBM Spectrum Virtualize GUI-based management software, the CLI, or through an application, such as IBM Spectrum Control, that uses the IBM Storwize V7000 Common Information Model object manager (CIMOM).

### 2.5.7 HyperSwap

IBM HyperSwap function is an HA feature that provides dual-site, active-active access to a volume. Active-active volumes have a copy in one site and a copy at another site. Data that is written to the volume is automatically sent to both copies so that either site can provide access to the volume if the other site becomes unavailable. Active-active relationships are made between the copies at each site. These relationships automatically run and switch direction according to which copy or copies are online and up to date.

Relationships can be grouped into consistency groups just like Metro Mirror and Global Mirror relationships. The consistency groups fail over consistently as a group based on the state of all copies in the group. An image that can be used if there is a disaster recovery (DR) is maintained at each site. When the system topology is set to HyperSwap, each IBM Storwize V7000 controller, and host map object in the system configuration must have a site attribute set to `1` or `2`.

This site must be the same site as the site of the controllers that provide the managed disks to that I/O group. When managed disks are added to storage pools, their site attributes must match. This requirement ensures that each copy in the active-active relationship is fully independent and is at a distinct site.

## 2.5.8 Dense Expansion Drawers

Dense Expansion Drawers, or just Dense Drawers, are optional disk expansion enclosure introduced as 5U rack-mounted. Each chassis features two expansion canisters, two power supplies, two expander modules and a total of 4 fan modules.

Each Dense Drawer can hold up 92 drives that are positioned in four rows of 14 and additional 3 rows of 12 mounted drives assemblies. There are two Secondary Expander Modules (SEM) that are centrally located in the chassis. One SEM addresses 54 drive ports the other addresses 38 drive ports.

The drive slots are numbered 1 - 14, starting from the left rear slot and working from left to right, back to front.

Each canister in the Dense Drawer chassis features two SAS ports numbered 1 and 2. The use of SAS port1 is mandatory, because the expansion enclosure must be attached to an IBM Storwize V7000 node or another expansion enclosure. SAS connector 2 is optional, because it is used to attach to more expansion enclosures.

Each IBM Storwize V7000 can support up to four Dense Drawers per SAS chain.

Figure 2-7 shows a Dense Expansion Drawer.

Figure 2-7 IBM Dense Expansion Drawer

## 2.5.9 Expansion Enclosures

There are two types of available Expansion Enclosures, IBM Storwize V7000 large form factor (LFF) Expansion Enclosure Model 12F and small form factor (SFF) 24F.

- IBM Storwize V7000 Gen2 LFF 12F includes the following components:
- Two expansion canisters
- 12 Gb SAS ports for control enclosure and Expansion Enclosure attachment
- Twelve slots for 3.5-inch SAS drives
- 2U, 19-inch rack mount enclosure with ac power supplies

IBM Storwize V7000 SFF Expansion Enclosure Model 24F includes the following components:

- Two expansion canisters
- 12 Gb SAS ports for control enclosure and expansion enclosure attachment
- Twenty-four slots for 2.5-inch SAS drives
- 2U, 19-inch rack mount enclosure with AC power supplies

The Expansion Enclosure is a 2U enclosure, containing the following components:

- 24 2.5 in. drives (HDDs or SSDs).
- 2 Storage Bridge Bay (SBB)-compliant enclosure services manager (ESM) canisters.
- 2 fan assemblies. These mount between the drive midplane and the Node Canisters. Each
- fan module is removable when the Node Canister is removed.
- 2 Power supplies.
- RS232 port on the back panel (3.5 mm stereo jack). This is used for configuration during manufacturing.

The front of an Expansion Enclosure is shown in Figure 2-8.

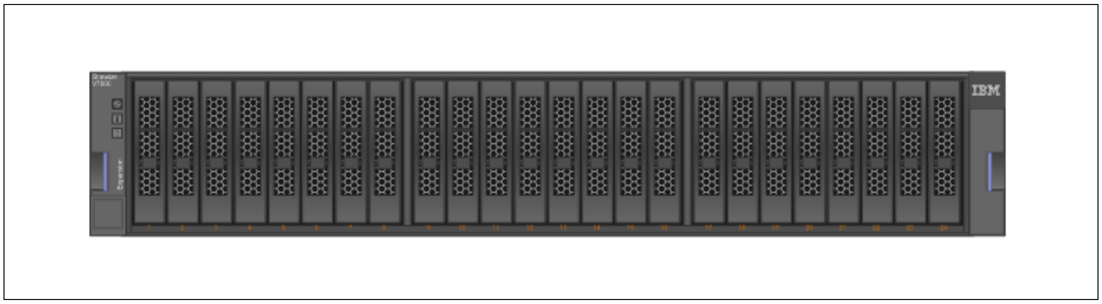


Figure 2-8   Front of IBM Storwize V7000 Expansion Enclosure

Figure 2-9 on page 27 shows a rear view of an expansion enclosure.

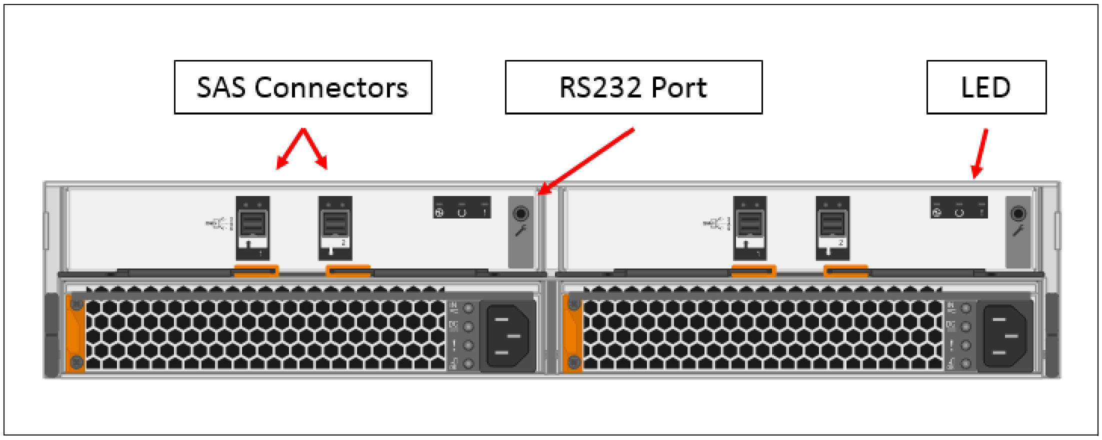


Figure 2-9   Rear of IBM Storwize V7000 expansion enclosure

## 2.5.10  RAID

The IBM Storwize V7000 setup contains several internal drive objects, but these drives cannot be directly added to the storage pools. Drives need to be included in a Redundant Array of Independent Disks (RAID) to provide protection against the failure of individual drives.

These drives are referred to as members of the array. Each array has a RAID level. RAID levels provide various degrees of redundancy and performance, and have various restrictions regarding the number of members in the array.

Apart from traditional disk arrays IBM Spectrum Virtualize V7.6 introduced distributed RAID. Distributed RAID improves recovery time of failed disk drives in an array by distributing spare capacity between primary disks, rather than dedicating a whole spare drive for replacement.

IBM Storwize V7000 supports hot spare drives. When an array member drive fails, the system automatically replaces the failed member with a hot spare drive and rebuilds the array to restore its redundancy. Candidate and spare drives can be manually exchanged with array members.

Each array has a set of goals that describe the location and performance of each array. A sequence of drive failures and hot spare takeovers can leave an array unbalanced, with members that do not match these goals. The system automatically rebalances such arrays when the appropriate drives are available.

## 2.5.11 Read Intensive Flash Drives

Generally, there are two types of SSDs for Enterprise Storage, the Multi-level cell (MLC) and Single-level cell (SLC).

The most common SSD technology is MLC. They are found in consumer products such as portable eletronic devices. However, they are also strongly present in some enterprise storage products. Enterprise class SSDs are built on mid to high-endurance multi-level cell flash technology, known as mainstream endurance SSD.

MLC SSDs uses multi cell to store data and feature the Wear Levelling method, which is the process to evenly spread the data across all memory cells on the SSD. This method helps to eliminate potential hotspots caused by repetitive Write-Erase cycles. SLC SSDs use a single cell to store one bit of data, and that makes them generally faster.

To support particular business demands, IBM Spectrum Virtualize has qualified the use of RI SSDs with applications where the read operations are signifcantly high.

Read Intensive (RI) SSDs are available as an optional purchase product to IBM SAN Volume Controller and IBM Storwize Family.

For more information about Read Intensive SSDs, the certified models and IBM Spectrum Virtualize code level to support RI SSDs, see this website:

`http://www-03.ibm.com/systems/support/storage/ssic/interoperability.wss`

## 2.5.12 Managed disks

A managed disk (MDisk) is the unit of storage that IBM Spectrum Virtualize virtualizes. This unit could be a logical volume on an external storage array presented to the IBM Storwize V7000, or a RAID consisting of internal drives. IBM Spectrum Virtualize can then allocate these MDisks into various storage pools. An MDisk is not visible to a host system on the storage area network, because it is internal or zoned only to the IBM Storwize V7000 system.

An MDisk has four modes:

- Array

  Array mode MDisks are constructed from drives using the RAID function. Array MDisks are always associated with storage pools.

- Unmanaged

  Unmanaged MDisks are not being used by the system. This situation might occur when an MDisk is first imported into the system, for example.

- Managed

  Managed MDisks are assigned to a storage pool and provide extents so that volumes can use it.

- Image

  Image MDisks are assigned directly to a volume with a one-to-one mapping of extents between the MDisk and the volume. This situation is normally used when importing logical volumes into the clustered system that already have data on them, which ensures that the data is preserved as it is imported into the clustered system.

## 2.5.13 Quorum disks

A quorum disk is an MDisk that contains a reserved area for use exclusively by the system. In IBM Storwize V7000, internal drives can be considered as quorum candidates. The clustered system uses quorum disks to break a tie when exactly half the nodes in the system remain after a SAN failure.

The clustered system automatically forms the quorum disk by taking a small amount of space from an MDisk. It allocates space from up to three different MDisks for redundancy, although only one quorum disk is active.

If the environment has multiple storage systems, to avoid the possibility of losing all of the quorum disks because of a failure of a single storage system, you should allocate the quorum disk on different storage systems. It is possible to manage the quorum disks by using the CLI.

## 2.5.14 IP Quorum

In a HyperSwap configuration, there must be a third, independent site to house quorum devices. To use a quorum disk as the quorum device, this third site must use Fibre Channel connectivity together with an external storage system. Sometimes, Fibre Channel connectivity is not possible. In a local environment, no extra hardware or networking, such as Fibre Channel or SAS-attached storage, is required beyond what is normally always provisioned within a system.

To use an IP-based quorum application as the quorum device for the third site, no Fibre Channel connectivity is used. Java applications are run on hosts at the third site. However, there are strict requirements on the IP network, and some disadvantages with using IP quorum applications.

Unlike quorum disks, all IP quorum applications must be reconfigured and redeployed to hosts when certain aspects of the system configuration change. These aspects include adding or removing a node from the system, or when node service IP addresses are changed.

For stable quorum resolutions, an IP network must provide the following requirements:

- Connectivity from the hosts to the service IP addresses of all nodes. If IP quorum is configured incorrectly, the network must also deal with possible security implications of exposing the service IP addresses, because this connectivity can also be used to access the service GUI.
- Port 1260 is used by IP quorum applications to communicate from the hosts to all nodes.
- The maximum round-trip delay must not exceed 80 ms, which means 40 ms each direction.
- A minimum bandwidth of 2 MBps is ensured for node-to-quorum traffic.

Even with IP quorum applications at the third site, quorum disks at site one and site two are required, because they are used to store metadata. To provide quorum resolution, use the **mkquorumapp** command to generate a Java application that is copied from the system and run on a host at a third site. The maximum number of applications that can be deployed is five. Currently, supported Java runtime environments (JREs) are IBM Java 7.1 and IBM Java 8.

### 2.5.15 Storage pool

A storage pool is a collection of MDisks that are grouped to provide capacity for volumes. All MDisks in the pool are split into extents with the same size. Volumes are then allocated out of the storage pool and are mapped to a host system.

**IBM Storwize V7000 object names:** The names must begin with a letter, which cannot be numeric. The name can be a maximum of 63 characters. Valid characters are uppercase letters (A-Z), lowercase letters (a-z), digits (0 - 9), underscore (_), period (.), hyphen (-), and space. The names must not begin or end with a space.

MDisks can be added to a storage pool at any time to increase the capacity of the storage pool. MDisks can belong in only one storage pool, and only MDisks in unmanaged mode can be added to the storage pool. When an MDisk is added to the storage pool, the mode changes from unmanaged to managed (and vice versa when you remove it).

Each MDisk in the storage pool is divided into several extents. The size of the extent is selected by the administrator at creation time of the storage pool, and cannot be changed later. The size of the extent ranges from 16 megabytes (MB) - 8 GB. The default extent size in V7.8 is 1024 MB.

The extent size has a direct effect on the maximum volume size and storage capacity of the clustered system. A system can manage 4 million (4 x 1024 x 1024) extents. For example, a system with a 16 MB extent size can manage up to 16 MB x 4 MB = 64 terabytes (TB) of storage.

Use the same extent size for all storage pools in a clustered system, which is a prerequisite for supporting volume migration between two storage pools. If the storage pool extent sizes are not the same, you must use volume mirroring to copy volumes between storage pools.

**Default extent size:** The IBM Storwize V7000 GUI has a default extent size value of 1024 MB when you define a new storage pool, and can only be changed using the CLI.

Figure 2-10 shows the storage pool components.

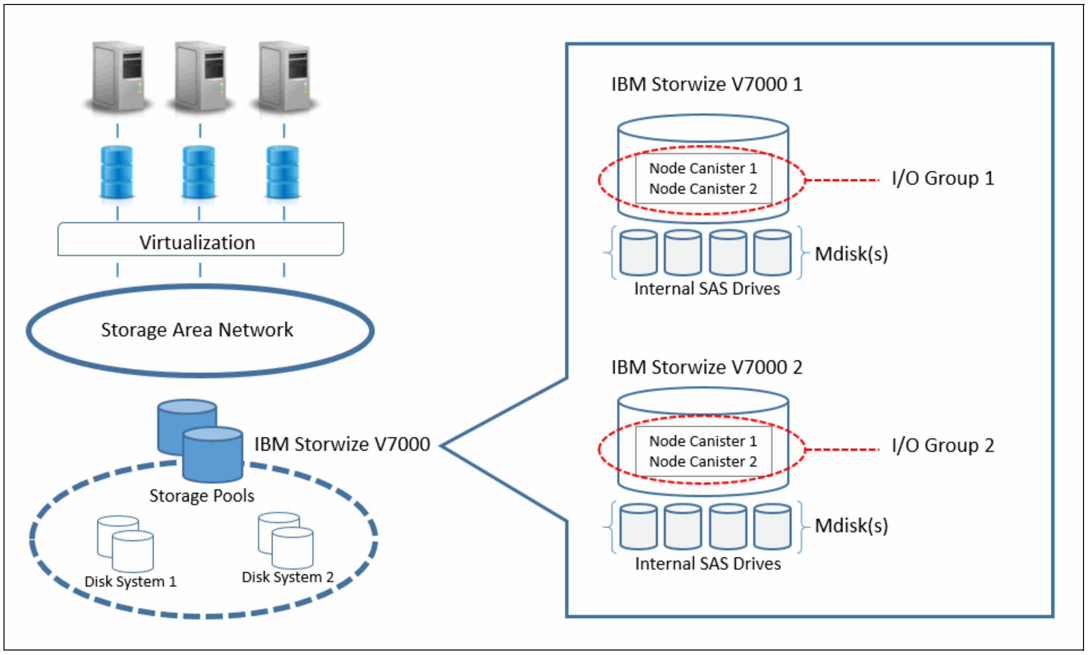


*Figure 2-10 IBM Spectrum Virtualize virtualization components*

## 2.5.16 Volumes

A volume is a logical disk that is presented to a host system by the clustered system. In our virtualized environment, the host system has a volume mapped to it by IBM Storwize V7000. IBM Spectrum Virtualize translates this volume into several extents, which are allocated across MDisks. The advantage of storage virtualization is that the host is “decoupled” from the underlying storage, so the virtualization appliance can move the extents without affecting the host system.

The host system cannot directly access the underlying MDisks in the same manner as it can access RAID arrays in a traditional storage environment.

There are three types of volumes:

- Striped

  A striped volume is allocated one extent in turn from each MDisk in the storage pool. This process continues until the space required for the volume has been satisfied.

  It is also possible to supply a list of MDisks to use.

  Figure 2-11 shows how a striped volume is allocated, assuming that 10 extents are required.

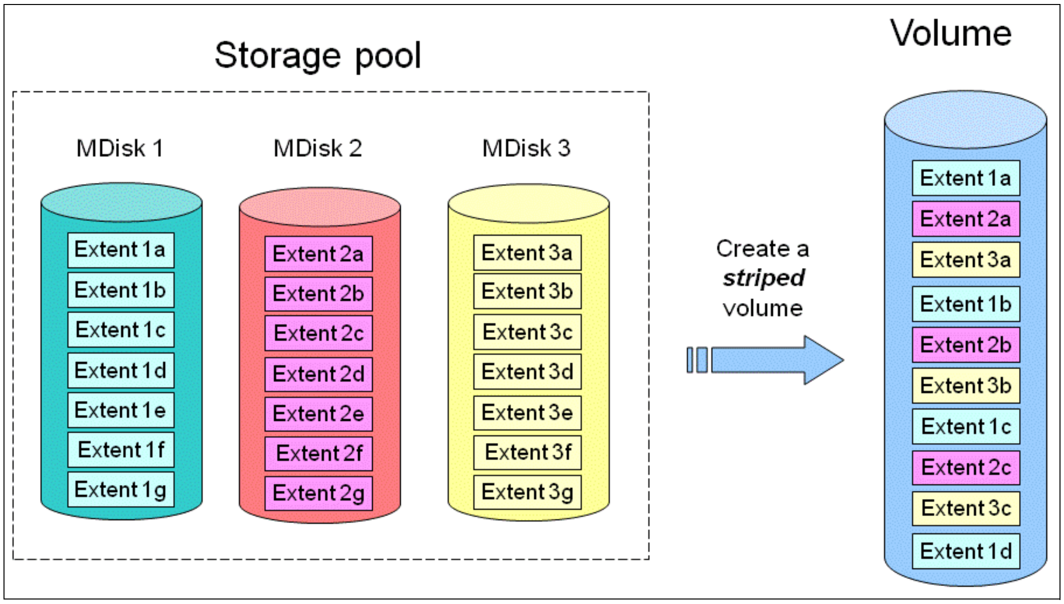


*Figure 2-11 Striped volume*

- Sequential

  A sequential volume is where the extents are allocated one after the other, from one MDisk to the next MDisk (Figure 2-12).

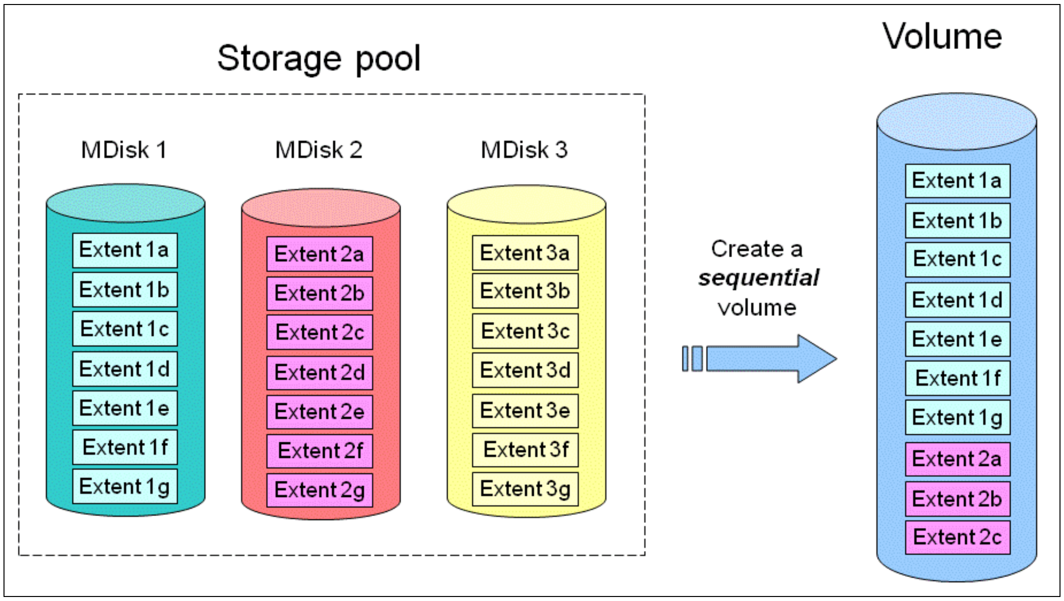


*Figure 2-12 Sequential volume*

- Image mode

  Image mode volumes are special volumes that have a direct relationship with one MDisk. They are used to migrate existing data into and out of the clustered system.

  When the image mode volume is created, a direct mapping is made between extents that are on the MDisk and the extents that are on the volume. The logical block address (LBA) $x$ on the MDisk is the same as the LBA $x$ on the volume, which ensures that the data on the MDisk is preserved as it is brought into the clustered system (Figure 2-13).

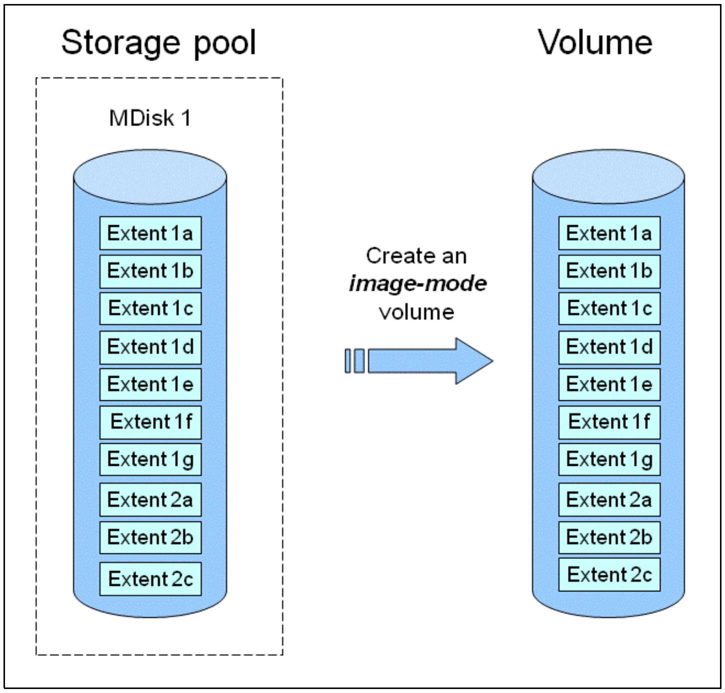


*Figure 2-13 Image mode volume*

Some virtualization functions are not available for image mode volumes, so it is often useful to migrate the volume into a new storage pool. After it is migrated, the MDisk becomes a managed MDisk.

If you add an MDisk containing data to a storage pool, any data on the MDisk is lost. Ensure that you create image mode volumes from MDisks that contain data before adding MDisks to the storage pools.

## 2.5.17 Easy Tier

IBM Easy Tier is a performance optimization function that automatically migrates or moves the extents of a volume across two or more tiers. Easy Tier monitors the host I/O activity and latency on the extent of all volumes with the Easy Tier function turned on, in a multitiered storage pool, over a 24-hour period.

After this period of time, the IBM Easy Tier creates an extent migration plan based on this activity, and then dynamically moves high activity (or *hot*) extents to a higher disk tier within the storage pool. It also moves extent activity that has dropped off (or *cooled*) from the high-tiered MDisk back to a lower-tiered MDisk.

The Easy Tier function can be turned on or off at the storage pool and the volume level. The *automatic storage pool balancing* is an integrated part of the IBM Easy Tier engine and is enabled by default on all pools. Storage Pool Balancing and IBM Easy Tier do not require additional licenses within IBM Spectrum Virtualize running on IBM Storwize V7000.

It is possible to demonstrate the potential benefit of IBM Easy Tier in the environment before installing different drives. When IBM Easy Tier is turned on, the algorithm produces the statistic file that can be offloaded from IBM Storwize V7000.

The IBM Storage Tier Advisor Tool (STAT) can be used to create a summary report using the statistic file offloaded from IBM Storwize V7000.

The STAT tool can be found in the following website:

https://ibm.biz/BdEfrX

## 2.5.18 Encryption

IBM Storwize V7000 provides optional encryption of data at rest, which protects against the potential exposure of sensitive user data and user metadata that is stored on discarded, lost, or stolen storage devices. Encryption of system data and system metadata is not required, so system data and metadata are not encrypted.

Planning for encryption involves purchasing a licensed function and then activating and enabling the function on the system.

License is required to encrypt data that is stored on drives. When encryption is activated and enabled on the system, valid encryption keys must be present on the system when the system unlocks the drives or the user generates a new key.

Within IBM Spectrum Virtualize V7.8, the encryption keys can be either managed by IBM Security Key Lifecycle Manager version or USB Flash drives; but not both. IBM Security Key Lifecycle Manager is an IBM solution to provide infrastructure and processes to locally create, distribute, backup, and manage the lifecycle of keys and certificates.

Before activating and enabling encryption, you must determine the method of accessing key information during times when the system requires an encryption key to be present.

When Security Key Lifecycle Manager is used as a key manager for encryption, you can run into a deadlock situation if the key servers are running on encrypted storage provided by the IBM Storwize V7000.

To avoid a deadlock situation, ensure the IBM Storwize V7000 node canisters are able to “talk” to an encryption server to get the unlock key after a power-on or restart scenario.

Data encryption is protected by the Advanced Encryption Standard (AES) algorithm that uses a 256-bit symmetric encryption key in XTS mode, as defined in the Institute of Electrical and Electronics Engineers (IEEE) 1619-2007 standard as XTS-AES-256. That data encryption key is itself protected by a 256-bit AES key wrap when stored in non-volatile form.

## 2.5.19 iSCSI

iSCSI is an alternative means of attaching hosts and external storage controllers to the IBM Storwize V7000. Within IBM Spectrum Virtualize release 7.7, IBM introduced software capabilities to allow the underlying virtualized storage to attached to IBM Storwize V7000 via iSCSI protocol.

In the simplest terms, iSCSI enables the transport of SCSI commands and data over an Internet Protocol network, based on IP routers and Ethernet switches. iSCSI is a block-level protocol that encapsulates SCSI commands into Transmission Control Protocol/Internet Protocol (TCP/IP) packets and uses an existing IP network, rather than requiring expensive FC host bus adapters (HBAs) and a SAN fabric infrastructure.

The major functions of iSCSI include encapsulation and the reliable delivery of CDB transactions between initiators and targets through the Internet Protocol network, especially over a potentially unreliable IP network.

Every iSCSI node in the network, must have an iSCSI name and address as described below:

- An *iSCSI name* is a location-independent, permanent identifier for an iSCSI node. An iSCSI node has one iSCSI name, which stays constant for the life of the node. The terms *initiator name* and *target name* also refer to an iSCSI name.
- An *iSCSI address* specifies not only the iSCSI name of an iSCSI node, but a location of that node. The address consists of a host name or IP address, a TCP port number (for the target), and the iSCSI name of the node. An iSCSI node can have any number of addresses, which can change at any time, particularly if they are assigned by way of Dynamic Host Configuration Protocol (DHCP). An SVC node represents an iSCSI node and provides statically allocated IP addresses.

### 2.5.20 Real-time Compression

IBM Real-time Compression (RtC) is an attractive solution to address increasing data storage requirements, power, cooling and floor space. When applied, IBM Real-time Compression can significantly save storage space because more data is stored in rack space so fewer storage enclosures are required to store a data set. IBM Real-time Compression provides the following benefits:

- Compression for active primary data. IBM Real-time Compression can be used with active primary data.
- Compression for replicated/mirrored data. Remote volume copies can be compressed in addition to the volumes at the primary storage tier. This process reduces storage requirements in Metro Mirror and Global Mirror destination volumes also.
- No changes to the existing environment are required. IBM Real-time Compression is part of the storage system.
- Overall savings in operational expenses. More data is stored in a rack space, so fewer storage expansion enclosures are required to store a data set. This reduced rack space has the following benefits:
  - Reduced power and cooling requirements. More data is stored in a system, therefore requiring less power and cooling per gigabyte or used capacity.
  - Reduced software licensing for additional functions in the system. More data stored per enclosure reduces the overall spending on licensing.
- Disk space savings are immediate. The space reduction occurs when the host writes the data. This process is unlike other compression solutions, in which some or all of the reduction is realized only after a post-process compression batch job is run.

### 2.5.21 IP replication

IP replication was introduced in V7.2 and allows data replication between IBM Spectrum Virtualize family members. IP replication uses IP-based ports of the cluster node canisters.

IP replication function is transparent to servers and applications in the same way that traditional FC-based mirroring is. All remote mirroring modes (Metro Mirror, Global Mirror, and Global Mirror with changed volumes) are supported.

The configuration of the system is straightforward and IBM Storwize family systems normally find each other in the network and can be selected from the GUI.

IP replication includes Bridgeworks SANSlide network optimization technology, and is available at no additional charge. Remember, remote mirror is a chargeable option but the price does not change with IP replication. Existing remote mirror users have access to the function at no additional charge.

IP connections that are used for replication can have long latency (the time to transmit a signal from one end to the other), which can be caused by distance or by many “hops” between switches and other appliances in the network. Traditional replication solutions transmit data, wait for a response, and then transmit more data, which can result in network utilization as low as 20% (based on IBM measurements).

Bridgeworks SANSlide technology, which is integrated with the IBM Storwize family, requires no separate appliances and so requires no additional cost and no configuration steps. It uses artificial intelligence (AI) technology to transmit multiple data streams in parallel, adjusting automatically to changing network environments and workloads.

SANSlide improves network bandwidth utilization up to 3x. Therefore, customers can deploy a less costly network infrastructure, or take advantage of faster data transfer to speed replication cycles, improve remote data currency, and enjoy faster recovery.

### 2.5.22 IBM Storwize V7000 copy services

IBM Spectrum Virtualize supports the following copy services:

- Synchronous remote copy
- Asynchronous remote copy
- FlashCopy
- Transparent Cloud Tiering

Starting with V6.3 (now IBM Spectrum Virtualize), copy services functions are implemented within a single IBM Storwize V7000 or between multiple members of the IBM Spectrum Virtualize family. The Copy Services layer sits above and operates independently of the function or characteristics of the underlying disk subsystems used to provide storage resources to an IBM Storwize V7000.

### 2.5.23 Synchronous or asynchronous remote copy

The general application of remote copy seeks to maintain two copies of data. Often, the two copies are separated by distance, but not always. The remote copy can be maintained in either synchronous or asynchronous modes.

With IBM Spectrum Virtualize, Metro Mirror and Global Mirror are the IBM branded terms for the functions that are synchronous remote copy and asynchronous remote copy.

Synchronous remote copy ensures that updates are committed at both the primary and the secondary volumes before the application considers the updates complete. Therefore, the secondary volume is fully up to date if it is needed in a failover.

However, the application is fully exposed to the latency and bandwidth limitations of the communication link to the secondary volume. In a truly remote situation, this extra latency can have a significant adverse effect on application performance.

Special configuration guidelines exist for SAN fabrics and IP networks that are used for data replication. There must be considerations in regards the distance and available bandwidth of the intersite links.

A function of Global Mirror designed for low bandwidth has been introduced in IBM Spectrum Virtualize. It uses change volumes that are associated with the primary and secondary volumes. These volumes are used to record changes to the remote copy volume, the FlashCopy relationship that exists between the secondary volume and the change volume, and between the primary volume and the change volume.

This function is called *Global Mirror cycling mode*. Figure 2-14 shows an example of this function where you can see the relationship between volumes and change volumes.

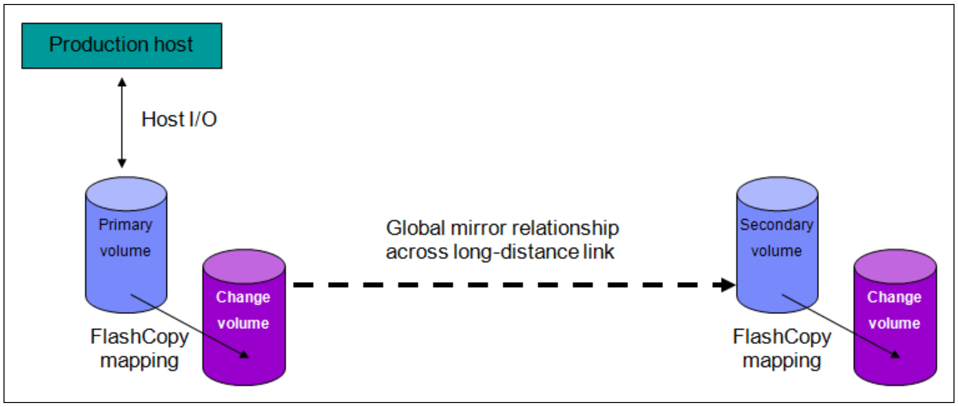


*Figure 2-14 Global Mirror with change volumes*

In asynchronous remote copy, the application acknowledges that the write is complete before the write is committed at the secondary volume. Therefore, on a failover, certain updates (data) might be missing at the secondary volume. The application must have an external mechanism for recovering the missing updates, if possible. This mechanism can involve user intervention. Recovery on the secondary site involves starting the application on this recent backup, and then rolling forward or backward to the most recent commit point.

## 2.5.24 FlashCopy and Transparent Cloud Tiering

FlashCopy and Cloud Backup are used to make a copy of a source volume on a target volume. After the copy operation has started, the original content of the target volume is lost and the target volume has the contents of the source volume as they existed at a single point in time. Although the copy operation takes time, the resulting data at the target appears as though the copy was made instantaneously.

### FlashCopy

FlashCopy is sometimes described as an instance of a time-zero (T0) copy or a point-in-time (PiT) copy technology.

FlashCopy can be performed on multiple source and target volumes. FlashCopy permits the management operations to be coordinated so that a common single point in time is chosen for copying target volumes from their respective source volumes.

With IBM Spectrum Virtualize, multiple target volumes can undergo FlashCopy from the same source volume. This capability can be used to create images from separate points in time for the source volume, and to create multiple images from a source volume at a common point in time. Source and target volumes can be thin-provisioned volumes.

Reverse FlashCopy enables target volumes to become restore points for the source volume without breaking the FlashCopy relationship, and without waiting for the original copy operation to complete. IBM Spectrum Virtualize supports multiple targets, and therefore multiple rollback points.

Most clients aim to integrate the FlashCopy feature for point in time copies and quick recovery of their applications and databases. An IBM solution to this is provided by IBM Spectrum Protect, which is described on the following website:

https://ibm.biz/BdsB5z

### Transparent Cloud Tiering

IBM Spectrum Transparent Cloud Tiering is a new function introduced in IBM Spectrum Virtualize V7.8. Transparent Cloud Tiering is an alternative solution for data protection, backup and restore that interfaces to Cloud Service Providers, such as IBM Softlayer, Amazon S3 and OpenStack Swift. Transparent Cloud Tiering is charged as additional priced software per IBM Storwize V7000 Controller.

The Transparent Cloud Tiering function helps organizations to reduce costs related to power and cooling when off-site data protection is required to send sensitive data out of the main site.

Transparent Cloud Tiering uses IBM FlashCopy techniques that provide full and incremental snapshots of one or more volumes. Snapshots are encrypted and compressed before being uploaded to the cloud. Reverse operations are also supported within that function. When a set of data is transferred out to cloud, the volume snapshot is stored as an object storage.

Cloud object storage uses innovative approach and cost-effective solution to store large amount of unstructured data and delivers mechanisms to provide security services, high availability and reliability.

The management GUI (graphical user interface) provides an easy-to-use initial setup, advanced security settings and audit logs that records all backup and restore to cloud.

To know more about cloud object storage, visit:

https://ibm.biz/Bdsc7m

# 2.6 Business Continuity

Business continuity and continuous application availability are among the top requirements for many organizations. Advances in virtualization, storage, and networking have made enhanced business continuity possible.

Information technology solutions can now manage both planned and unplanned outages, and provide the flexibility and cost efficiencies that are available from cloud-computing models.

## 2.6.1 Business Continuity with HyperSwap

The HyperSwap high availability feature in the IBM Spectrum Virtualize software allows business continuity in the event of hardware failure, power failure, connectivity failure, or disasters such as fire or flooding. The HyperSwap feature is available on the IBM SAN Volume Controller, IBM Storwize V7000, IBM Storwize V7000 Unified, and IBM Storwize V5000 products.

The HyperSwap feature provides highly available volumes accessible through two sites at up to 300km apart. A fully independent copy of the data is maintained at each site. When data is written by hosts at either site, both copies are synchronously updated before the write operation is completed. The HyperSwap feature will automatically optimize itself to minimize data transmitted between sites and to minimize host read and write latency.

Some of the key features of HyperSwap are:

- Works with SVC and IBM Storwize V7000, V5000, V7000 Unified hardware as well.
- Uses intra-cluster synchronous remote copy (Metro Mirror) capabilities along with existing change volume and access I/O group technologies.
- Makes a host's volumes accessible across two IBM V7000 / V5000 Storwize or SVC I/O groups in a clustered system using the Metro Mirror relationship under the covers, look like one volume to the host.
- Works with the standard multipathing drivers that are available on a wide variety of host types, with no additional host support required to access the highly available volume.

# 2.7 Management and support tools

The IBM Spectrum Virtualize system can be managed through the included management software that runs on the IBM Storwize V7000 hardware.

## 2.7.1 IBM Assist On-site and remote service

The IBM Assist On-site tool is a remote desktop-sharing solution that is offered through the IBM website. With it, the IBM service representative can remotely view your system to troubleshoot a problem.

You can maintain a chat session with the IBM service representative so that you can monitor this activity and either understand how to fix the problem yourself or allow the representative to fix it for you.

To use the IBM Assist On-site tool, the master console must be able to access the Internet. The following website provides further information about this tool:

`https://ibm.biz/BdsBNB`

When you access the website, you sign in and enter a code that the IBM service representative provides to you. This code is unique to each IBM Assist On-site session. A plug-in is downloaded on to your master console to connect you and your IBM service representative to the remote service session. The IBM Assist On-site tool contains several layers of security to protect your applications and your computers. The plug-in is removed after the next reboot.

You can also use security features to restrict access by the IBM service representative. Your IBM service representative can provide you with more detailed instructions for using the tool.

## 2.7.2 Event notifications

IBM Storwize V7000 can use Simple Network Management Protocol (SNMP) traps, syslog messages, and a Call Home email to notify you and the IBM Support Center when significant events are detected. Any combination of these notification methods can be used simultaneously.

Each event that IBM Storwize V7000 detects is assigned a notification type of `Error`, `Warning`, or `Information`. You can configure the IBM Storwize V7000 to send each type of notification to specific recipients.

### 2.7.3 Simple Network Management Protocol traps

SNMP is a standard protocol for managing networks and exchanging messages. The IBM Spectrum Virtualize can send SNMP messages that notify personnel about an event. You can use an SNMP manager to view the SNMP messages that IBM Spectrum Virtualize sends. You can use the management GUI or the IBM Storwize V7000 CLI to configure and modify your SNMP settings.

You can use the Management Information Base (MIB) file for SNMP to configure a network management program to receive SNMP messages that are sent by the IBM Spectrum Virtualize.

### 2.7.4 Syslog messages

The syslog protocol is a standard protocol for forwarding log messages from a sender to a receiver on an IP network. The IP network can be either IPv4 or IPv6.

IBM Storwize V7000 can send syslog messages that notify personnel about an event. The event messages can be sent in either expanded or concise format. You can use a syslog manager to view the syslog messages that IBM Storwize V7000 sends.

IBM Spectrum Virtualize uses the User Datagram Protocol (UDP) to transmit the syslog message. You can use the management GUI or the IBM Storwize V7000 CLI to configure and modify your syslog settings.

### 2.7.5 Call Home email

The Call Home feature transmits operational and error-related data to you and IBM through a Simple Mail Transfer Protocol (SMTP) server connection in the form of an event notification email. When configured, this function alerts IBM service personnel about hardware failures and potentially serious configuration or environmental issues. You can use the Call Home function if you have a maintenance contract with IBM or if the IBM Storwize V7000 is within the warranty period.

To send email, at least one SMTP server must be configured. The system support as many as five more SMTP servers for backup purposes. The SMTP server must accept the relaying of email from the IBM Storwize V7000 clustered system IP address.

Use the management GUI or the IBM Storwize V7000 CLI to configure the email settings, including contact information and email recipients. Set the reply address to a valid email address.

Send a test email to check that all connections and infrastructure are set up correctly. The Call Home function can be disabled at any time by using the management GUI or the IBM Storwize V7000 CLI.

## 2.8 Useful IBM Storwize V7000 websites

See the following IBM Storwize V7000 web pages for more information:

- Support page:

  `http://ibm.co/1nyBrTn`
- IBM Storwize V7000 Unified and IBM Storwize V7000 Disk Systems:

http://www.ibm.com/systems/storage/news/center/storwize_v7000/index.html

- List of supported hardware:

  http://www-01.ibm.com/support/docview.wss?uid=ssg1S1009559

- Configuration Limits and Restrictions:

  http://www-01.ibm.com/support/docview.wss?uid=ssg1S1009561

- IBM Knowledge Center:

  https://www.ibm.com/support/knowledgecenter/STVLF4_7.8.0/spectrum.virtualize.780.doc/svirt_ichome_780.html

3

# Planning

In this chapter we describe the required steps when you plan the installation of an IBM Storwize V7000

We also review the implications for your storage network and describe performance considerations.

This chapter includes the following topics:

- General planning rules
- Physical planning
- Logical planning
- Performance considerations

## 3.1 General planning rules

> **Important:** At the time of writing, the statements provided in this book are correct, but they might change. Always verify any statements that are made in this book with the IBM Storwize V7000 supported hardware list, device driver, firmware, and recommended software levels that are available at the following website:
>
> `http://www.ibm.com/support/docview.wss?uid=ssg1S1009559`

For more information about the topics that are described, see *IBM System Storage SAN Volume Controller and Storwize V7000 Best Practices and Performance Guidelines*, SG24-7521:

`http://www.redbooks.ibm.com/abstracts/sg247521.html`

Complete the following tasks when you are planning for the Storwize V7000:

- Collect and document the number of hosts (application servers) to attach to the Storwize V7000, the traffic profile activity (read or write, sequential, or random), and the performance requirements, which are input/output (I/O) operations per second (IOPS).
- Collect and document the following storage requirements and capacities:
  - The total back-end storage that is present in the environment to be provisioned on the Storwize V7000
  - The total back-end new storage to be provisioned on the Storwize V7000
  - The required virtual storage capacity that is used as a fully managed virtual disk (volume) and used as a space-efficient (SE) volume
  - The required storage capacity for local mirror copy (volume mirroring)
  - The required storage capacity for point-in-time copy (IBM FlashCopy)
  - The required storage capacity for remote copy (Metro Mirror and Global Mirror)
  - The required storage capacity for compressed volumes
  - The required storage capacity for encrypted volumes
  - Per host:
    - Storage capacity
    - Logical unit number (LUN) quantity
    - Sizes
- Define the local and remote SAN fabrics and systems in the cluster if a remote copy or a secondary site is needed.
- Design the SAN according to the requirement for high availability (HA) and best performance. Consider the total number of ports and the bandwidth that is needed between the host and the Storwize V7000, the Storwize V7000 and the disk subsystem, between the Storwize V7000 nodes, and for the inter-switch link (ISL) between the local and remote fabric.

> **Note:** Check and carefully count the required ports for extended links. Especially in an Hyperswap (HS) environment, you might need many of the higher-cost longwave gigabit interface converters (GBICs).

- Design the iSCSI network according to the requirements for high availability (HA) and best performance. Consider the total number of ports and bandwidth that is needed between the host and the Storwize V7000.
- Determine the Storwize V7000 service Internet Protocol (IP) address.
- Determine the IP addresses for the Storwize V7000 system and for the host that connects through Internet Small Computer System Interface (iSCSI).
- Determine the IP addresses for IP replication.
- Define a naming convention for the Storwize V7000 nodes, host, and storage subsystem.
- Define the managed disks (MDisks) in the disk subsystem.
- Define the storage pools. The storage pools depend on the disk subsystem that is in place and the data migration requirements.
- Plan the logical configuration of the volume within the I/O Groups and the storage pools to optimize the I/O load between the hosts and the Storwize V7000.
- Plan for the physical location of the equipment in the rack.

Storwize V7000 planning can be categorized into the following types:

- Physical planning
- Logical planning

We describe these planning types in the following sections.

## 3.2 Physical planning

You must consider several key factors when you are planning the physical site of a Storwize V7000 installation. The physical site must have the following characteristics:

- Power, cooling, and location requirements are present for the Storwize V7000
- You must plan for two separate power sources if you have a redundant ac-power switch, which is available as an optional feature.
- The Storwize V7000 requires two Electronic Industries Alliance (EIA) units for each control enclosure or expansion enclosure.
- Other hardware devices can be in the same Storwize V7000 rack, such as IBM Storwize Virtualize (SVC), IBM Storwize V3700, SAN switches, an Ethernet switch, and other devices.
- You must consider the maximum power rating of the rack; do not exceed it. For more information about the power requirements, see the following website:

  `https://ibm.biz/BdsBU7`

### 3.2.1 Cable connections

Create a cable connection table or documentation that follows your environment’s documentation procedure to track all of the following connections that are required for the setup:

- Nodes
- Ethernet
- iSCSI or Fibre Channel over Ethernet (FCoE) connections
- FC ports

# 3.3 Logical planning

For logical planning, we describe the following topics:

- Management IP addressing plan
- SAN zoning and SAN connections
- iSCSI IP addressing plan
- ISCSI external storage attachment
- External storage configuration planning
- Host failover
- IP mirroring
- Planning for external storage virtualization
- Storwize V7000 clustered system configuration
- Storage pool configuration
- Volume configuration
- Host attachment planning
- Host mapping (LUN masking)
- NPIV planning
- Advanced Copy Services
- SAN boot support
- Data migration from a non-virtualized storage subsystem
- Storwize V7000 configuration backup procedure

## 3.3.1 Management IP addressing plan

To plan your installation, you need to consider the Transmission Control Protocol/Internet Protocol (TCP/IP) address requirements of the Storwize V7000 Gen2/Gen2+, and the requirements to access other services. You must also plan the address allocation and the Ethernet router, gateway, and firewall configuration to provide the required access and network security.

Figure 3-1 on page 47 shows the TCP/IP ports and services that are used by the Storwize V7000 Gen2.

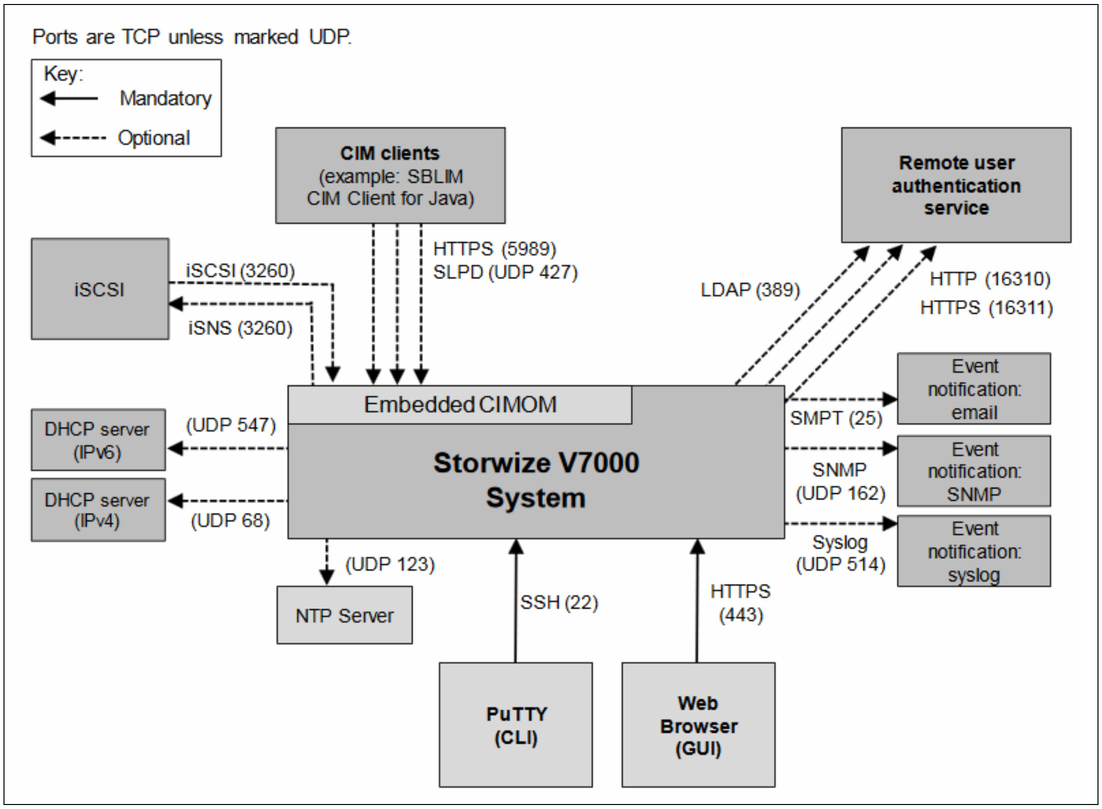


*Figure 3-1 TCP/IP ports*

Table 3-1 shows the list of ports and services used by the Storwize V7000 Gen2.

*Table 3-1 TCP/IP ports and services listing*

| Service | Traffic direction | Protocol | Port | Service type |
|---|---|---|---|---|
| Email Simple Mail Transfer Protocol (SMTP) notification and inventory reporting | Outbound | TCP | 25 | optional |
| Simple Network Management Protocol (SNMP) event notification | Outbound | User Datagram Protocol (UDP) | 162 | optional |
| Syslog event notification | Outbound | UDP | 514 | optional |
| IPv4 Dynamic Host Configuration Protocol (DHCP) node service address | Outbound | UDP | 68 | optional |
| IPv6 DHCP (node service address) | Outbound | UDP | 547 | optional |
| Network Time Protocol (NTP) server | Outbound | UDP | 123 | optional |
| Secure Shell (SSH) for command-line interface (CLI) access | Inbound | TCP | 22 | mandatory |
| Hypertext Transfer Protocol Secure (HTTPS) for graphical user interface (GUI) access | Inbound | TCP | 443 | mandatory |
| Common Information Model object manager (CIMOM) HTTPS | Inbound | TCP | 5989 | optional |

| Service | Traffic direction | Protocol | Port | Service type |
|---|---|---|---|---|
| CIMOM Service Location Protocol Daemon (SLPD) | Inbound | UDP | 427 | optional |
| Remote user authentication service (HTTP) | Outbound | TCP | 16310 | optional |
| Remote user authentication service (HTTPS) | Outbound | TCP | 16311 | optional |
| Remote user authentication service: Lightweight Directory Access Protocol (LDAP) | Outbound | TCP | 389 | optional |
| Internet Small Computer System Interface (iSCSI) | Inbound | TCP | 3260 | optional |
| iSCSI Internet Storage Name Service (iSNS) | Outbound | TCP | 3260 | optional |
| IP Partnership management IP communication | Inbound | TCP | 3260 | optional |
| IP Partnership management IP communication | Outbound | TCP | 3260 | optional |
| IP Partnership data path connections | Inbound | TCP | 3265 | optional |
| IP Partnership data path connections | Outbound | TCP | 3265 | optional |

## Prerequisites

Ensure that the Storwize V7000 Gen2 (control and expansion enclosures) has been physically installed with the correct cabling. Verify that Ethernet and Fibre Channel (FC) connectivity has been correctly configured. Before configuring the IBM Storwize V7000 Gen2, ensure that the following information is available:

- Licenses

  The licenses indicate whether the client is permitted to use IBM Easy Tier, IBM FlashCopy, External Virtualization, Encryption, Remote Copy, and IBM Real-time Compression.

- For IPv4 addressing:
  - Cluster IPv4 address. This address includes one IP address for management of the Storwize V7000 Gen2 System.
  - Service IPv4 addresses. These addresses include at least two IPv4 addresses for the service interfaces (one for each node canister).
  - IPv4 subnet mask.
  - Gateway IPv4 address.
- For IPv6 addressing:
  - Cluster IPv6 address. This address includes one address for management of the Storwize V7000 Gen2 System.
  - Service IPv6 addresses. These addresses include at least two IPv6 addresses for the service interface (one for each node canister).
  - IPv6 prefix.
  - Gateway IPv6 address.

## GUI requirements

To access the IBM Storwize V7000 management GUI, direct a web browser to the system management IP address after the system initialization described in Chapter 4, “Initial configuration” on page 85.

> **Important:** Ensure that your web browser is supported and has the appropriate settings enabled. For more information, see the IBM Knowledge Center for Storwize V7000:
>
> http://ibmurl.hursley.ibm.com/NXSN

To access the management GUI, you direct a web browser to the system management IP address.

Only the Storwize V7000 Gen2/Gen2+ has a *Technician port*. The technician port is the Ethernet port labeled T on the rear of the node canister. The port broadcasts a Dynamic Host Configuration Protocol (DHCP) service so that a notebook or computer is automatically assigned an IP address on connection to the port.

After the cluster configuration has been completed, the Technician port automatically routes the connected user directly to the service GUI.

> **Note:** The default IP address for the Technician port on a Storwize V7000 is 192.168.0.1. If the Technician port is connected to a switch, it is disabled and an error is logged.

Each Storwize V7000 node requires one Ethernet cable to connect it to an Ethernet switch or hub. The cable must be connected to port 1. A 10/100/1000 megabit (Mb) Ethernet connection is required for each cable. Both Internet Protocol Version 4 (IPv4) and Internet Protocol Version 6 (IPv6) are supported.

For more information, see Storwize V7000 Infocenter:

http://ibmurl.hursley.ibm.com/NXSN

> **Note:** For increased redundancy, an optional second Ethernet connection is supported for each Storwize V7000 node. This cable is connected to Ethernet port 2.

To ensure system failover operations, Ethernet port 1 on all nodes must be connected to the same set of subnets. If used, Ethernet port 2 on all nodes must also be connected to the same set of subnets. However, the subnets for Ethernet port 1 do not have to be the same as Ethernet port 2.

Each Storwize V7000 cluster has a Cluster Management IP address, in addition to a Service IP address for each node in the cluster. See Example 3-1 for details.

*Example 3-1 Management IP address sample*

```
management IP add. 10.11.12.120
node 1 service IP add. 10.11.12.121
node 2 service IP add. 10.11.12.122
```

Each node in a Storwize V7000 system needs to have at least one Ethernet connection.

When accessing the Storwize V7000 through the GUI or Secure Shell (SSH), choose one of the available IP addresses to which to connect. No automatic failover capability is available. If one network is down, use an IP address on the alternate network. Clients might be able to use the intelligence in domain name servers (DNSs) to provide partial failover.

### 3.3.2 SAN zoning and SAN connections

External storage controllers virtualized by IBM Storwize V7000 must be connected through SAN switches. A direct connection between the IBM Storwize V7000 and storage controllers or hosts ports is not supported.

Ensure that the switches or directors are at the firmware levels supported by the IBM Storwize V7000, and that the IBM Storwize V7000 port login maximums listed in the restriction document will not be exceeded. The configuration restrictions can be found by navigating to the Support home page at the following address:

http://www.ibm.com/support

The suggested SAN configuration is composed of a minimum of two fabrics. To provide redundancy if one of the fabrics goes offline, the ports on external storage systems that will be virtualized by IBM Storwize V7000 and its ports are evenly split between the two fabrics.

The SAN fabric is zoned to allow the Storwize V7000 to “see” each other’s nodes and the disk subsystems, and for the hosts to see the Storwize V7000 nodes. The hosts cannot directly see or operate LUNs on the disk subsystems that are assigned to the Storwize V7000 system. The Storwize V7000 nodes within a Storwize V7000 system must see each other and all of the storage that is assigned to the Storwize V7000 system.

The zoning capabilities of the SAN switch are used to create three distinct zones. V7.8 supports 2 Gbps, 4 Gbps, 8 Gbps, and 16 Gbps FC fabric, depending on the hardware platform and on the switch where the Storwize V7000 is connected. In an environment where you have a fabric with multiple-speed switches, the preferred practice is to connect the Storwize V7000 and the disk subsystem to the switch operating at the highest speed.

All Storwize V7000 nodes in the Storwize V7000 clustered system are connected to the same SANs, and they present volumes to the hosts. These volumes are created from storage pools that are composed of MDisks that are presented by the disk subsystems.

**Ports:** IBM Storwize V7000 supports a maximum of 16 ports or WWPNs from a given external storage system that will be virtualized.

Figure 3-2 on page 51 is an example of how to cable the devices to the SAN. Refer to this example as we describe the storage to node zoning.

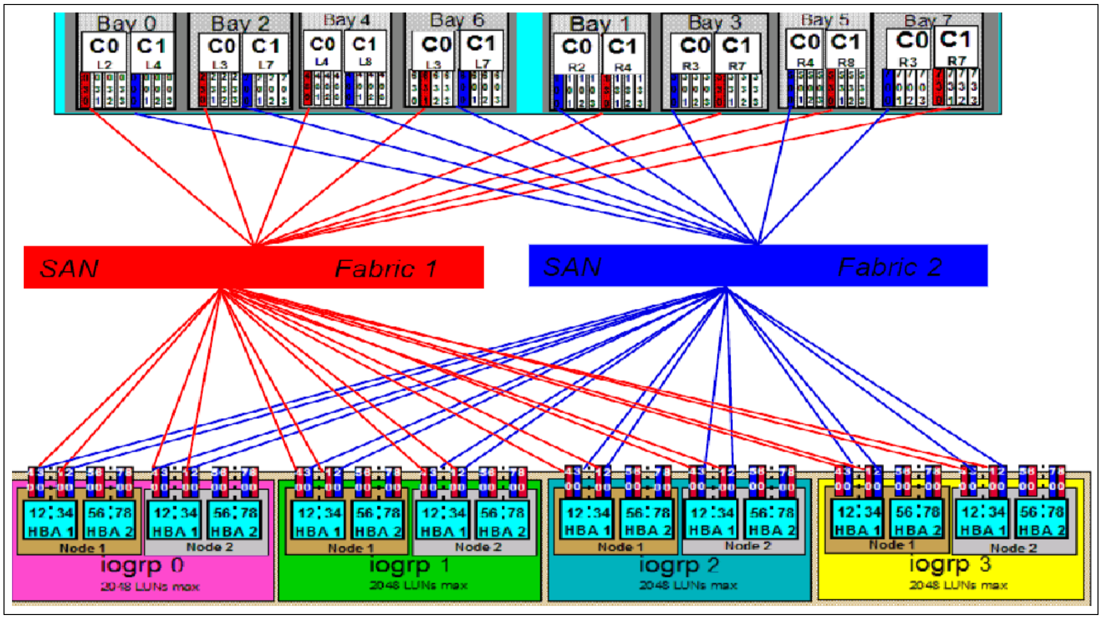


*Figure 3-2 Storage to zoning with 8 ports in an existing environment*

Figure 3-2 refers to IBM SAN Volume Controller 2145-CG8/CF8 nodes using 4 ports per node but the same zoning method is also applicable to 2145-DH8 and IBM Storwize V7000 node zoning. If you are planning to use more than 4 ports per IBM Storwize V7000 canister, you must do additional cabling. But the zoning concept with redundancy remains the same.

Each node has four FC ports standard, and optionally four additional FC ports (or two 10 gigabits per second (Gbps) Ethernet ports for FCoE use). Each I/O group has two nodes. Therefore, with no zoning in a dual SAN environment, the number of paths to a volume are as follows:

- With a standard four-port FC HBA, it would be four multiplied by the number of host ports. For example, if a host has two ports, you multiply two times four resulting in eight paths, which are the maximum supported.
- With standard four-port FC HBA and an optional second four-port HBA, it would be eight multiplied by the number of host ports. For example, if a host has two ports, you multiply two times eight resulting in 16 paths, which exceed the limit of eight and is not supported.
- With standard four-port FC HBA and an optional two-port FCoE HBA, it would be six multiplied by the number of host ports. For example, if a host has two ports, you multiply two times six resulting in 12 paths, which exceed the limit of eight and is not supported.

This rule exists to limit the number of paths that must be resolved by the multipathing device driver. More paths do not equate to better performance or higher availability. For optimum performance and availability, limit a host with two FC ports to only four paths: One path to each node on each SAN.

To restrict the number of paths to a host, zone the switches so that each HBA port is zoned with one port from each node in each I/O group that it accesses volumes from. If a host has multiple HBA ports, zone each port to a different set of Storwize V7000 node ports to maximize performance and redundancy. This also applies to a host with a Converged Network Adapter (CNA) that accesses volumes using FCoE.

### 3.3.3 iSCSI IP addressing plan

- Storwize V7000 uses the built-in Ethernet ports for iSCSI traffic. If the optional 10 Gbps Ethernet feature is installed, you can connect host systems through the two 10 Gbps Ethernet ports per node.
- Storwize V7000 supports the Challenge Handshake Authentication Protocol (CHAP) authentication methods for iSCSI.
- iSCSI IP addresses can fail over to the partner node in the I/O Group if a node fails. This design reduces the need for multipathing support in the iSCSI host.
- iSCSI IP addresses can be configured for one or more nodes.
- iSCSI Simple Name Server (iSNS) addresses can be configured in the SVC.
- Note that `iqn.1986-03.com.ibm:2076.<cluster_name>.<node_name>` is the iSCSI qualified name (IQN) for a Storwize V7000 node. Because the IQN contains the clustered system name and the node name, it is important not to change these names after iSCSI is deployed.
- Each node can be given an iSCSI alias, as an alternative to the IQN.
- The IQN of the host to a Storwize V7000 host object is added in the same way that you add FC worldwide port names (WWPNs).
- Host objects can have WWPNs and IQNs.

Standard iSCSI host connection procedures can be used to discover and configure the Storwize V7000 as an iSCSI target.

### 3.3.4 iSCSI external storage attachment

V7.7 of IBM Spectrum Virtualize introduced support for external storage controllers attached through iSCSI.

For more information about supported iSCSI Storage Controllers see:

http://www.ibm.com/support/docview.wss?uid=ssg1S1009559

#### *Online Data Migration support for the following use cases:*

These are all the supported possibilites to migrate your data between different kind of attached storage controllers.

- From (3rd party) iSCSI Controller to internal SAS storage pool
- From (3rd party) iSCSI Controller to FC/FCoE storage pool
- From (3rd party) iSCSI Controller to iSCSI storage pool
- From (3rd party) FC/FCoE storage Controller to iSCSI storage pool

### 3.3.5 External storage configuration planning

External storage systems provide redundancy through various Redundant Array of Independent Disks (RAID) levels, which prevents a single physical disk failure from causing a managed disk (MDisk), storage pool, or associated host volume, from going offline. To minimize the risk of data loss, only virtualize storage systems where LUNs are configured using a RAID level other than RAID 0 (for example RAID 1, RAID 10, RAID 0+1, RAID 5, or RAID 6).

Verify that the storage controllers to be virtualized by IBM Storwize V7000 meet the requirements, and that you have read any configuration restrictions by navigating to the IBM Support Portal website:

http://www.ibm.com/support

Ensure that the firmware or microcode levels of the storage controllers to be virtualized are supported by IBM Storwize V7000.

IBM Storwize V7000 must have exclusive access to the LUNs from the external storage system mapped to it. LUN access cannot be shared between IBM Storwize V7000s, between an IBM Storwize V7000 and other storage virtualization platforms, or between an IBM Storwize V7000 and hosts. However, different LUNs can be mapped from one external storage system to an IBM Storwize V7000 and other hosts in the SAN through different storage ports.

Make sure to configure the storage subsystem LUN masking settings to map all LUNs to all the WWPNs in the IBM Storwize V7000 storage system.

Be sure to go to the IBM Storwize V7000 page and review the *Configuring and servicing external storage system* topic before you prepare the external storage systems for discovery by the IBM Storwize V7000 system. This website is at the following location:

http://ibmurl.hursley.ibm.com/NXSN

You can also go to this link:

http://ibmurl.hursley.ibm.com/NXSO

### Guidelines for virtualizing external storage

When virtualizing external storage with the IBM Storwize V7000, follow these guidelines:

- Avoid splitting arrays into multiple LUNs at the external storage system level. When possible, create a single LUN per array for mapping to the IBM Storwize V7000.
- Except for IBM Easy Tier, do not mix MDisks that vary in performance or reliability in the same storage pool. Always put similarly sized MDisks into one storage pool. For more information about Easy Tier, see Chapter 10, “Advanced features for storage efficiency” on page 363.
- Do not leave volumes in image mode. Only use image mode to import or export existing data into or out of the IBM Storwize V7000. Migrate such data from image mode MDisks to other storage pools to benefit from storage virtualization.
- Using the copy services in Storwize V7000 gives you a unified method to manage data integrity across heterogeneous storage systems.
- The Easy Tier function is included with the IBM Storwize V7000 system, and the external storage system could benefit from this powerful storage tiering function to remove hot spots and improve overall performance.

### Port designation recommendations

The port to local node communication is used for mirroring write cache and metadata exchange between nodes, and is critical to the stable operation of the cluster. The Storwize V7000 Gen2 and Gen2+ nodes with their 8-port,12-port and 16-port configurations provide an opportunity to isolate the port to local node traffic from other cluster traffic on dedicated ports. Therefore, the nodes provide a level of protection against malfunctioning devices and workloads that could compromise the performance of the shared ports.

Additionally, there is a benefit in isolating remote replication traffic on dedicated ports, and ensuring that problems that affect the cluster-to-cluster interconnect do not adversely affect ports on the primary cluster. If not, these problems affect the performance of workloads running on the primary cluster.

### Zoning considerations for Metro Mirror and Global Mirror

With Metro Mirror and Global Mirror configurations, more zones are required that contain only the local nodes and the remote nodes. It is valid for the local hosts to see the remote nodes or for the remote hosts to see the local nodes. Any zone that contains the local and the remote back-end storage systems and local nodes or remote nodes, or both, is not valid.

For best results in Metro Mirror and Global Mirror configurations where the round-trip latency between systems is less than 80 milliseconds, zone each node so that it can communicate with at least one Fibre Channel port on each node in each remote system. This configuration maintains redundancy of the fault tolerance of port and node failures within local and remote systems.

However, to accommodate the limitations of some switch vendors on the number of ports or WWNNs that are allowed in a zone, you can further reduce the number of ports or WWNNs in a zone. Such a reduction can result in reduced redundancy and extra workload being placed on other system nodes, and on the Fibre Channel links between the nodes of a system.

If the round-trip latency between systems is greater than 80 milliseconds, stricter configuration requirements apply:

- Use SAN zoning and port masking to ensure that two Fibre Channel ports on each node that is used for replication are dedicated for replication traffic.
- Apply SAN zoning to provide separate intersystem zones for each local-to-remote input/output (I/O) group pair that is used for replication. See the information about long-distance links for Metro Mirror and Global Mirror partnerships for further details.

The minimum configuration requirement is to zone both nodes in one I/O group to both nodes in one I/O group at the secondary site. The I/O group maintains fault tolerance of a node or port failure at either the local or remote site location.

It does not matter which I/O groups at either site are zoned, because I/O traffic can be routed through other nodes to get to the destination. However, if an I/O group that is doing the routing contains the nodes that are servicing the host I/O, there is no additional burden or latency for those I/O groups. There is no added burden because the I/O group nodes are directly connected to the remote system.

If only a subset of the I/O groups within a system is using Metro Mirror and Global Mirror, you can restrict the zoning so that only those nodes can communicate with nodes in remote systems. You can have nodes that are not members of any system zoned to detect all the systems. You can then add a node to the system in case you must replace a node.

For more information about zoning refer to *IBM System Storage SAN Volume Controller and Storwize V7000 Best Practices and Performance Guidelines*, SG24-7521.

You must also observe the following guidelines:

- LUNs (MDisks) must have exclusive access to a single Storwize V7000 clustered system and cannot be shared between other Storwize V7000 clustered systems or hosts.
- A storage controller can present LUNs to the Storwize V7000 (as MDisks) and to other hosts in the SAN. However, in this case, it is better to avoid Storwize V7000 and hosts to share storage ports.

- Mixed port speeds are not possible for intracluster communication. All node ports within a clustered system must be running at the same speed.
- ISLs are not to be used for intracluster node communication or node-to-storage controller access.
- The switch configuration in a Storwize V7000 fabric must comply with the switch manufacturer's configuration rules, which can impose restrictions on the switch configuration. For example, a switch manufacturer might limit the number of supported switches in a SAN. Operation outside of the switch manufacturer's rules is not supported.
- HBAs in dissimilar hosts or dissimilar HBAs in the same host must be in separate zones. For example, IBM AIX and Microsoft hosts must be in separate zones. In this case, *dissimilar* means that the hosts are running separate operating systems or are using separate hardware platforms. Therefore, various levels of the same operating system are regarded as *similar*. This requirement is a SAN interoperability issue, rather than a Storwize V7000 requirement.
- Host zones are to contain only one initiator (HBA) each, and as many Storwize V7000 node ports as you need, depending on the high availability and performance that you want from your configuration.

**Important:** Be aware of the following considerations:

- The use of ISLs for intracluster node communication can negatively affect the availability of the system because of the high dependency on the quality of these links to maintain heartbeat and other system management services. Therefore, we strongly advise that you use them only as part of an interim configuration to facilitate SAN migrations, and not as part of the designed solution.
- The use of ISLs for Storwize V7000 node to storage controller access can lead to port congestion, which can negatively affect the performance and resiliency of the SAN. Therefore, we strongly advise that you use them only as part of an interim configuration to facilitate SAN migrations, and not as part of the designed solution.
- The use of mixed port speeds for intercluster communication can lead to port congestion, which can negatively affect the performance and resiliency of the SAN. Therefore, it is not supported.

You can use the **`lsfabric`** command to generate a report that displays the connectivity between nodes and other controllers and hosts. This report is helpful for diagnosing SAN problems.

## 3.3.6 Host failover

From a host perspective, a multipathing I/O driver is not required to handle a Storwize V7000 node failover. In a Storwize V7000 node restart, the host reconnects to the IP addresses of the iSCSI target node that reappear after several seconds on the ports of the partner node.

A host multipathing I/O driver for iSCSI *is* required in the following situations:

- To protect a host from network link failures, including port failures on the Storwize V7000 nodes
- To protect a host from an HBA failure (if two HBAs are in use)
- To protect a host from network failures, if the host is connected through two HBAs to two separate networks
- To provide load balancing on the server's HBA and the network links

The commands for the configuration of the iSCSI IP addresses were separated from the configuration of the cluster IP addresses.

The following commands are new commands that are used for managing iSCSI IP addresses:

- The **`lsportip`** command lists the iSCSI IP addresses that are assigned for each port on each Storwize V7000 node in the cluster.
- The **`cfgportip`** command assigns an IP address to each node's Ethernet port for iSCSI I/O.

The following commands are new commands that are used for managing the cluster IP addresses:

- The **`lssystemip`** command returns a list of the cluster management IP addresses that are configured for each port.

The **`chsystemip`** command modifies the IP configuration parameters for the cluster.

The parameters for remote services (Secure Shell (SSH) and web services) remain associated with the cluster object. During a Storwize V7000 code upgrade, the configuration settings for the clustered system are applied to the node Ethernet port 1.

For iSCSI-based access, the use of redundant network connections, and separating iSCSI traffic by using a dedicated network or VLAN, prevents any NIC, switch, or target port failure from compromising the host server's access to the volumes.

Because both onboard Ethernet ports of a Storwize V7000 node can be configured for iSCSI, we advise that you dedicate Ethernet port 1 for Storwize V7000 management and port 2 for iSCSI usage. By using this approach, port 2 can be connected to a dedicated network segment or VLAN for iSCSI. Because the Storwize V7000 does not support the use of VLAN tagging to separate management and iSCSI traffic, you can assign the correct LAN switch port to a dedicated VLAN to separate Storwize V7000 management and iSCSI traffic.

## 3.3.7 IP mirroring

One of the most important new functions of V7.2 and later in the Storwize family is IP replication, which enables the use of lower-cost Ethernet connections for remote mirroring. The capability is available as a licensable option (Metro Mirror or Global Mirror) on all Storwize family systems. The new function is transparent to servers and applications in the same way that traditional FC-based mirroring is transparent. All remote mirroring modes (Metro Mirror, Global Mirror, and Global Mirror with Change Volumes) are supported.

The configuration of the system is straightforward. Storwize family systems normally can find each other in the network, and can be selected from the GUI. IP replication includes Bridgeworks SANSlide network optimization technology, and is available at no additional charge. Remote mirror is a licensable option but the price does not change with IP replication. Existing remote mirror users have access to the new function at no additional charge.

IP connections that are used for replication can have a long *latency* (the time to transmit a signal from one end to the other), which can be caused by distance or by many hops between switches and other appliances in the network. Traditional replication solutions transmit data, wait for a response, and then transmit more data, which can result in network usage as low as 20% (based on IBM measurements). This situation gets worse as the latency gets longer.

Bridgeworks SANSlide technology that is integrated with the IBM Storwize family requires no separate appliances; therefore, no other costs and configuration are necessary. It uses

Artificial Intelligence (AI) technology to transmit multiple data streams in parallel, and adjusts automatically to changing network environments and workloads.

Because SANSlide does not use compression, it is independent of application or data type. Most importantly, SANSlide improves network bandwidth usage up to 3x, so clients might be able to deploy a less costly network infrastructure. Also, they could use faster data transfer to speed replication cycles, improve remote data currency, and recover more quickly.

**Note:** The limiting factor of the distance is the round-trip time. The maximum supported round-trip time between sites is 80 milliseconds (ms) for a 1 Gbps link. For a 10 Gbps link, the maximum supported round-trip time between sites is 10 ms.

### Key features of IP mirroring

IBM offers the new, enhanced function of IP-based Remote Copy services, which are primarily targeted to small and midrange environments, where clients typically cannot afford the cost of FC-based replication between sites.

IP-based replication offers the following new features:

- Remote Copy modes support Metro Mirror, Global Mirror, and Global Mirror with Change Volumes.
- All platforms that support Remote Copy are supported.
- The configuration uses automatic path configuration through the discovery of a remote cluster. You can configure any Ethernet port (10Gb/1Gb) for replication that uses Remote Copy port groups.
- Dedicated Ethernet ports for replication.
- CHAP-based authentication is supported.
- Licensing is the same as the existing Remote Copy.
- High availability features auto-failover support across redundant links.
- Performance is based on a vendor-supplied IP connectivity solution, which has experience in offering low bandwidth, high latency long-distance IP links. Support is for 80 ms round-trip time at a 1 Gbps link.

Figure 3-3 shows the schematic way to connect two sides through IP mirroring.

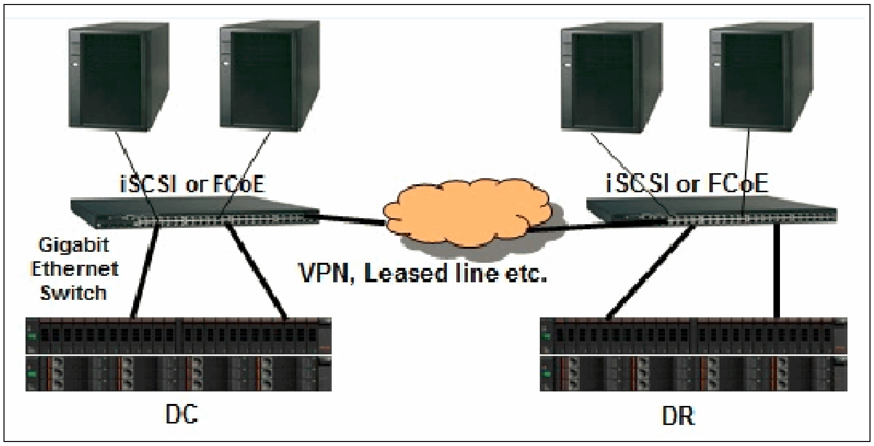


*Figure 3-3 IP mirroring*

Figure 3-4 and Figure 3-5 on page 58 show configuration possibilities for connecting two sites through IP mirroring. Figure 3-4 shows the configuration with single links.

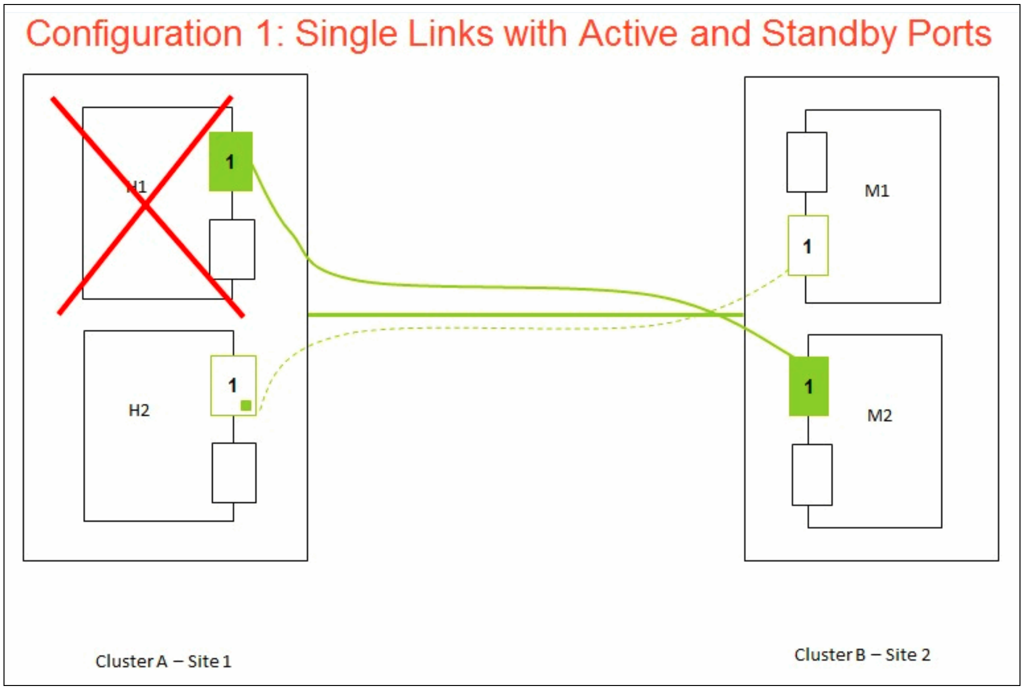


*Figure 3-4 Single link configuration*

The administrator must configure at least one port on each site to use with the link. Configuring more than one port means that replication continues, even if a node fails. Figure 3-5 shows a redundant IP configuration with two links.

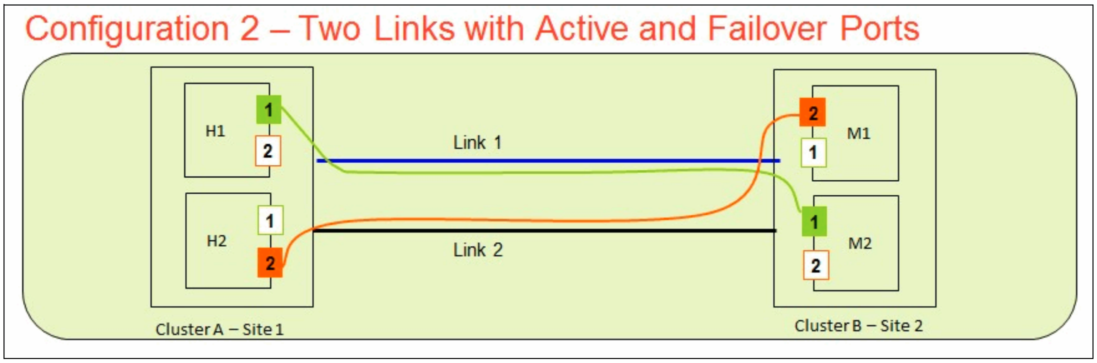


*Figure 3-5 Two links with active and failover ports*

As shown in Figure 3-5, the following replication group setup for dual redundant links is used:

- Replication Group 1: Four IP addresses, each on a different node (green)
- Replication Group 2: Four IP addresses, each on a different node (orange)

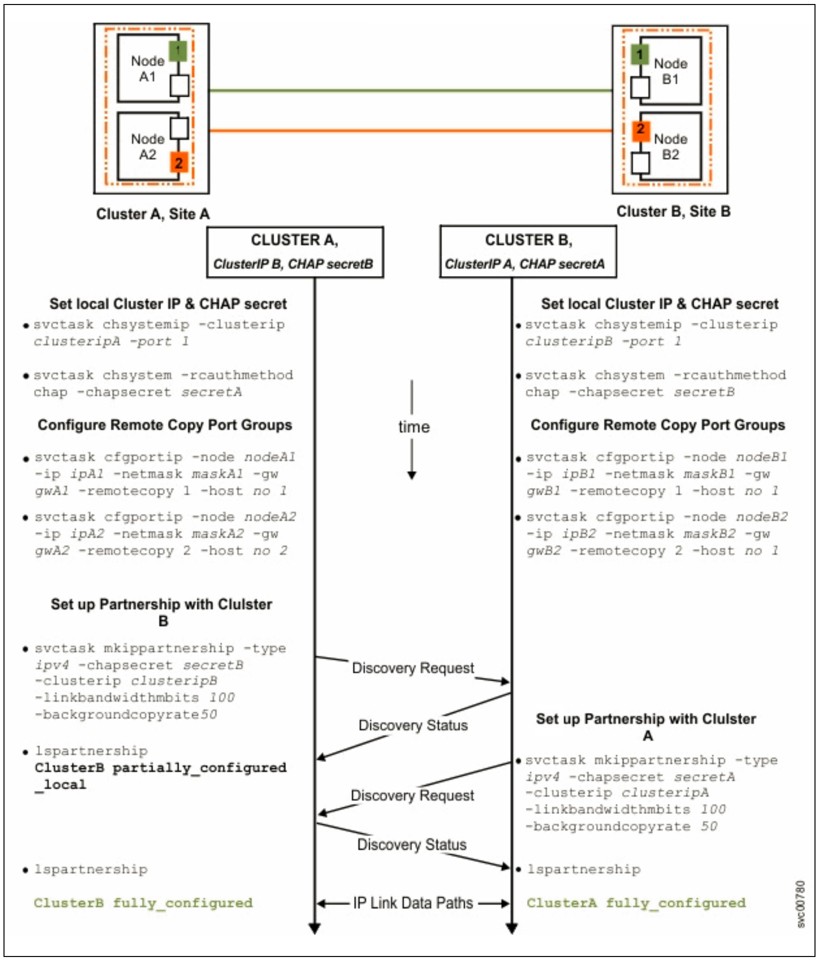
The following simultaneous IP replication sessions can be used at any time:

- Possible user configuration of each Ethernet port:
  - Not used for IP replication (default)
  - Used for IP replication, link 1
  - Used for IP replication, link 2

- IP replication status for each Ethernet port:
  - Not used for IP replication
  - Active (solid box)
  - Standby (outline box)

Figure 3-6 on page 59 shows the configuration of an IP partnership. You can obtain the requirements to set up an IP partnership on the following website:

http://ibmurl.hursley.ibm.com/NXSN



**Set local Cluster IP & CHAP secret**

```
• svctask chsystemip -clusterip
  clusteripA -port 1

• svctask chsystem -rcauthmethod
  chap -chapsecret secretA
```

**Configure Remote Copy Port Groups**

```
• svctask cfgportip -node nodeA1
  -ip ipA1 -netmask maskA1 -gw
  gwA1 -remotecopy 1 -host no 1
• svctask cfgportip -node nodeA2
  -ip ipA2 -netmask maskA2 -gw
  gwA2 -remotecopy 2 -host no 2
```

**Set up Partnership with Clulster B**

```
• svctask mkippartnership -type
  ipv4 -chapsecret secretB
  -clusterip clusteripB
  -linkbandwidthmbits 100
  -backgroundcopyrate50

• lspartnership
  ClusterB partially_configured
  _local

• lspartnership

  ClusterB fully_configured
```

**Set local Cluster IP & CHAP secret**

```
• svctask chsystemip -clusterip
  clusteripB -port 1

• svctask chsystem -rcauthmethod
  chap -chapsecret secretB
```

**Configure Remote Copy Port Groups**

```
• svctask cfgportip -node nodeB1
  -ip ipB1 -netmask maskB1 -gw
  gwB1 -remotecopy 1 -host no 1
• svctask cfgportip -node nodeB2
  -ip ipB2 -netmask maskB2 -gw
  gwB2 -remotecopy 2 -host no 1
```

**Set up Partnership with Clulster A**

```
• svctask mkippartnership -type
  ipv4 -chapsecret secretA
  -clusterip clusteripA
  -linkbandwidthmbits 100
  -backgroundcopyrate 50

• lspartnership

  ClusterA fully_configured
```

*Figure 3-6 IP partnership configuration*

## Terminology for IP replication

This section lists the following terminology for IP replication:

**Discovery** This term refers to the process by which two Storwize V7000 clusters exchange information about their IP address configuration. For IP-based partnerships, only IP addresses that are configured for Remote Copy are discovered.

For example, the first discovery occurs, and then the user runs the **`mkippartnership`** command in the command-line interface (CLI). Later discoveries might occur as a result of user activities (configuration changes), or as a result of hardware failures (for example, node failure and port failure).

**Remote Copy port group** This term indicates the settings of local and remote Ethernet ports (on local and partnered Storwize V7000 systems) that can access each other through a long-distance IP link.

For a successful partnership to be established between two Storwize V7000 clusters, at least two ports must be in the same Remote Copy port group, one from the local cluster and one from the partner cluster. More than two ports from the same system in a group can exist to enable Transmission Control Protocol (TCP) connection failover in a local and partnered node or port failure.

**Remote Copy port group ID** This numeric value indicates to which group the port belongs. Zero (0) is used to indicate that a port is not used for Remote Copy. For two Storwize V7000 clusters to form a partnership, both clusters must have at least one port that is configured with the same group ID, and they must be accessible to each other.

**RC logic** RC logic is a bidirectional full-duplex data path between two SVC clusters that are Remote Copy partners. This path is between an IP address pair, one local and one remote.

An RC login carries Remote Copy traffic that consists of host WRITEs, background copy traffic during initial sync within a relationship, periodic updates in Global Mirror with changed volumes relationships, and so on.

**Path configuration** Path configuration is the act of setting up RC logins between two partnered Storwize V7000 systems. The selection of IP addresses to be used for RC logins is based on certain rules that are specified in the Preferred practices section. Most of those rules are driven by constraints and requirements from a vendor-supplied link management library.

A simple algorithm is run by each Storwize V7000 system to arrive at the list of RC logins that must be established. Local and remote Storwize V7000 clusters are expected to arrive at the same IP address pairs for RC login creation, even though they run the algorithm independently.

## Preferred practices

The following preferred practices are suggested for IP replication:

- Configure two physical links between sites for redundancy.
- Configure Ethernet ports that are dedicated for Remote Copy. Do not allow iSCSI host attach for these Ethernet ports.
- Configure remote copy port group IDs on both nodes for each physical link to survive node failover.
- A minimum of four nodes are required for dual redundant links to work across node failures. If a node failure occurs on a two-node system, one link is lost.
- Do not zone in two Storwize V7000 systems over FC/FCoE when an IP partnership exists.
- Configure CHAP secret-based authentication, if required.
- The maximum supported round-trip time between sites is 80 ms for a 1 Gbps link.
- The maximum supported round-trip time between sites is 10 ms for a 10 Gbps link.

- For IP partnerships, the suggested method of copying is Global Mirror with changed volumes, because of the performance benefits. Also, Global Mirror and Metro Mirror might be more susceptible to the loss of synchronization.
- The amount of inter-cluster heartbeat traffic is 1 megabits per second (Mbps) per link.
- The minimum bandwidth requirement for the inter-cluster link is 10 Mbps. However, this bandwidth scales up with the amount of host I/O that you choose to use.

For more information, see *IBM SAN Volume Controller and Storwize Family Native IP Replication*, REDP-5103:

http://www.redbooks.ibm.com/abstracts/redp5103.html

## 3.3.8 Planning for external storage virtualization

This chapter describes how to plan for virtualizing external storage with IBM Storwize V7000. Virtualizing the storage infrastructure with IBM Storwize V7000 makes your storage environment more flexible, cost-effective, and easy to manage. The combination of IBM Storwize V7000 and an external storage system provides more storage capacity benefits from the powerful software functions within the IBM Storwize V7000.

The external storage systems that are incorporated into the IBM Storwize V7000 environment can be new or existing systems. The data on existing storage systems can be easily migrated to the IBM Storwize V7000 managed environment, as described in Chapter 9, “Storage migration” on page 347.

Back-end storage subsystem configuration planning must be applied to all storage controllers that are attached to the Storwize V7000.

For more information about supported storage subsystems, see this website:

http://www.ibm.com/support/docview.wss?uid=ssg1S1009559

Apply the following general guidelines for back-end storage subsystem configuration planning:

- In the SAN, storage controllers that are used by the Storwize V7000 clustered system must be connected through SAN switches. Direct connection between the Storwize V7000 and the storage controller is not supported.
- Multiple connections are allowed from the redundant controllers in the disk subsystem to improve data bandwidth performance. It is not mandatory to have a connection from each redundant controller in the disk subsystem to each counterpart SAN, but it is a preferred practice. Therefore, canister A in a Storwize V3700 subsystem can be connected to SAN A only, or to SAN A and SAN B. Also, canister B in the Storwize V3700 subsystem can be connected to SAN B only, or to SAN B and SAN A.
- HyperSwap configurations are supported by certain rules and configuration guidelines.

### Storwize V7000 zones

The switch fabric must be zoned so that the Storwize V7000 nodes can detect the back-end storage systems and the front-end host bus adapters (HBAs). Typically, the front-end host HBAs and the back-end storage systems are not in the same zone. The exception to this is where split host and split storage system configuration is in use.

All nodes in a system must be able to detect the same ports on each back-end storage system. Operation in a mode where two nodes detect a different set of ports on the same storage system is degraded, and the system logs errors that request a repair action. This can occur if inappropriate zoning is applied to the fabric or if inappropriate logical unit number (LUN) masking is used.

This rule has important implications for back-end storage systems that impose exclusive rules for mappings between HBA worldwide node names (WWNNs) and storage partitions.

Each Storwize V7000 port must be zoned so that it can be used for internode communications. When configuring switch zoning, you can zone some Storwize V7000 node ports to a host or to back-end storage systems.

When configuring zones for communication between nodes in the same system, the minimum configuration requires that all Fibre Channel ports on a node detect at least one Fibre Channel port on each other node in the same system. You cannot reduce the configuration in this environment.

It is critical that you configure storage systems and the storage area network (SAN) so that a system cannot access logical units (LUs) that a host or another system can also access. You can achieve this configuration with storage system LUN mapping and masking. If a node can detect a storage system through multiple paths, use zoning to restrict communication to those paths that do not travel over inter-switch links (ISLs).

If you do not have a storage subsystem that supports the Storwize V7000 round-robin algorithm, ensure that the number of MDisks per storage pool is a multiple of the number of storage ports that are available. This approach ensures sufficient bandwidth to the storage controller, and an even balance across storage controller ports.

## Host zones

The configuration rules for host zones are different depending upon the number of hosts that access the system. For configurations of fewer than 64 hosts per system, Storwize V7000 supports a simple set of zoning rules that enable a small set of host zones to be created for different environments. For configurations of more than 64 hosts per system, Storwize V7000 supports a more restrictive set of host zoning rules. These rules apply for both Fibre Channel (FC) and Fibre Channel over Ethernet (FCoE) connectivity.

Zoning that contains host HBAs must ensure host HBAs in dissimilar hosts or dissimilar HBAs are in separate zones. *Dissimilar* hosts means that the hosts are running different operating systems or are different hardware platforms. Therefore, different levels of the same operating system are regarded as similar.

To obtain the best overall performance of the system and to prevent overloading, the workload to each Storwize V7000 port must be equal. This can typically involve zoning approximately the same number of host Fibre Channel ports to each Storwize V7000 Fibre Channel port.

### *Systems with fewer than 64 hosts*

For systems with fewer than 64 hosts that are attached, zones that contain host HBAs must contain no more than 40 initiators, including the Storwize V7000 ports that act as initiators. A configuration that exceeds 40 initiators is not supported. A valid zone can be 32 host ports plus 8 Storwize V7000 ports.

When it is possible, place each HBA port in a host that connects to a node into a separate zone. Include exactly one port from each node in the I/O groups that are associated with this host. This type of host zoning is not mandatory, but is preferred for smaller configurations.

**Important:** If the switch vendor recommends fewer ports per zone for a particular SAN, the rules that are imposed by the vendor take precedence over Storwize V7000 rules.

To obtain the best performance from a host with multiple FC ports, the zoning must ensure that each FC port of a host is zoned with a different group of Storwize V7000 ports.

### *Systems with more than 64 hosts*

Each HBA port must be in a separate zone, and each zone must contain exactly one port from each Storwize V7000 node in each I/O group that the host accesses. A host can be associated with more than one I/O group, and therefore can access volumes from different I/O groups in a SAN. However, this reduces the maximum number of hosts that can be used in the SAN.

For example, if the same host uses volumes in two different I/O groups, this expends one of the 256 iSCSI hosts in each I/O group, or one of the 512 FC, FCoE, or serial-attached SCSI (SAS) hosts in each I/O group. If each host accesses volumes in every I/O group, there can be only 256 iSCSI hosts, or 512 FC, FCoE, or SAS hosts, in the configuration, as shown in Figure 3-7.

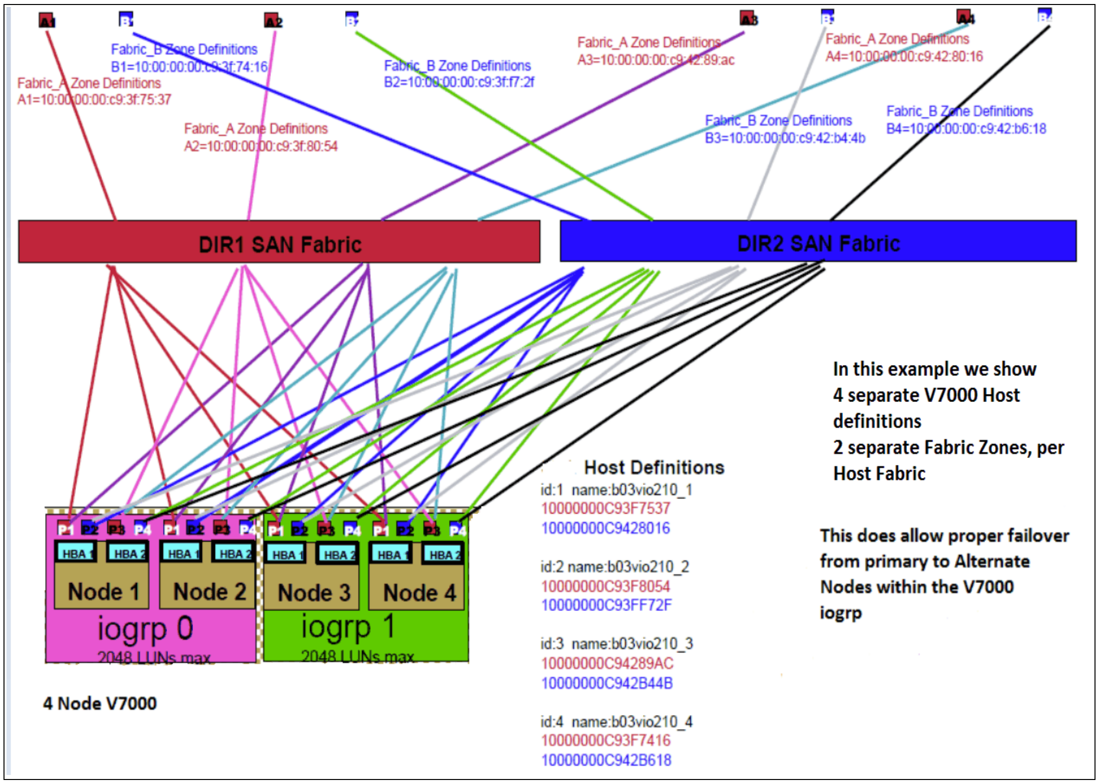

A1
B1
A2
B2
A3
B3
A4
B4
Fabric_B Zone Definitions
B1=10:00:00:00:c9:3f:74:16
Fabric_A Zone Definitions
A1=10:00:00:00:c9:3f:75:37
Fabric_A Zone Definitions
A2=10:00:00:00:c9:3f:80:54
Fabric_B Zone Definitions
B2=10:00:00:00:c9:3f:f7:2f
Fabric_A Zone Definitions
A3=10:00:00:00:c9:42:89:ac
Fabric_A Zone Definitions
A4=10:00:00:00:c9:42:80:16
Fabric_B Zone Definitions
B3=10:00:00:00:c9:42:b4:4b
Fabric_B Zone Definitions
B4=10:00:00:00:c9:42:b6:18
DIR1 SAN Fabric
DIR2 SAN Fabric
P1 P2 P3 P4
HBA 1 HBA 2
Node 1
Node 2
Node 3
Node 4
iogrp 0
2048 LUNs max
iogrp 1
2048 LUNs max
4 Node V7000
Host Definitions
id:1 name:b03vio210_1
10000000C93F7537
10000000C9428016
id:2 name:b03vio210_2
10000000C93F8054
10000000C93FF72F
id:3 name:b03vio210_3
10000000C94289AC
10000000C942B44B
id:4 name:b03vio210_4
10000000C93F7416
10000000C942B618
In this example we show
4 separate V7000 Host
definitions
2 separate Fabric Zones, per
Host Fabric
This does allow proper failover
from primary to Alternate
Nodes within the V7000
iogrp

*Figure 3-7 Storwize V7000 Zoning to multiple Host HBAs*

## 3.3.9 Storwize V7000 clustered system configuration

To ensure high availability in Storwize V7000 installations, consider the following guidelines when you design a SAN with the Storwize V7000:

- All nodes in a clustered system must be in the same LAN segment, because the nodes in the clustered system must assume the same clustered system or service IP address. Ensure that the network configuration allows any of the nodes to use these IP addresses. If you plan to use the second Ethernet port on each node, two LAN segments can be used. However, port 1 of every node must be in one LAN segment, and port 2 of every node must be in the other LAN segment.

- To maintain application uptime in the unlikely event of an individual Storwize V7000 node failing, Storwize V7000 nodes are always deployed in pairs (I/O Groups). If a node fails or is removed from the configuration, the remaining node operates in a degraded mode, but it is still a valid configuration. The remaining node operates in write-through mode, which means that the data is written directly to the disk subsystem (the cache is disabled for write I/O).
- The FC SAN connections between the Storwize V7000 node and the switches are optical fiber. These connections can run at 2 Gbps, 4 Gbps, 8 Gbps, or 16 Gbps, depending on your Storwize V7000 and switch hardware. The Storwize V7000 nodes auto-negotiate the connection speed with the switch.
- The Storwize V7000 node ports must be connected to the FC fabric only. Direct connections between the SVC and the host, or the disk subsystem, are unsupported.
- Two Storwize V7000 clustered systems cannot have access to the same LUNs within a disk subsystem. Configuring zoning so that two Storwize V7000 clustered systems have access to the same LUNs (MDisks) will likely result in data corruption.
- The Storwize V7000 uses three MDisks as quorum disks or three internal drives for the clustered system. A preferred practice for redundancy is to have each quorum disk in a separate storage subsystem, where possible. The current locations of the quorum disks can be displayed by using the **`lsquorum`** command, and relocated by using the **`chquorum`** command.

## 3.3.10 Storage pool configuration

The storage pool is at the center of the many-to-many relationship between the MDisks and the volumes. It acts as a container from which managed disks contribute chunks of physical disk capacity that are known as *extents*, and from which volumes are created.

MDisks in the Storwize V7000 are LUNs that are assigned from the underlying disk subsystems or internal arrays to the Storwize V7000, and can be managed or unmanaged. A managed MDisk is an MDisk that is assigned to a storage pool. Consider the following points:

- A storage pool is a collection of MDisks. An MDisk can be contained only within a single storage pool.
- Since software V7.5, the Storwize V7000 supports up to 1024 storage pools.
- The number of volumes that can be allocated from a storage pool is unlimited. However, with 7.8 an I/O Group is limited to 10K, and the clustered system limit is 10K.
- Volumes are associated with a single storage pool, except in cases where a volume is being migrated or mirrored between storage pools.

The Storwize V7000 supports extent sizes of 16 mebibytes (MiB), 32 MiB, 64 MiB, 128 MiB, 256 MiB, 512 MiB, 1024 MiB, 2048 MiB, 4096 MiB, and 8192 MiB. The extent size is a property of the storage pool and is set when the storage pool is created. All MDisks in the storage pool have the same extent size, and all volumes that are allocated from the storage pool have the same extent size.

The extent size of a storage pool cannot be changed. If you want another extent size, the storage pool must be deleted and a new storage pool configured.

Table 3-2 lists all of the available extent sizes in a Storwize V7000.

*Table 3-2 Extent size and total storage capacities per system*

| Extent size (MiB) | Total storage capacity manageable per system |
|---|---|
| 16 | 64 tebibytes (TiB) |
| 32 | 128 TiB |
| 64 | 256 TiB |
| 128 | 512 TiB |
| 256 | 1 pebibytes (PiB) |
| 512 | 2 PiB |
| 1024 | 4 PiB |
| 2048 | 8 PiB |
| 4096 | 16 PiB |
| 8192 | 32 PiB |

Consider the following information about storage pools:

- Maximum clustered system capacity is related to the extent size:
  - 16 MiB extent = 64 TiB and doubles for each increment in extent size, for example, 32 MiB = 128 TiB. We strongly advise 128 MiB - 256 MiB. The IBM Storage Performance Council (SPC) benchmarks used a 256 MiB extent.
  - Pick the extent size and then use that size for all storage pools.
  - You cannot migrate volumes between storage pools with separate extent sizes. However, you can use volume mirroring to create copies between storage pools with separate extent sizes.
- Storage pool reliability, availability, and serviceability (RAS) considerations:
  - It might make sense to create multiple storage pools; however, you must ensure that a host only gets volumes that are built from one storage pool. If the storage pool goes offline, it affects only a subset of all of the hosts that make up the Storwize V7000. However, creating multiple storage pools can cause a high number of storage pools, which can approach the Storwize V7000 limits.
  - If you do not isolate hosts to storage pools, create one large storage pool. Creating one large storage pool assumes that the physical disks are all the same size, speed, and RAID level.
  - The storage pool goes offline if an MDisk is unavailable, even if the MDisk has no data on it. Do not put MDisks into a storage pool until they are needed.
  - Create at least one separate storage pool for all of the image mode volumes.
  - Ensure that all of the host-persistent reserves are removed from LUNs that are given to the Storwize V7000.
- Storage pool performance considerations

  It might make sense to create multiple storage pools if you are attempting to isolate workloads to separate disk spindles. Storage pools with too few MDisks cause an MDisk overload, so having more spindle counts in a storage pool is better to meet workload requirements.
- The storage pool and Storwize V7000 cache relationship

The Storwize V7000 employs cache partitioning to limit the potentially negative effect that a poorly performing storage controller can have on the clustered system. The partition allocation size is based on the number of configured storage pools. This design protects against individual controller overloading, and against failures from using write cache and degrading the performance of the other storage pools in the clustered system.

Table 3-3 shows the limit of the write-cache data.

*Table 3-3 Limit of the cache data*

| Number of storage pools | Upper limit |
|---|---|
| 1 | 100% |
| 2 | 66% |
| 3 | 40% |
| 4 | 30% |
| 5 or more | 25% |

Consider the rule that no single partition can occupy more than its upper limit of cache capacity with write data. These limits are upper limits, and they are the points at which the Storwize V7000 cache starts to limit incoming I/O rates for volumes that are created from the storage pool. If a particular partition reaches this upper limit, the net result is the same as a global cache resource that is full. That is, the host writes are serviced on a one-out-one-in basis because the cache destages writes to the back-end disks.

However, only writes that are targeted at the full partition are limited. All I/O that is destined for other (non-limited) storage pools continues as normal. The read I/O requests for the limited partition also continue normally. However, because the Storwize V7000 is destaging write data at a rate that is greater than the controller can sustain (otherwise, the partition does not reach the upper limit), read response times are also likely affected.

## 3.3.11 Volume configuration

An individual volume is a member of one storage pool and one I/O Group. When a volume is created, you first identify the performance that you want, availability needs, and cost requirements for that volume, and then select the storage pool accordingly.

The storage pool defines which MDisks that are provided by the disk subsystem make up the volume. The I/O Group, which is made up of two nodes, defines which Storwize V7000 nodes provide I/O access to the volume.

**Important:** No fixed relationship exists between I/O Groups and storage pools.

Perform volume allocation that is based on the following considerations:

- Optimize performance between the hosts and the Storwize V7000 by attempting to distribute volumes evenly across available I/O Groups and nodes within the clustered system.
- Reach the level of performance, reliability, and capacity that you require by using the storage pool that corresponds to your needs. (You can access any storage pool from any node.) Choose the storage pool that fulfills the demands for your volumes regarding performance, reliability, and capacity.

- I/O Group considerations:
  - When you create a volume, it is associated with one node of an I/O Group. By default, when you create a volume, it is associated with the next node by using a round-robin algorithm. You can specify a *preferred access node*, which is the node through which you send I/O to the volume rather than by using the round-robin algorithm. A volume is defined for an I/O Group.
  - Even if you have eight paths for each volume, all I/O traffic flows only toward one node (the preferred node). Therefore, only four paths are used by the IBM Subsystem Device Driver (SDD). The other four paths are used only in a failure of the preferred node, or when a concurrent code upgrade is running.
- Creating image mode volumes:
  - Use image mode volumes when an MDisk already has data on it from a non-virtualized disk subsystem. When an image mode volume is created, it directly corresponds to the MDisk from which it is created. Therefore, volume logical block address (LBA) *x* = MDisk LBA *x*. The capacity of image mode volumes defaults to the capacity of the supplied MDisk.
  - When you create an image mode disk, the MDisk must have a mode of *unmanaged*, which does not belong to any storage pool. (A capacity of 0 is not allowed.) Image mode volumes can be created in sizes with a minimum granularity of 512 bytes, and they must be at least one block (512 bytes) in size.
- Creating managed mode volumes with sequential or striped policy

  When a managed mode volume with sequential or striped policy is created, you must use several MDisks that contain extents that are free and equal to or greater than the size of the volume that you want to create. Sufficient extents might be available on the MDisk, but a contiguous block that is large enough to satisfy the request might not be available.
- Thin-provisioned volume considerations:
  - When the thin-provisioned volume is created, you must understand the usage patterns by the applications or group users that are accessing this volume. You also must consider the actual size of the data, the rate of data creation, and modifying or deleting existing data.
  - The following operating modes for thin-provisioned volumes are available:
    - *Autoexpand volumes* allocate storage from a storage pool on demand with minimal required user intervention. However, a malfunctioning application can cause a volume to expand until it uses all of the storage in a storage pool.
    - *Non-autoexpand volumes* have a fixed amount of assigned storage. In this case, the user must monitor the volume and assign more capacity when required. A malfunctioning application can cause only the volume that it uses to fill up.
  - Depending on the initial size for the real capacity, the grain size and a warning level can be set. If a volume goes offline through a lack of available physical storage for autoexpand, or because a volume that is marked as non-autoexpand was not expanded in time, a danger exists of data being left in the cache until storage is made available. This situation is not a data integrity or data loss issue, but you must not rely on the Storwize V7000 cache as a backup storage mechanism.

> **Important:** Keep a warning level on the used capacity so that it provides adequate time to respond and provision more physical capacity.
>
> Warnings must not be ignored by an administrator.
>
> Use the autoexpand feature of the thin-provisioned volumes.

  - When you create a thin-provisioned volume, you can choose the grain size for allocating space in 32 kibibytes (KiB), 64 KiB, 128 KiB, or 256 KiB chunks. The grain size that you select affects the maximum virtual capacity for the thin-provisioned volume. The default grain size is 256 KiB, which is the advised option. If you select 32 KiB for the grain size, the volume size cannot exceed 260,000 gibibytes (GiB). The grain size cannot be changed after the thin-provisioned volume is created.

    Generally, smaller grain sizes save space but require more metadata access, which can adversely affect performance. If you are not going to use the thin-provisioned volume as a FlashCopy source or target volume, use 256 KiB to maximize performance. If you are going to use the thin-provisioned volume as a FlashCopy source or target volume, specify the same grain size for the volume and for the FlashCopy function.
  - Thin-provisioned volumes require more I/Os because of directory accesses. For truly random workloads with 70% read and 30% write, a thin-provisioned volume requires approximately one directory I/O for every user I/O.
  - The directory is two-way write-back-cached (as with the Storwize V7000 fastwrite cache), so certain applications perform better.
  - Thin-provisioned volumes require more processor processing, so the performance per I/O Group can also be reduced.
  - A thin-provisioned volume feature that is called *zero detect* provides clients with the ability to reclaim unused allocated disk space (zeros) when they are converting a fully allocated volume to a thin-provisioned volume by using volume mirroring.
- Volume mirroring guidelines:
  - Create or identify two separate storage pools to allocate space for your mirrored volume.
  - Allocate the storage pools that contain the mirrors from separate storage controllers.
  - If possible, use a storage pool with MDisks that share characteristics. Otherwise, the volume performance can be affected by the poorer performing MDisk.

### 3.3.12 Host attachment planning

Host attachment to Storwize V7000 through FC must be made through a SAN fabric, because direct host attachment to Storwize V7000 nodes is not supported. For Storwize V7000 configurations, using two redundant SAN fabrics is a preferred practice. Therefore, we advise that you have each host equipped with a minimum of two host bus adapters (HBAs), or at least a dual-port HBA with each HBA connected to a SAN switch in either fabric.

Storwize V7000 imposes no particular limit on the actual distance between the Storwize V7000 nodes and host servers. Therefore, a server can be attached to an edge switch in a core-edge configuration and the Storwize V7000 cluster is at the core of the fabric.

For host attachment, the Storwize V7000 supports up to three inter-switch link (ISL) hops in the fabric, which means that the server to the Storwize V7000 can be separated by up to five

FC links, four of which can be 10 km long (6.2 miles) if longwave small form factor pluggables (SFPs) are used.

The Storwize V7000 nodes contain shortwave SFPs, so they must be within the allowed distance depending on the speed of the switch to which they attach. Therefore, the configuration that is shown in Figure 3-8 is supported.

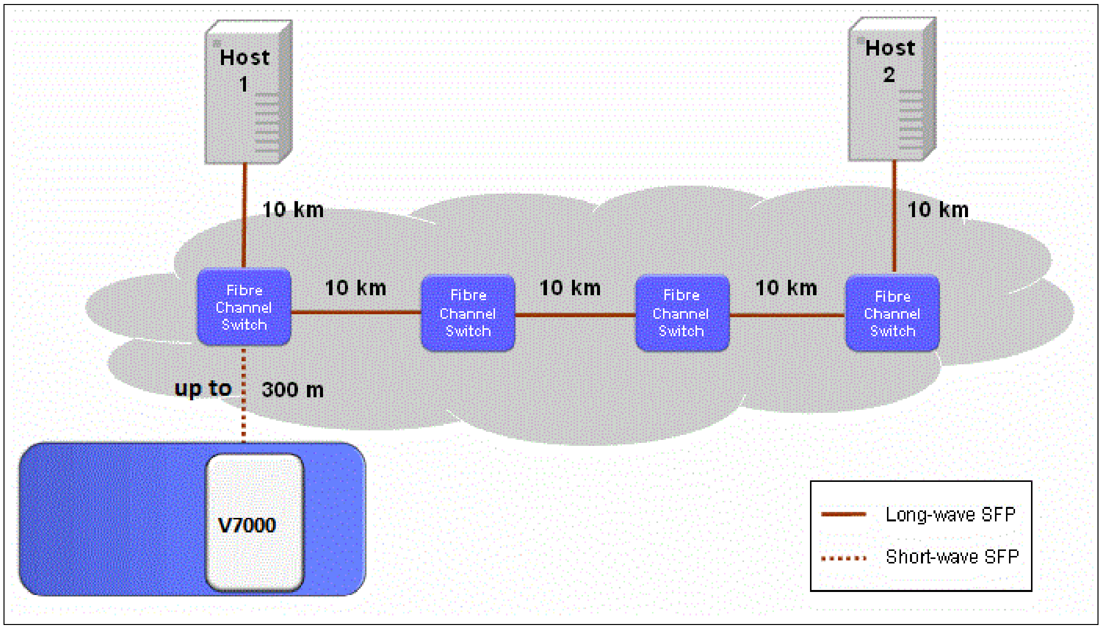


*Figure 3-8 Example of host connectivity*

In Figure 3-8, the optical distance between Storwize V7000 Node and Host 2 is slightly over 40 km (24.85 miles).

To avoid latencies that lead to degraded performance, we suggest that you avoid ISL hops whenever possible. In an optimal setup, the servers connect to the same SAN switch as the Storwize V7000 nodes.

**Note:** Prior to attaching host systems to Storwize V7000, refer to the Configuration Limits and Restrictions for IBM Storwize V7000 described in:

http://www.ibm.com/support/docview.wss?uid=ssg1S1009561

Access from a server to a Storwize V7000 cluster through the SAN fabric is defined by using switch zoning.

Consider the following rules for zoning hosts with the Storwize V7000:

- Homogeneous HBA port zones

  Switch zones that contain HBAs must contain HBAs from similar host types and similar HBAs in the same host. For example, IBM AIX and Microsoft Windows hosts must be in separate zones, and QLogic and Emulex adapters must also be in separate zones.

  **Important:** A configuration that breaches this rule is unsupported, because it can introduce instability to the environment.

- HBA to Storwize V7000 port zones

  Place each host's HBA in a separate zone with one or two Storwize V7000 ports. If two ports exist, use one from each node in the I/O Group. Do not place more than two Storwize V7000 ports in a zone with an HBA, because this design results in more than the recommended number of paths, as seen from the host multipath driver.

  > **Number of paths:** For *n* + 1 redundancy, use the following number of paths:
  >
  > - With two HBA ports, zone HBA ports to Storwize V7000 ports 1:2 for a total of four paths.
  > - With four HBA ports, zone HBA ports to Storwize V7000 ports 1:1 for a total of four paths.
  >
  > Optional (*n*+2 redundancy): With four HBA ports, zone HBA ports to Storwize V7000 ports 1:2 for a total of eight paths.
  >
  > Here, we use the term *HBA port* to describe the SCSI initiator and *V7000 port* to describe the SCSI target.

- Maximum host paths per logical unit (LU)

  For any volume, the number of paths through the SAN from the Storwize V7000 nodes to a host must not exceed eight. For most configurations, four paths to an I/O Group (four paths to each volume that is provided by this I/O Group) are sufficient.

  > **Important:** The maximum number of host paths per logical unit number (LUN) must not exceed eight.

- Balanced host load across HBA ports

  To obtain the best performance from a host with multiple ports, ensure that each host port is zoned with a separate group of Storwize V7000 ports.

- Balanced host load across Storwize V7000 ports

To obtain the best overall performance of the subsystem and to prevent overloading, the workload to each Storwize V7000 port must be equal. You can achieve this balance by zoning approximately the same number of host ports to each Storwize V7000 port.

Figure 3-9 on page 71 shows an overview of a configuration where servers contain two single-port HBAs each, and the configuration includes the following characteristics:

- Distribute the attached hosts equally between two logical sets per I/O Group, if possible. Connect hosts from each set to the same group of Storwize V7000 ports. This *port group* includes exactly one port from each Storwize V7000 node in the I/O Group. The zoning defines the correct connections.
- The port groups are defined in the following manner:
  - Hosts in host set one of an I/O Group are always zoned to the P1 and P4 ports on both nodes, for example, N1/N2 of I/O Group zero.
  - Hosts in host set two of an I/O Group are always zoned to the P2 and P3 ports on both nodes of an I/O Group.
- You can create aliases for these port groups (per I/O Group):
  - Fabric A: IOGRP0_PG1 → N1_P1;N2_P1,IOGRP0_PG2 → N1_P3;N2_P3
  - Fabric B: IOGRP0_PG1 → N1_P4;N2_P4,IOGRP0_PG2 → N1_P2;N2_P2

- Create host zones by always using the host port WWPN and the PG1 alias for hosts in the first host set. Always use the host port WWPN and the PG2 alias for hosts from the second host set. If a host must be zoned to multiple I/O Groups, add the PG1 or PG2 aliases from the specific I/O Groups to the host zone.

The use of this schema provides four paths to one I/O Group for each host, and helps to maintain an equal distribution of host connections on Storwize V7000 ports.

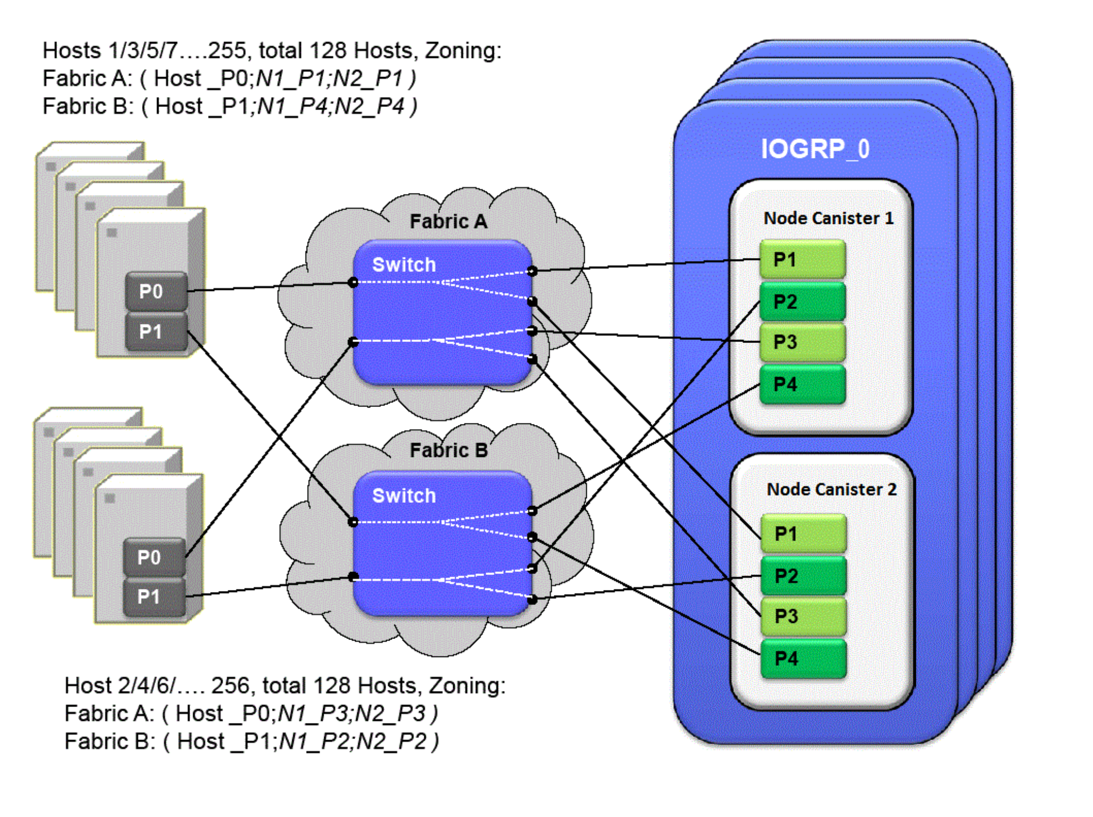

Hosts 1/3/5/7….255, total 128 Hosts, Zoning:
Fabric A: ( Host _P0;*N1_P1;N2_P1* )
Fabric B: ( Host _P1;*N1_P4;N2_P4* )
P0
P1
P0
P1
Fabric A
Switch
Fabric B
Switch
IOGRP_0
Node Canister 1
P1
P2
P3
P4
Node Canister 2
P1
P2
P3
P4
Host 2/4/6/…. 256, total 128 Hosts, Zoning:
Fabric A: ( Host _P0;*N1_P3;N2_P3* )
Fabric B: ( Host _P1;*N1_P2;N2_P2* )

*Figure 3-9 Overview of four-path host zoning*

When possible, use the minimum number of paths that are necessary to achieve a sufficient level of redundancy. For the Storwize V7000 environment, no more than four paths per I/O Group are required to accomplish this layout.

All paths must be managed by the multipath driver on the host side. If we assume that a server is connected through four ports to the Storwize V7000, each volume is seen through eight paths. With 125 volumes mapped to this server, the multipath driver must support handling up to 1000 active paths (8 x 125).

For more configuration and operational information about the IBM Subsystem Device Driver (SDD), see the *Multipath Subsystem Device Driver User's Guide*, S7000303:

http://www.ibm.com/support/docview.wss?uid=ssg1S7000303

For hosts that use four HBAs/ports with eight connections to an I/O Group, use the zoning schema that is shown in Figure 3-10. You can combine this schema with the previous four-path zoning schema.

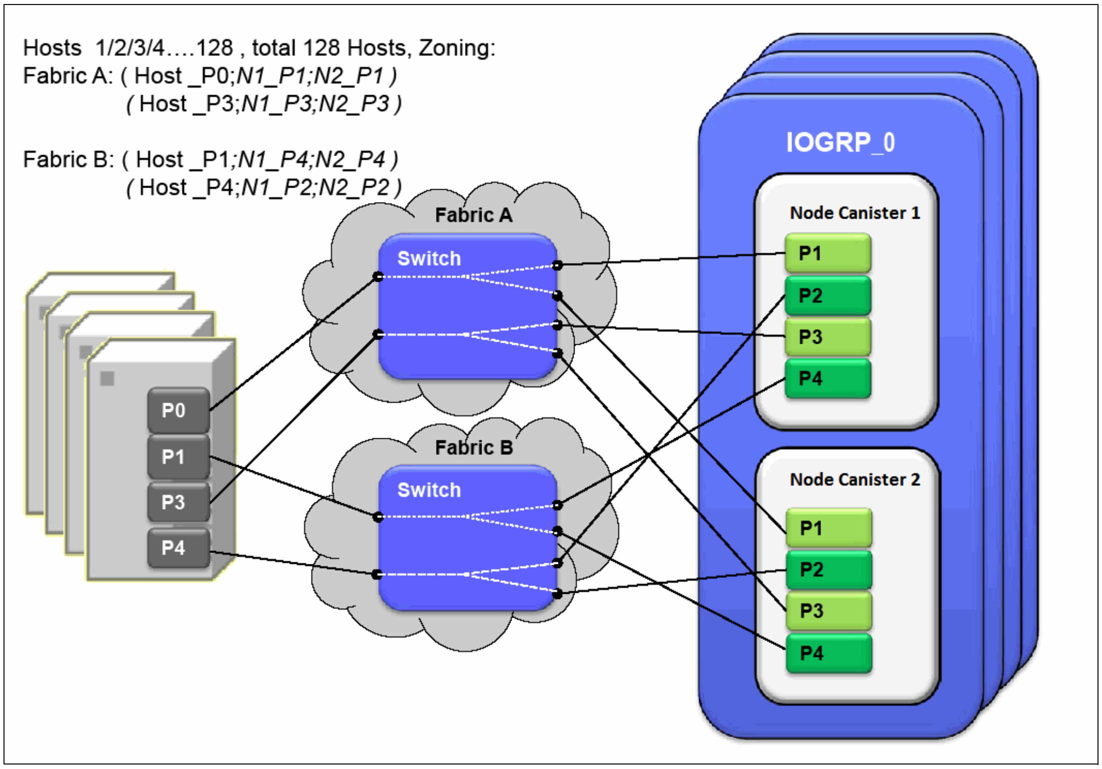


*Figure 3-10 Overview of eight-path host zoning*

For more information see Chapter 8, “Hosts” on page 283.

### 3.3.13 Host mapping (LUN masking)

For the host and application servers, the following guidelines apply:

- Each Storwize V7000 node presents a volume to the SAN through four ports. Because two nodes are used in normal operations to provide redundant paths to the same storage, a host with two HBAs can see multiple paths to each LUN that is presented by the Storwize V7000. Use zoning to limit the pathing from a minimum of two paths to the available maximum of eight paths, depending on the kind of HA and performance that you want in your configuration.

  It is best to use zoning to limit the pathing to four paths. The hosts must run a multipathing device driver to limit the pathing back to a single device. The multipathing driver that is supported and delivered by IBM Storwize V7000 is the IBM Subsystem Device Driver (SDD). Native Microsoft Multipath I/O (MPIO) drivers on selected hosts are supported.

  For more operating system-specific information about MPIO support, see this website:

  http://www.ibm.com/support/docview.wss?uid=ssg1S1009559

  The actual version of the Subsystem Device Driver Device Specific Module (SDDDSM) for IBM products is available at this website:

  http://www.ibm.com/support/docview.wss?uid=ssg1S4000350

- The number of paths to a volume from a host to the nodes in the I/O Group that owns the volume must not exceed eight, even if eight is not the maximum number of paths that are supported by the multipath driver (SDD supports up to 32). To restrict the number of paths to a host volume, the fabrics must be zoned so that each host FC port is zoned to no more than two ports from each Storwize V7000 node in the I/O Group that owns the volume.

**Multipathing:** We suggest the following number of paths per volume (*n*+1 redundancy):

- With two HBA ports, zone the HBA ports to the Storwize V7000 ports 1:2 for a total of four paths.
- With four HBA ports, zone the HBA ports to the Storwize V7000 ports 1:1 for a total of four paths.
- Optional (*n*+2 redundancy): With four HBA ports, zone the HBA ports to the Storwize V7000 ports 1:2 for a total of eight paths.

We use the term *HBA port* to describe the *SCSI Initiator*. We use the term *V7000 port* to describe the *SCSI target*.

The maximum number of host paths per volume must not exceed eight.

- If a host has multiple HBA ports, each port must be zoned to a separate set of Storwize V7000 ports to maximize high availability and performance.
- To configure greater than 256 hosts, you must configure the host to I/O Group mappings on the Storwize V7000. Each I/O Group can contain a maximum of 512 hosts, so it is possible to create 2048 host objects on an eight-node Storwize V7000 clustered system. Volumes can be mapped only to a host that is associated with the I/O Group to which the volume belongs.
- Port masking:

  You can use a *port mask* to control the node target ports that a host can access, which satisfies the following requirements:

  – As part of a security policy to limit the set of WWPNs that can obtain access to any volumes through a Storwize V7000 port

  – As part of a scheme to limit the number of logins with mapped volumes visible to a host multipathing driver, such as SDD, and therefore limit the number of host objects that are configured without resorting to switch zoning
- The port mask is an optional parameter of the **`mkhost`** and **`chhost`** commands. The port mask is four binary bits. Valid mask values range from `0000` (no ports enabled) to `1111` (all ports enabled). For example, a mask of `0011` enables port 1 and port 2. The default value is `1111` (all ports enabled).
- The Storwize V7000 supports connection to the Cisco MDS family and Brocade family. For more information, see this website:

  http://www.ibm.com/support/docview.wss?uid=ssg1S1009559

## 3.3.14 NPIV planning

For more information, see 8.3, “N-Port Virtualization ID (NPIV) Support” on page 285.

## 3.3.15 Advanced Copy Services

Storwize V7000 offers the following Advanced Copy Services:

- FlashCopy
- Metro Mirror
- Global Mirror

**Layers:** V6.3 introduced a new property that is called *layer* for the clustered system. This property is used when a copy services partnership exists between an SVC and an IBM Storwize V7000. There are two layers: *Replication* and *storage*. All SVC clustered systems are replication layers and cannot be changed. By default, the IBM Storwize V7000 is a storage layer, which must be changed by using the **chsystem** CLI command before you use it to make any copy services partnership with the SVC.

Figure 3-11 shows an example for replication and storage layers.

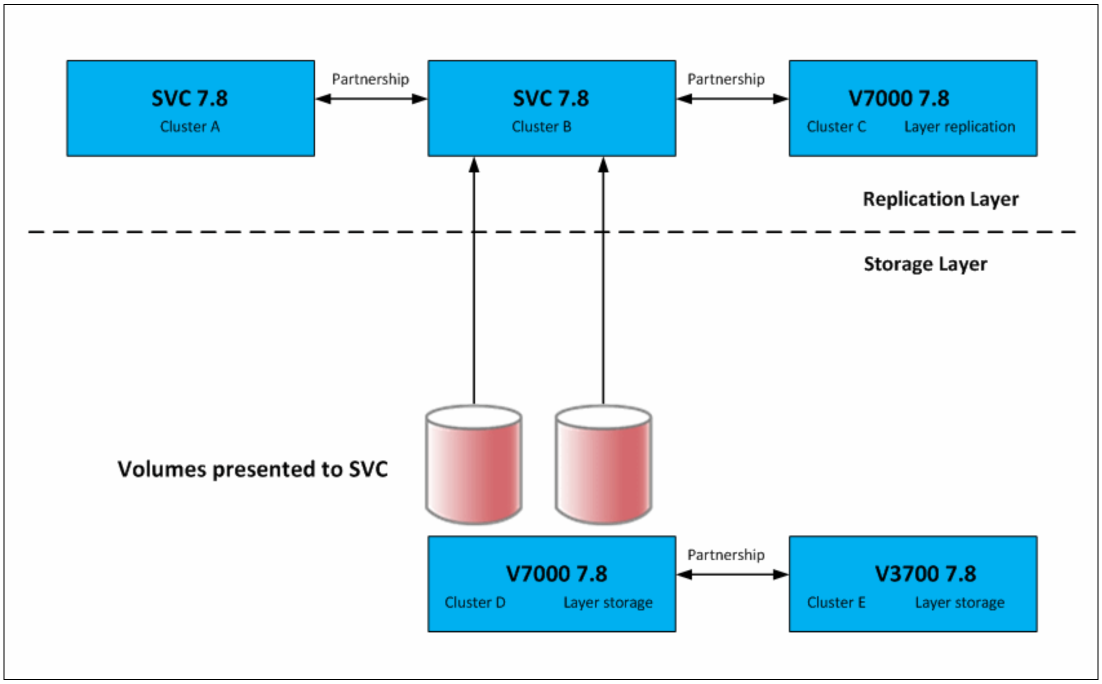


*Figure 3-11 Replication and storage layer*

SVC Advanced Copy Services must apply the guidelines that are described next.

### FlashCopy guidelines

Consider the following FlashCopy guidelines:

- Identify each application that must have a FlashCopy function implemented for its volume.
- FlashCopy is a relationship between volumes. Those volumes can belong to separate storage pools and separate storage subsystems.
- You can use FlashCopy for backup purposes by interacting with the Tivoli Storage Manager Agent, or for cloning a particular environment.
- Define which FlashCopy best fits your requirements:
  - No copy
  - Full copy
  - Thin-Provisioned
  - Incremental
- Define which FlashCopy rate best fits your requirement in terms of the performance and the amount of time to complete the FlashCopy. Table 3-4 shows the relationship of the background copy rate value to the attempted number of grains to be split per second.

- Define the grain size that you want to use. A *grain* is the unit of data that is represented by a single bit in the FlashCopy bitmap table. Larger grain sizes can cause a longer FlashCopy elapsed time and a higher space usage in the FlashCopy target volume. Smaller grain sizes can have the opposite effect. The data structure and the source data location can modify those effects.

  In an actual environment, check the results of your FlashCopy procedure in terms of the data that is copied at every run and in terms of elapsed time, comparing them to the new Storwize V7000 FlashCopy results. Eventually, adapt the grain per second and the copy rate parameter to fit your environment's requirements. See Table 3-4.

*Table 3-4 Grain splits per second*

| User percentage | Data copied per second | 256 KiB grain per second | 64 KiB grain per second |
|---:|---:|---:|---:|
| 1 - 10 | 128 KiB | 0.5 | 2 |
| 11 - 20 | 256 KiB | 1 | 4 |
| 21 - 30 | 512 KiB | 2 | 8 |
| 31 - 40 | 1 MiB | 4 | 16 |
| 41 - 50 | 2 MiB | 8 | 32 |
| 51 - 60 | 4 MiB | 16 | 64 |
| 61 - 70 | 8 MiB | 32 | 128 |
| 71 - 80 | 16 MiB | 64 | 256 |
| 81 - 90 | 32 MiB | 128 | 512 |
| 91 - 100 | 64 MiB | 256 | 1024 |

## Metro Mirror and Global Mirror guidelines

Storwize V7000 supports intracluster and intercluster Metro Mirror and Global Mirror. From the intracluster point of view, any single clustered system is a reasonable candidate for a Metro Mirror or Global Mirror operation. However, the intercluster operation needs at least two clustered systems that are separated by several moderately high-bandwidth links.

Technologies for extending the distance between two Storwize V7000 clustered systems can be broadly divided into the following categories:

- FC extenders
- SAN multiprotocol routers

Because of the more complex interactions that are involved, IBM explicitly tests products of this class for interoperability with the Storwize V7000. For more information about the current list of supported SAN routers in the supported hardware list, see this website:

http://www.ibm.com/support/docview.wss?uid=ssg1S1009559

IBM has tested several FC extenders and SAN router technologies with the Storwize V7000. You must plan, install, and test FC extenders and SAN router technologies with the Storwize V7000 so that the following requirements are met:

- The round-trip latency between sites must not exceed 80 - 250 ms. For Global Mirror, this limit allows a distance between the primary and secondary sites of up to 25,000 km (15,534.28 miles).

- If you use remote mirroring between systems with 80 - 250 ms round-trip latency, you must meet the following additional requirements:
  - There must be a Fibre Channel partnership between systems, not an IP partnership.
  - All systems in the partnership must have a minimum software level of 7.4.
  - The **`rcbuffersize`** setting must be set to 512 MB on each system in the partnership. This can be accomplished by running the **`chsystem -rcbuffersize 512`** command on each system. (Note that changing this setting is disruptive to Metro Mirror and Global Mirror operations. Use this command only before partnerships have been created between systems, or when all partnerships with the system have been stopped.).
  - Two Fibre Channel ports on each node that will be used for replication must be dedicated for replication traffic. This can be achieved using SAN zoning and port masking.
  - SAN zoning should be applied to provide separate intra-system zones for each local-remote I/O group pair that will be used for replication.
- The latency of long-distance links depends on the technology that is used to implement them. A point-to-point dark fiber-based link often provides a round-trip latency of 1 ms per 100 km (62.13 miles) or better. Other technologies provide longer round-trip latencies, which affect the maximum supported distance.
- The configuration must be tested with the expected peak workloads.
- When Metro Mirror or Global Mirror is used, a certain amount of bandwidth is required for the IBM Storwize V7000 intercluster heartbeat traffic. The amount of traffic depends on how many nodes are in each of the two clustered systems.

  Table 3-5 shows the amount of heartbeat traffic, in megabits per second, that is generated by various sizes of clustered systems.

*Table 3-5 Intersystem heartbeat traffic in Mbps*

| **Storwize V7000 System 1** | **Storwize V7000 System 2** | | | |
|---|---|---|---|---|
| | **2 nodes** | **4 nodes** | **6 nodes** | **8 nodes** |
| **2 nodes** | 5 | 6 | 6 | 6 |
| **4 nodes** | 6 | 10 | 11 | 12 |
| **6 nodes** | 6 | 11 | 16 | 17 |
| **8 nodes** | 6 | 12 | 17 | 21 |

- These numbers represent the total traffic between the two clustered systems when no I/O is taking place to mirrored volumes. Half of the data is sent by one clustered system, and half of the data is sent by the other clustered system. The traffic is divided evenly over all available intercluster links. Therefore, if you have two redundant links, half of this traffic is sent over each link during fault-free operation.
- The bandwidth between sites must, at the least, be sized to meet the peak workload requirements, in addition to maintaining the maximum latency that was specified previously. You must evaluate the peak workload requirement by considering the average write workload over a period of one minute or less, plus the required synchronization copy bandwidth.

  With no active synchronization copies and no write I/O disks in Metro Mirror or Global Mirror relationships, the Storwize V7000 protocols operate with the bandwidth that is indicated in Table 3-5 on page 76. However, you can determine the true bandwidth that is

required
for the link only by considering the peak write bandwidth to volumes that are participating in Metro Mirror or Global Mirror relationships. Then, add it to the peak synchronization copy bandwidth.

- If the link between the sites is configured with redundancy so that it can tolerate single failures, you must size the link so that the bandwidth and latency statements continue to be true, even during single failure conditions.
- The configuration is tested to simulate the failure of the primary site (to test the recovery capabilities and procedures), including eventual failback to the primary site from the secondary.
- The configuration must be tested to confirm that any failover mechanisms in the intercluster links interoperate satisfactorily with the Storwize V7000.
- The FC extender must be treated as a normal link.
- The bandwidth and latency measurements must be made by, or on behalf of, the client. They are not part of the standard installation of the Storwize V7000 by IBM. Make these measurements during installation and record the measurements. Testing must be repeated following any significant changes to the equipment that provides the intercluster link.

### Global Mirror guidelines

Consider the following guidelines:

- When Storwize V7000 Global Mirror is used, all components in the SAN must sustain the workload that is generated by application hosts and the Global Mirror background copy workload. Otherwise, Global Mirror can automatically stop your relationships to protect your application hosts from increased response times. Therefore, it is important to configure each component correctly.
- Use a SAN performance monitoring tool, such as IBM Tivoli Storage Productivity Center, which enables you to continuously monitor the SAN components for error conditions and performance problems. This tool helps you detect potential issues before they affect your disaster recovery solution.
- The long-distance link between the two clustered systems must be provisioned to allow for the peak application write workload to the Global Mirror source volumes and the client-defined level of background copy.
- The peak application write workload ideally must be determined by analyzing the Storwize V7000 performance statistics.
- Statistics must be gathered over a typical application I/O workload cycle, which might be days, weeks, or months, depending on the environment on which the Storwize V7000 is used. These statistics must be used to find the peak write workload that the link must support.
- Characteristics of the link can change with use. For example, latency can increase as the link is used to carry an increased bandwidth. The user must be aware of the link's behavior in such situations, and ensure that the link remains within the specified limits. If the characteristics are not known, testing must be performed to gain confidence of the link's suitability.
- Users of Global Mirror must consider how to optimize the performance of the long-distance link, which depends on the technology that is used to implement the link. For example, when you are transmitting FC traffic over an IP link, you might want to enable jumbo frames to improve efficiency.

- The use of Global Mirror and Metro Mirror between the same two clustered systems is supported.
- Support exists for cache-disabled volumes to participate in a Global Mirror relationship; however, this design is not a preferred practice.
- The **gmlinktolerance** parameter of the remote copy partnership must be set to an appropriate value. The default value is 300 seconds (5 minutes), which is appropriate for most clients.
- During SAN maintenance, the user must choose to perform one of the following actions:
  - Reduce the application I/O workload during maintenance (so that the degraded SAN components can manage the new workload)
  - Disable the **gmlinktolerance** feature
  - Increase the **gmlinktolerance** value (which means that application hosts might see extended response times from Global Mirror volumes)
  - Stop the Global Mirror relationships

  If the **gmlinktolerance** value is increased for maintenance lasting *x* minutes, it must be reset only to the normal value *x* minutes after the end of the maintenance activity.

  If **gmlinktolerance** is disabled during maintenance, it must be re-enabled after the maintenance is complete.
- Starting with software V7.6, you can use the **chsystem** command to set the maximum replication delay for the system. This value ensures that the single slow write operation does not affect the entire primary site.

  You can configure this delay for all relationships or consistency groups that exist on the system by using the **maxreplicationdelay** parameter on the **chsystem** command. This value indicates the amount of time (in seconds) that a host write operation can be outstanding before replication is stopped for a relationship on the system. If the system detects a delay in replication on a particular relationship or consistency group, only that relationship or consistency group is stopped.

  In systems with many relationships, a single slow relationship can cause delay for the remaining relationships on the system. This setting isolates the potential relationship with delays so that you can investigate the cause of these issues. When the maximum replication delay is reached, the system generates an error message that indicates the relationship that exceeded the maximum replication delay.
- Global Mirror volumes must have their preferred nodes evenly distributed between the nodes of the clustered systems. Each volume within an I/O Group has a preferred node property that can be used to balance the I/O load between nodes in that group.

  Figure 3-12 on page 79 shows the correct relationship between volumes in a Metro Mirror or Global Mirror solution.

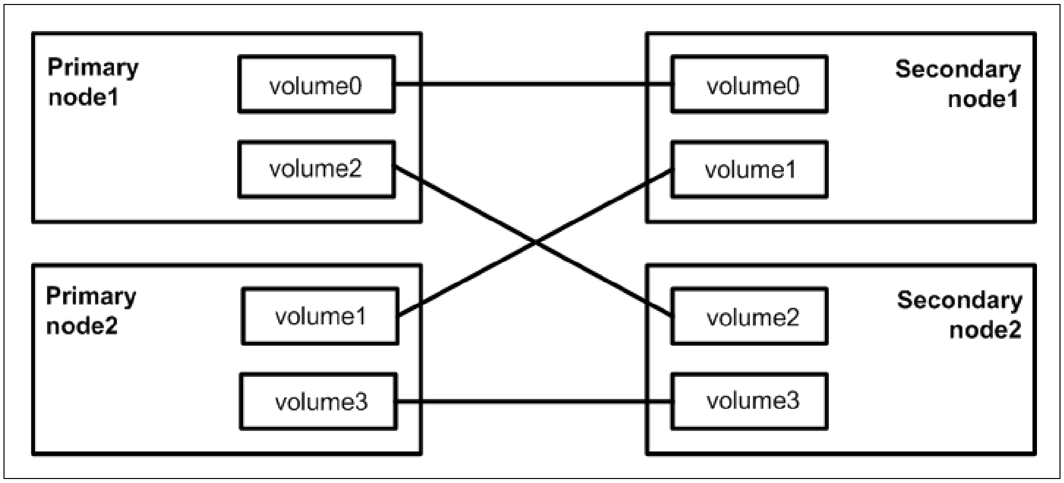


*Figure 3-12 Correct volume relationship*

- The capabilities of the storage controllers at the secondary clustered system must be provisioned to allow for the peak application workload to the Global Mirror volumes, plus the client-defined level of background copy, plus any other I/O being performed at the secondary site. The performance of applications at the primary clustered system can be limited by the performance of the back-end storage controllers at the secondary clustered system to maximize the amount of I/O that applications can perform to Global Mirror volumes.
- A complete review must be performed before Serial Advanced Technology Attachment (SATA) for Metro Mirror or Global Mirror secondary volumes is used. The use of a slower disk subsystem for the secondary volumes for high-performance primary volumes can mean that the Storwize V7000 cache might not be able to buffer all the writes, and flushing cache writes to SATA might slow I/O at the production site.
- Storage controllers must be configured to support the Global Mirror workload that is required of them. You can dedicate storage controllers to only Global Mirror volumes. You can also configure the controller to ensure sufficient quality of service (QoS) for the disks that are used by Global Mirror. Alternatively, you can ensure that physical disks are not shared between Global Mirror volumes and other I/O, for example, by not splitting an individual RAID array.
- MDisks within a Global Mirror storage pool must be similar in their characteristics, for example, RAID level, physical disk count, and disk speed. This requirement is true of all storage pools, but maintaining performance is important when Global Mirror is used.
- When a consistent relationship is stopped, for example, by a persistent I/O error on the intercluster link, the relationship enters the `consistent_stopped` state. I/O at the primary site continues, but the updates are not mirrored to the secondary site. Restarting the relationship begins the process of synchronizing new data to the secondary disk. While this synchronization is in progress, the relationship is in the `inconsistent_copying` state.

  Therefore, the Global Mirror secondary volume is not in a usable state until the copy completes and the relationship returns to a Consistent state. For this reason, it is highly advisable to create a FlashCopy of the secondary volume before the relationship is restarted. When started, the FlashCopy provides a consistent copy of the data, even while the Global Mirror relationship is copying.

  If the Global Mirror relationship does not reach the Synchronized state (for example, if the intercluster link experiences further persistent I/O errors), the FlashCopy target can be used at the secondary site for disaster recovery purposes.

- If you plan to use a Fibre Channel over IP (FCIP) intercluster link, it is important to design and size the pipe correctly.

  Example 3-2 shows a best-guess bandwidth sizing formula.

  *Example 3-2 Wide area network (WAN) link calculation example*

  ```
  Amount of write data within 24 hours times 4 to allow for peaks
  Translate into MB/s to determine WAN link needed
  Example:
  250 GB a day
  250 GB * 4 = 1 TB
  24 hours * 3600 secs/hr. = 86400 secs
  1,000,000,000,000/ 86400 = approximately 12 MB/s,
  Which means OC3 or higher is needed (155 Mbps or higher)
  ```

- If compression is available on routers or WAN communication devices, smaller pipelines might be adequate. Workload is probably not evenly spread across 24 hours. If extended periods of high data change rates exist, consider suspending Global Mirror during that time frame.
- If the network bandwidth is too small to handle the traffic, the application write I/O response times might be elongated. For the Storwize V7000, Global Mirror must support short-term “Peak Write” bandwidth requirements. Storwize V7000 Global Mirror is much more sensitive to a lack of bandwidth than the DS8000.
- You must also consider the initial sync and resync workload. The Global Mirror partnership’s background copy rate must be set to a value that is appropriate to the link and secondary back-end storage. The more bandwidth that you give to the sync and resync operation, the less workload can be delivered by the Storwize V7000 for the regular data traffic.
- Do not propose Global Mirror if the data change rate exceeds the communication bandwidth or if the round-trip latency exceeds 80 - 250 ms. A round-trip latency that is greater than 250 ms requires Storage Customer Opportunity REquest (SCORE)/request for price quotation (RPQ) submission.

## 3.3.16 SAN boot support

The IBM Storwite V7000 supports SAN boot or start-up for AIX, Microsoft Windows Server, and other operating systems. Because SAN boot support can change, check the following websites regularly:

http://www.ibm.com/support/docview.wss?uid=ssg1S1009559
http://www.ibm.com/systems/support/storage/ssic/interoperability.wss

For more detailed information see Appendix B, “CLI setup and SAN boot” on page 675.

## 3.3.17 Data migration from a non-virtualized storage subsystem

Data migration is an important part of a Storwize V7000 implementation. Therefore, you must accurately prepare a data migration plan. You might need to migrate your data for one of the following reasons:

- To redistribute workload within a clustered system across the disk subsystem
- To move workload onto newly installed storage
- To move workload off old or failing storage, ahead of decommissioning it
- To move workload to rebalance a changed workload

- To migrate data from an older disk subsystem to Storwize V7000-managed storage
- To migrate data from one disk subsystem to another disk subsystem

Because multiple data migration methods are available, choose the method that best fits your environment, operating system platform, type of data, and application’s service level agreement (SLA).

We can define data migration as belonging to the following groups:

- Based on operating system Logical Volume Manager (LVM) or commands
- Based on special data migration software
- Based on the Storwize V7000 data migration feature

With data migration, apply the following guidelines:

- Choose which data migration method best fits your operating system platform, type of data, and SLA.
- Check the following interoperability matrix for the storage subsystem to which your data is being migrated:

  http://www.ibm.com/support/docview.wss?uid=ssg1S1009559
- Choose where you want to place your data after migration in terms of the storage pools that relate to a specific storage subsystem tier.
- Check whether enough free space or extents are available in the target storage pool.
- Decide whether your data is critical and must be protected by a volume mirroring option or whether it must be replicated in a remote site for disaster recovery.
- Prepare offline all of the zone and LUN masking and host mappings that you might need to minimize downtime during the migration.
- Prepare a detailed operation plan so that you do not overlook anything at data migration time.
- Run a data backup before you start any data migration. Data backup must be part of the regular data management process.
- You might want to use the Storwize V7000 as a data mover to migrate data from a non-virtualized storage subsystem to another non-virtualized storage subsystem. In this case, you might have to add checks that relate to the specific storage subsystem that you want to migrate. Be careful when you are using slower disk subsystems for the secondary volumes for high-performance primary volumes because the Storwize V7000 cache might not be able to buffer all the writes and flushing cache writes to SATA might slow I/O at the production site.

### 3.3.18 Storwize V7000 configuration backup procedure

Save the configuration externally when changes, such as adding new nodes and disk subsystems, are performed on the clustered system. Saving the configuration is a crucial part of Storwize V7000 management, and various methods can be applied to back up your Storwize V7000 configuration. The preferred practice is to implement an automatic configuration backup by applying the configuration backup command.

For more information, see Chapter 13, “RAS, monitoring, and troubleshooting” on page 605.

## 3.4 Performance considerations

Although storage virtualization with the Storwize V7000 improves flexibility and provides simpler management of a storage infrastructure, it can also provide a substantial performance advantage for various workloads. The Storwize V7000 caching capability and its ability to stripe volumes across multiple disk arrays are the reasons why performance improvement is significant when it is implemented with midrange disk subsystems. This technology is often provided only with high-end enterprise disk subsystems.

> **Tip:** Technically, almost all storage controllers provide both striping (RAID 5 or RAID 10) and a form of caching. The real benefit is the degree to which you can stripe the data across all MDisks in a storage pool. Therefore, you have the maximum number of active spindles at one time. The caching is secondary. The Storwize V7000 provides additional caching to the caching that midrange controllers provide (usually a few GB). Enterprise systems have much larger caches.

To ensure the performance that you want and verify the capacity of your storage infrastructure, undertake a performance and capacity analysis to reveal the business requirements of your storage environment. When this analysis is done, you can use the guidelines in this chapter to design a solution that meets the business requirements.

When you are considering performance for a system, always identify the bottleneck and, therefore, the limiting factor of a specific system. You must also consider the component for whose workload you identify a limiting factor. The component might not be the same component that is identified as the limiting factor for other workloads.

When you are designing a storage infrastructure with the Storwize V7000 or implementing a Storwize V7000 in an existing storage infrastructure, you must consider the performance and capacity of the SAN and disk subsystems, and the known or expected workload.

### 3.4.1 SAN

The Storwize V7000 now has the following models:

- 2076-124 (Gen1)
- 2076-524 (Gen2)
- 2076-624 (Gen2+)

All of these models can connect to 4 Gbps, 8 Gbps, and 16 Gbps switches. From a performance point of view, connecting the Storwize V7000 to 8 Gbps or 16 Gbps switches is better. Correct zoning on the SAN switch brings together security and performance. Implement a dual HBA approach at the host to access the Storwize V7000.

### 3.4.2 Disk subsystems

From a performance perspective, the following guidelines relate to connecting to a Storwize V7000:

- Connect all storage ports to the switch up to a maximum of 16, and zone them to all of the Storwize V7000 ports.
- Zone all ports on the disk back-end storage to all ports on the Storwize V7000 nodes in a clustered system.

- Also, ensure that you configure the storage subsystem LUN-masking settings to map all LUNs that are used by the Storwize V7000 to all the Storwize V7000 WWPNs in the clustered system.

The Storwize V7000 is designed to handle large quantities of multiple paths from the back-end storage.

In most cases, the Storwize V7000 can improve performance, especially on mid-sized to low-end disk subsystems, older disk subsystems with slow controllers, or uncached disk systems, for the following reasons:

- The Storwize V7000 can stripe across disk arrays, and it can stripe across the entire set of supported physical disk resources.
- The 2076-524 has 32GB of cache and 2076-624 has 32 GB of cache (upgrade to 64 GB possible).

The Storwize V7000 large cache and advanced cache management algorithms also allow it to improve on the performance of many types of underlying disk technologies. The Storwize V7000 capability to manage (in the background) the destaging operations that are incurred by writes (in addition to still supporting full data integrity) has the potential to be important in achieving good database performance.

Depending on the size, age, and technology level of the disk storage system, the total cache that is available in the Storwize V7000 can be larger, smaller, or about the same as the cache that is associated with the disk storage.

Because hits to the cache can occur in the upper (Storwize V7000) or the lower (disk controller) level of the overall system, the system as a whole can use the larger amount of cache wherever it is located. Therefore, if the storage control level of the cache has the greater capacity, expect hits to this cache to occur, in addition to hits in the Storwize V7000 cache.

Also, regardless of their relative capacities, both levels of cache tend to play an important role in enabling sequentially organized data to flow smoothly through the system. The Storwize V7000 cannot increase the throughput potential of the underlying disks in all cases, because this increase depends on the underlying storage technology and the degree to which the workload exhibits hotspots or sensitivity to cache size or cache algorithms.

### 3.4.3 Storwize V7000

The Storwize V7000 clustered system is scalable up to eight nodes, and the performance is nearly linear when more nodes are added into a Storwize V7000 clustered system until it becomes limited by other components in the storage infrastructure. Although virtualization with the Storwize V7000 provides a great deal of flexibility, it does not diminish the necessity to have a SAN and disk subsystems that can deliver the performance that you want.

Essentially, Storwize V7000 performance improvements are gained by having as many MDisks as possible, which creates a greater level of concurrent I/O to the back end without overloading a single disk or array.

Assuming that no bottlenecks exist in the SAN or on the disk subsystem, you must follow specific guidelines when you perform the following tasks:

- Creating a storage pool
- Creating volumes
- Connecting to or configuring hosts that must receive disk space from a Storwize V7000 clustered system

For more information about performance and preferred practices for the SVC, see *IBM System Storage SAN Volume Controller and Storwize V7000 Best Practices and Performance Guidelines*, SG24-7521:

`http://www.redbooks.ibm.com/abstracts/sg247521.html`

## 3.4.4 IBM Real-time Compression

IBM Real-time Compression (RtC) technology in storage systems is based on the Random Access Compression Engine (RACE) technology. It is implemented in IBM SAN Volume Controller and the IBM Storwize family, IBM FlashSystem V840 systems, IBM FlashSystem V9000 systems, and IBM XIV (IBM Spectrum Accelerate™). This technology plays a key role in storage capacity savings and investment protection.

Although the technology is easy to implement and manage, it is helpful to understand the internal processes and I/O workflow to ensure a successful implementation of any storage solution.

To learn more, see the following resources:

- IBM Real-time Compression Redbooks publications
- Chapter 10, “Advanced features for storage efficiency” on page 363

General suggestions:

- Best results can be achieved if the data compression ratio stays at 25% or above. Volumes can be scanned with the built-in Comprestimator, and the appropriate decision can be made.
- More concurrency within the workload gives a better result than single-threaded sequential I/O streams.
- I/O is de-staged to RACE from the upper cache in 64 KiB pieces, and best results will be achieved if the I/O size does not exceed this size.
- Volumes used for only one purpose usually have the same work patterns. Mixing database, virtualization and general-purpose data within the same volume could make the workload inconsistent. These might have no stable I/O size and no specific work pattern, and a below-average compression ratio, making these volumes hard to investigate in case of a performance degradation. Real-time Compression development advises not mixing data types within the same volume whenever possible.
- Pre-compressed data is best not to recompress, so volumes with compressed data can stay as uncompressed volumes.
- Volumes with encrypted data have a very low compression ratio, and are not good candidates for compression.

## 3.4.5 Performance monitoring

Performance monitoring must be a part of the overall IT environment. For the Storwize V7000 and other IBM storage subsystems, the official IBM tool to collect performance statistics and provide a performance report is IBM Tivoli Storage Productivity Center.

# 4

# Initial configuration

This chapter describes the first steps to configure the IBM Storwize V7000 system. It provides step-by-step instructions on how to create the cluster, define its basic settings, and add extra control and expansion enclosures.

Additional features such as user authentication and secure communications are also covered. These features are optional and do not need to be configured during the initial configuration.

This chapter includes the following topics:

- Prerequisites
- System initialization
- System setup
- Configuring user authentication
- Configuring secure communications

## 4.1 Prerequisites

Before initializing and setting up the Storwize V7000 ensure the following prerequisites are fulfilled:

- Storwize V7000 control and expansion enclosures have been physically installed with the correct cabling. The Ethernet and Fibre Channel connectivity has been correctly configured.
- The system is powered on. For information about powering on the system, see the IBM Knowledge Centre for Storwize V7000:

  `https://ibm.biz/BdsKSp`
- Your web browser is supported and has the appropriate settings enabled. For a list of the supported browsers and settings, see the IBM Knowledge Centre for Storwize V7000:

  `https://ibm.biz/BdsKSg`
- You have the following information available:
  - For IPv4 addressing:
    - Cluster IPv4 address. This address is used for the management of the system.
    - Service IPv4 addresses. These addresses are used for the service interfaces and you need one address for each node canister.
    - IPv4 subnet mask.
    - IPv4 gateway.
  - For IPv6 addressing:
    - Cluster IPv6 address. This address is used for the management of the system.
    - Service IPv6 addresses. These addresses are used for the service interfaces and you need one address for each node canister.
    - IPv6 prefix.
    - IPv6 gateway.
  - Licenses. The licenses indicate your entitlement to use licensed functions, which include: Remote Copy, External Virtualization, Real-time Compression, and Transparent Cloud Tiering.
  - Physical location of the system.
  - Name, email address and phone number of the storage administrator that IBM can contact if necessary.
  - Network Time Protocol (NTP) server IP address. This is optional. It is necessary only if you want to use a NTP service instead of manually entering date and time.
  - Simple Mail Transfer Protocol (SMTP) email server IP address. This is optional. It is necessary only if you want to enable *call home*.

## 4.2 System initialization

This section provides step-by-step instructions on how to create the Storwize V7000 cluster.

**Attention:** Do not repeat the instructions for system initialization on more than one node canister. After system initialization completes, use the management GUI to add more control enclosure to the system. See “Adding a control enclosure” on page 103 for information about how to perform this task.

Start by connecting a personal computer (PC) or notebook to the technician port, located on the rear of the Storwize V7000 node canister. Figure 4-1 shows the location of the technician port.

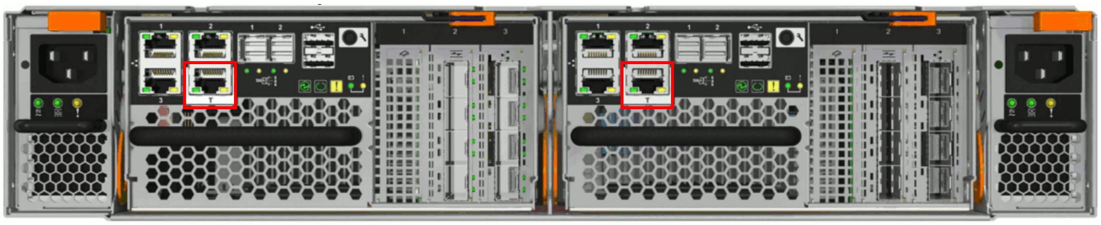


*Figure 4-1 Location of the technician port*

The technician port provides a DHCP IPv4 address, so you must ensure that your PC or notebook Ethernet port is configured for DHCP if you want the IP to be automatically assigned. If your PC or notebook does not have DHCP, you can set a static IP on the Ethernet port as `192.168.0.2`.

The default IP address for a new node canister is `192.168.0.1`.

**Note:** The Storwize V7000 does *not* provide IPv6 IP addresses for the technician port.

## 4.2.1 System initialization wizard

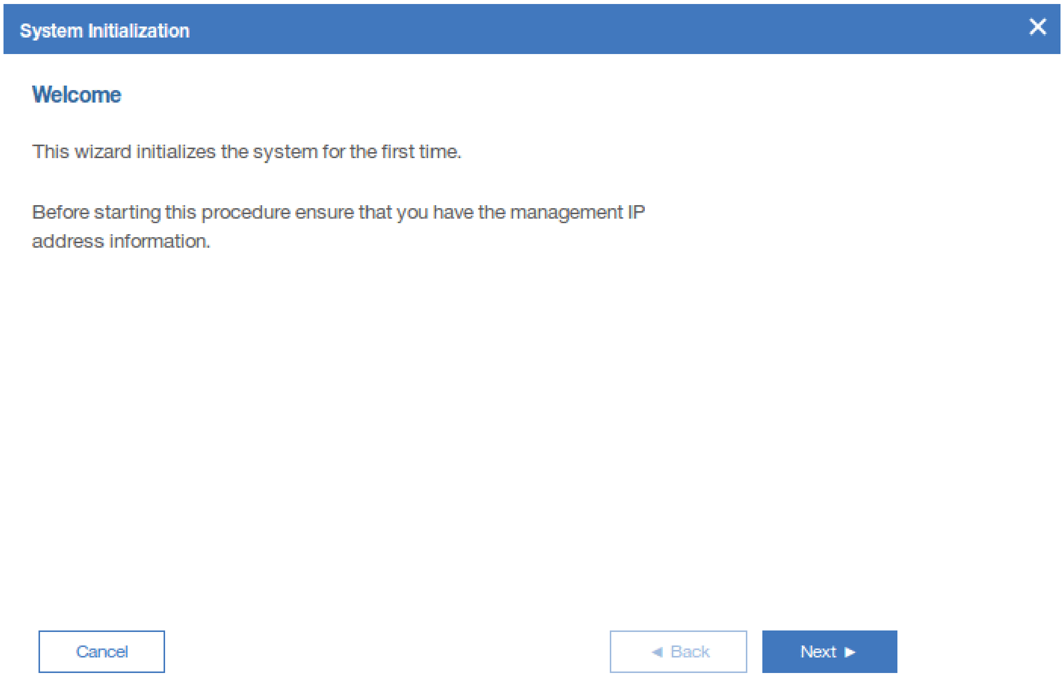
After connecting your PC or notebook to the technician port, validate that you have a valid IPv4 DCHP address (for example, `192.168.0.12`) and then follow these steps for initializing the system:

1. Open a supported browser and browse to *`http://install`*. The browser is automatically redirected to the System Initialization wizard. You can use the IP address `192.168.0.1` if you are not automatically redirected.

   **Note:** If the system cannot be initialized, you are redirected to the Service Assistant interface. Refer to the error codes shown to troubleshoot the problem.

   **Note:** During the system initialization you are prompted to accept untrusted certificates because the system certificates are self-signed. You can accept these because they are not harmful.

2. The welcome dialog box opens, as shown in Figure 4-2. Click **Next** to start the procedure.



System Initialization

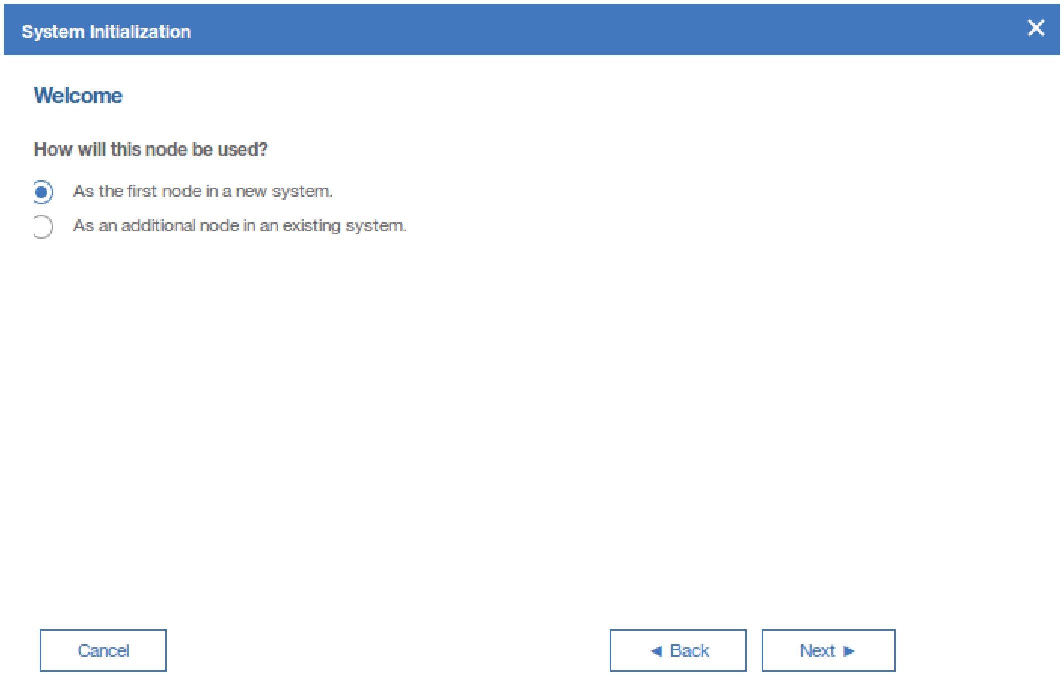
Welcome

This wizard initializes the system for the first time.

Before starting this procedure ensure that you have the management IP address information.

*Figure 4-2 System initialization: welcome*

3. Select the first option, as shown in Figure 4-3. Click **Next**.



System Initialization

Welcome

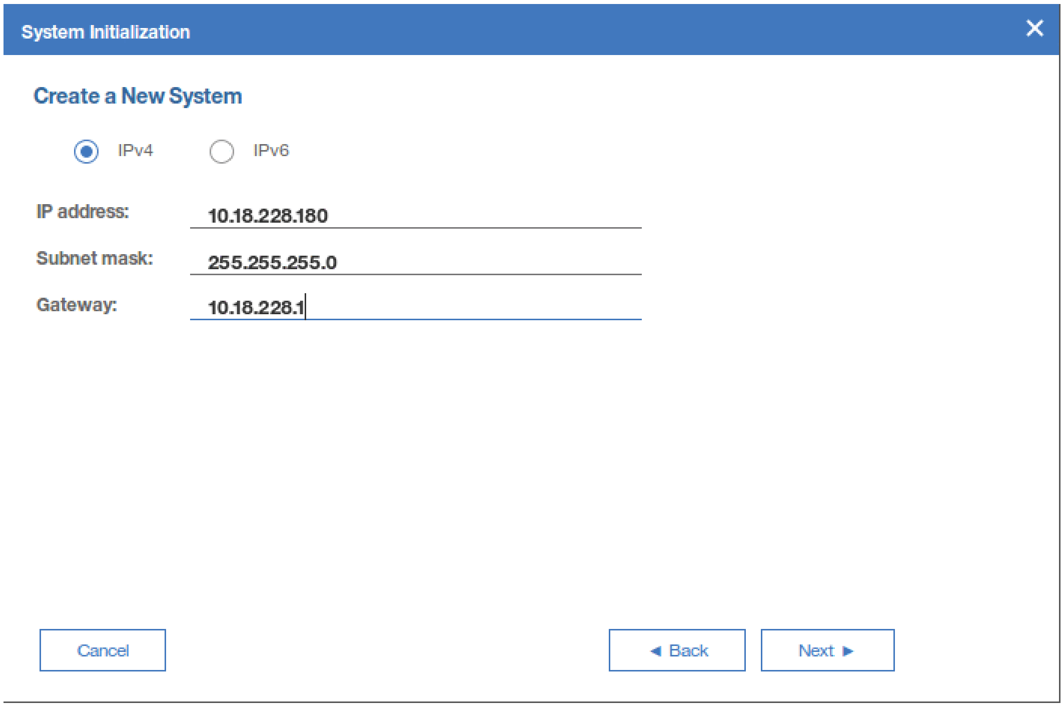
How will this node be used?

● As the first node in a new system.

○ As an additional node in an existing system.

*Figure 4-3 System initialization: configuring the first node in a new system*

4. Enter the IP address details for the new system. You can choose between an IPv4 or IPv6 address. In this example an IPv4 address is set, as shown in Figure 4-4. Click **Next**.



IP address: 10.18.228.180

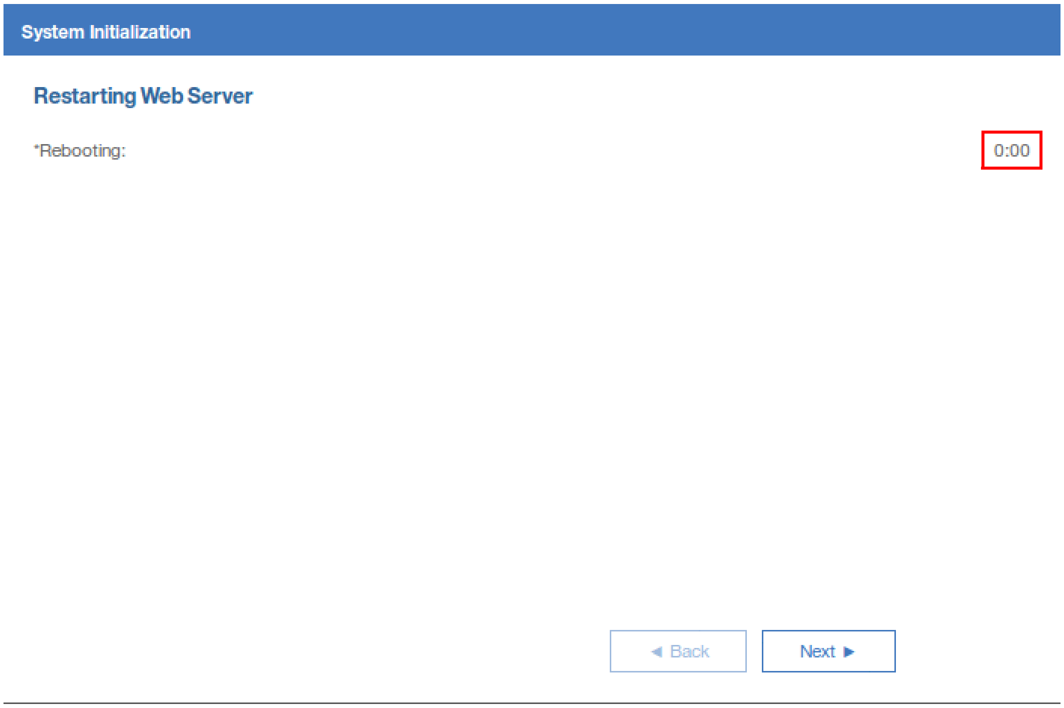
Subnet mask: 255.255.255.0

Gateway: 10.18.228.1

*Figure 4-4 System initialization: setting the system IP address*

This address is given to Ethernet port 1. After system initialization you can specify additional IP addresses for port 1 and port 2 until both ports have an IPv4 address and an IPv6 address.

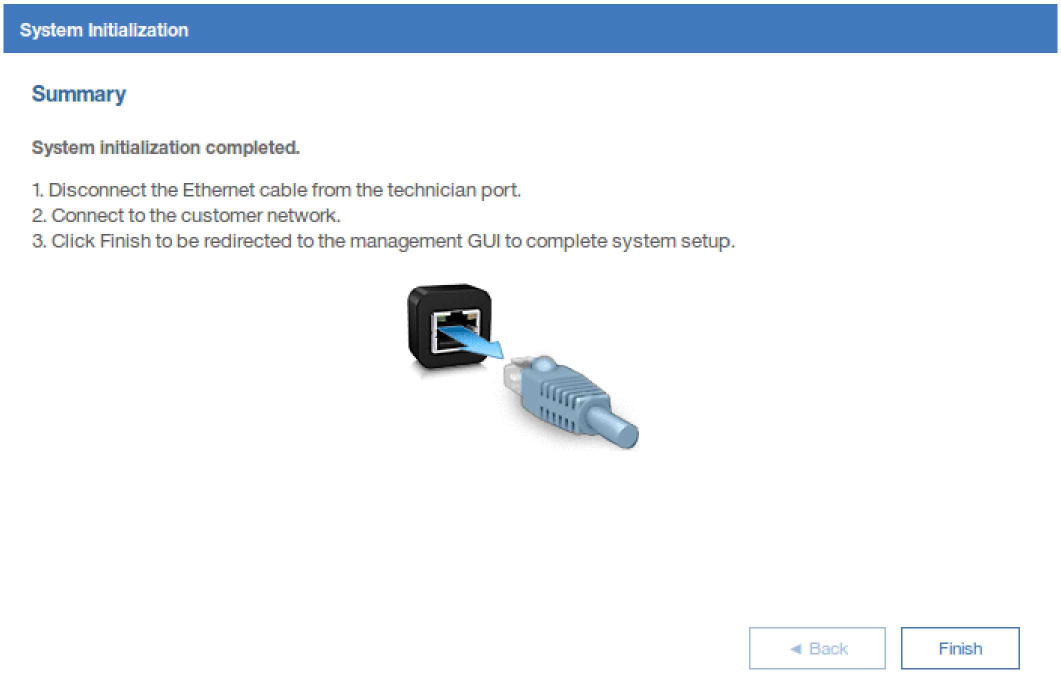
5. The web server restarts, as shown in Figure 4-5. Wait until the timer reaches the end and click **Next**.



*Rebooting: 0:00

*Figure 4-5 System initialization: web server restart*

6. After the system initialization is complete, follow the on-screen instructions listed below, that are shown in Figure 4-6 on page 90:
   a. Disconnect the Ethernet cable from the technician port and from your PC or notebook.
   b. Connect the PC or notebook to the same network as the system.
   c. Click **Finish** to be redirected to the management GUI to complete the system setup.



System Initialization

Summary

System initialization completed.

1. Disconnect the Ethernet cable from the technician port.
2. Connect to the customer network.
3. Click Finish to be redirected to the management GUI to complete system setup.

*Figure 4-6 System initialization: summary*

> **Note:** You can connect to the system IP address from any management console that is connected to the same network as the system. Enter the system IP address on a supported browser to access the management GUI.

# 4.3 System setup

This section provides step-by-step instructions on how to define the basic settings of the cluster with the system setup wizard and on how to add additional control and expansion enclosures.

## 4.3.1 System setup wizard

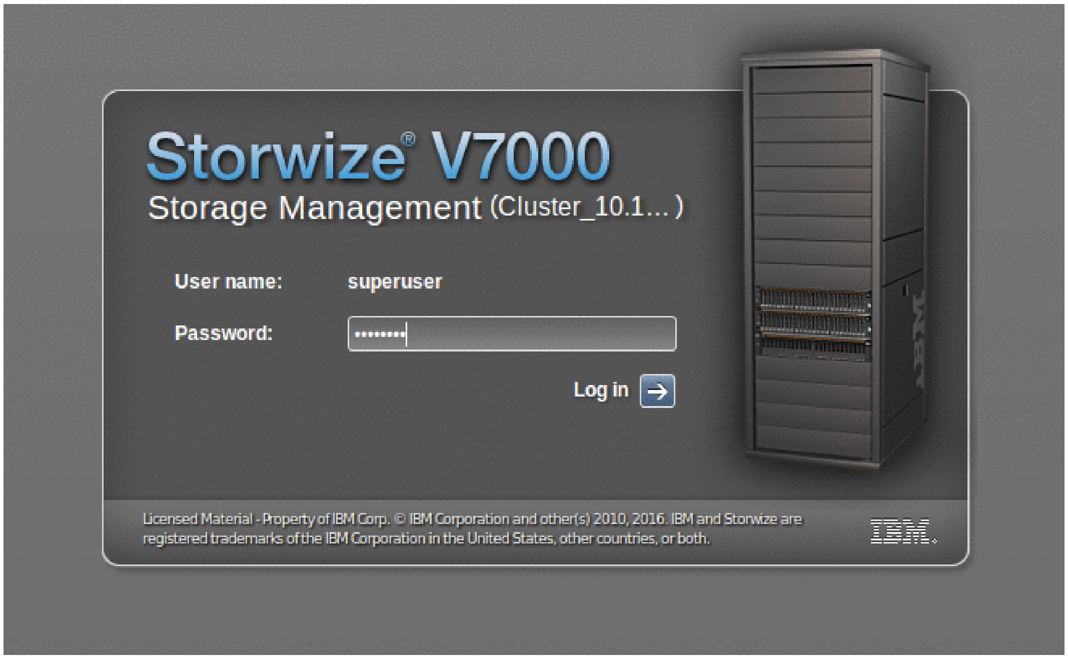
Whether you are redirected from your PC or notebook after completing system initialization, or you browse to the management IP address manually, you must complete the system setup wizard to define the basic settings of the system.

> **Note:** The first time you connect to the management GUI you are prompted to accept untrusted certificates because the system certificates are self-signed. You can accept these because they are not harmful.
>
> You can install certificates signed by a third-party certificate authority after you complete system setup. See “Configuring secure communications” on page 118 for instructions on how to perform this task.

Follow the steps below to successfully complete the system setup wizard:

1. Log in to system with the superuser account, as shown in Figure 4-7. Click **Log in**.

> **Important:** The default password for the superuser account is `passw0rd` (zero and not O).



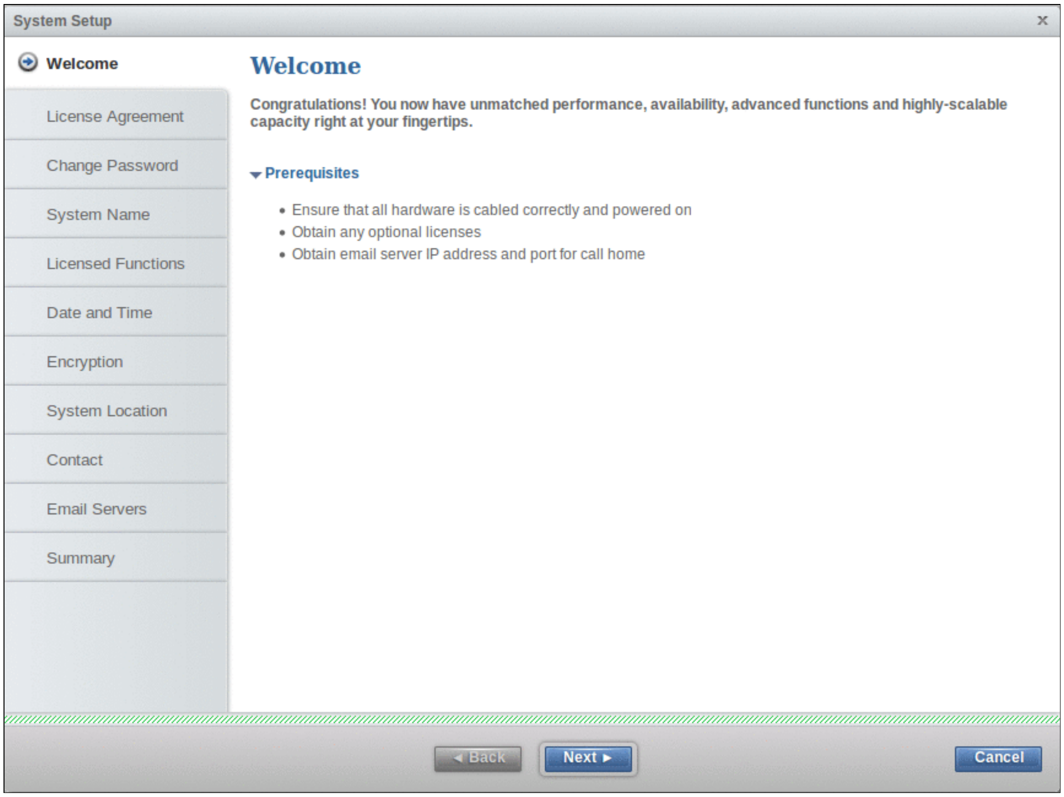
User name: superuser

Password: ••••••••

Licensed Material - Property of IBM Corp. © IBM Corporation and other(s) 2010, 2016. IBM and Storwize are registered trademarks of the IBM Corporation in the United States, other countries, or both.

*Figure 4-7 System setup: logging in for the first time*

2. The welcome dialog shown in Figure 4-8 opens. Verify the prerequisites and click **Next**.



Welcome

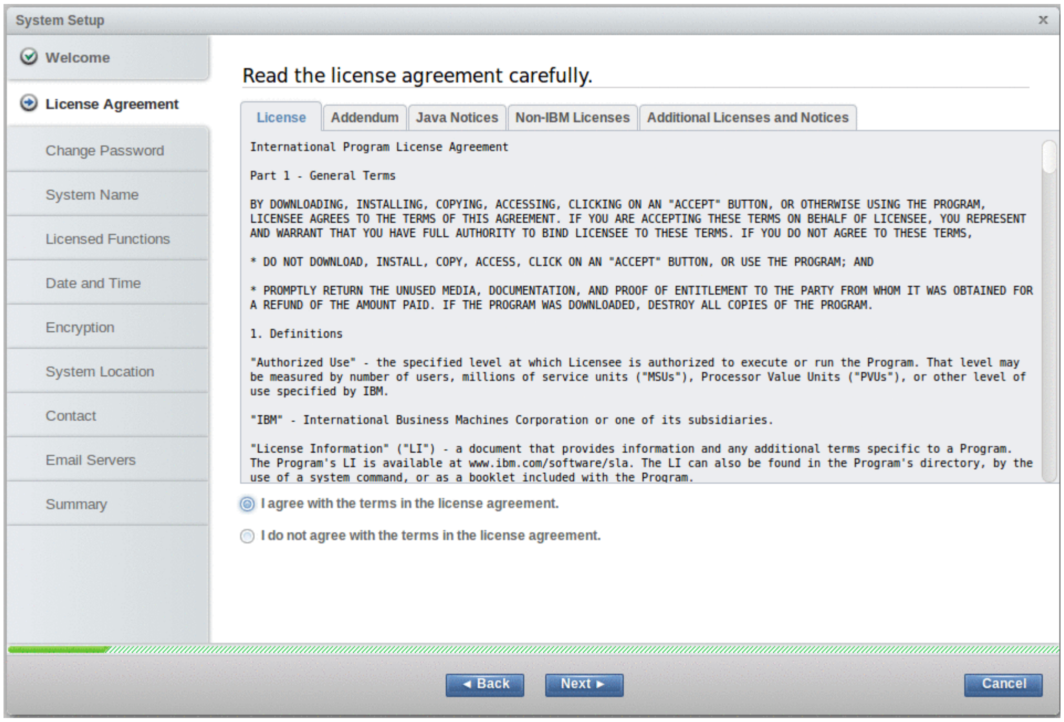
Congratulations! You now have unmatched performance, availability, advanced functions and highly-scalable capacity right at your fingertips.

▾ Prerequisites

- Ensure that all hardware is cabled correctly and powered on
- Obtain any optional licenses
- Obtain email server IP address and port for call home

*Figure 4-8 System setup: welcome*

3. Carefully read the license agreement. Select **I agree with the terms in the license agreement** when you are ready, as shown in Figure 4-9 on page 92. Click **Next**.



Read the license agreement carefully.

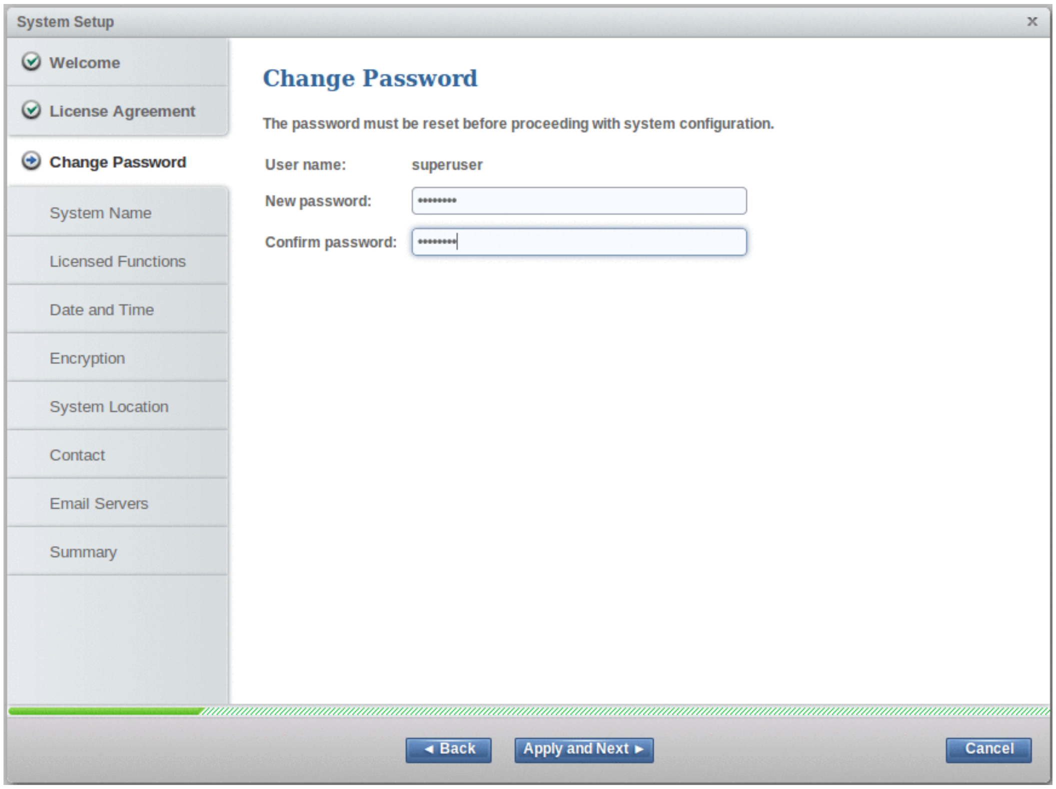
```
International Program License Agreement

Part 1 - General Terms

BY DOWNLOADING, INSTALLING, COPYING, ACCESSING, CLICKING ON AN "ACCEPT" BUTTON, OR OTHERWISE USING THE PROGRAM,
LICENSEE AGREES TO THE TERMS OF THIS AGREEMENT. IF YOU ARE ACCEPTING THESE TERMS ON BEHALF OF LICENSEE, YOU REPRESENT
AND WARRANT THAT YOU HAVE FULL AUTHORITY TO BIND LICENSEE TO THESE TERMS. IF YOU DO NOT AGREE TO THESE TERMS,

* DO NOT DOWNLOAD, INSTALL, COPY, ACCESS, CLICK ON AN "ACCEPT" BUTTON, OR USE THE PROGRAM; AND

* PROMPTLY RETURN THE UNUSED MEDIA, DOCUMENTATION, AND PROOF OF ENTITLEMENT TO THE PARTY FROM WHOM IT WAS OBTAINED FOR
A REFUND OF THE AMOUNT PAID. IF THE PROGRAM WAS DOWNLOADED, DESTROY ALL COPIES OF THE PROGRAM.

1. Definitions

"Authorized Use" - the specified level at which Licensee is authorized to execute or run the Program. That level may
be measured by number of users, millions of service units ("MSUs"), Processor Value Units ("PVUs"), or other level of
use specified by IBM.

"IBM" - International Business Machines Corporation or one of its subsidiaries.

"License Information" ("LI") - a document that provides information and any additional terms specific to a Program.
The Program's LI is available at www.ibm.com/software/sla. The LI can also be found in the Program's directory, by the
use of a system command, or as a booklet included with the Program.
```

● I agree with the terms in the license agreement.

○ I do not agree with the terms in the license agreement.

*Figure 4-9 System setup: license agreement*

4. Enter a new password for superuser, as shown in Figure 4-10. The password length is 6 - 63 characters and it cannot begin or end with a space. Click **Apply and Next**.



Change Password

The password must be reset before proceeding with system configuration.

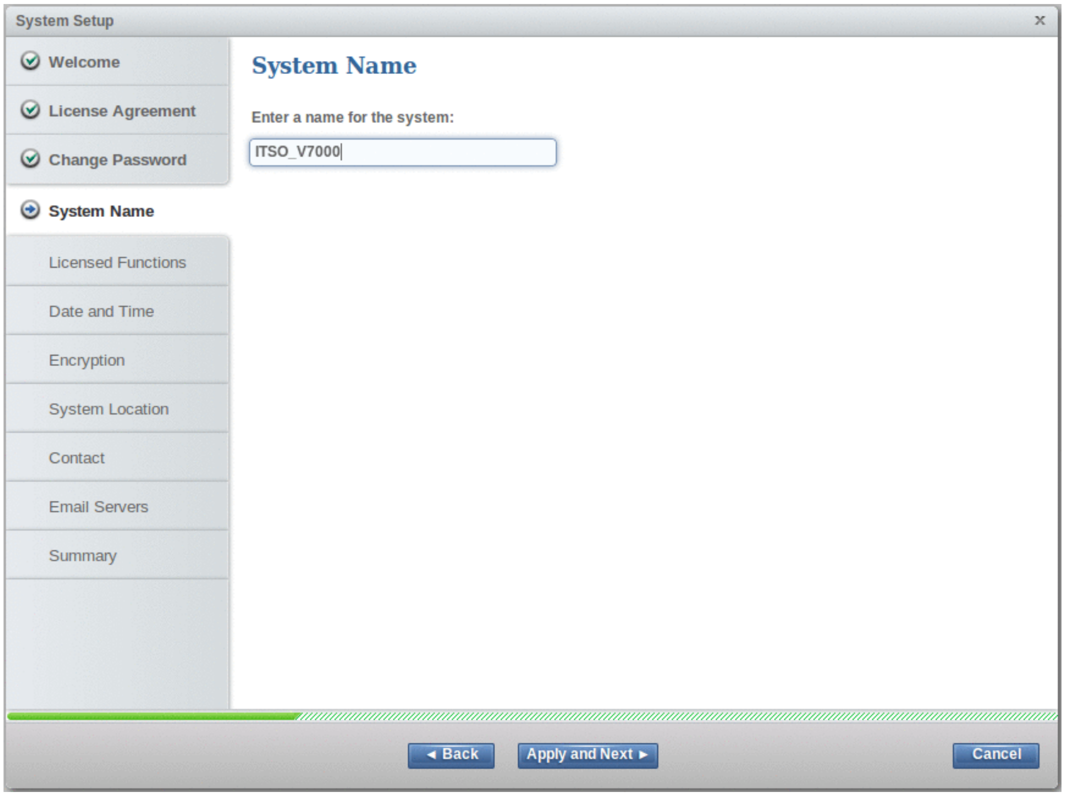
User name: superuser

New password: ••••••••

Confirm password: ••••••••

*Figure 4-10 System setup: changing the password for superuser*

5. Enter the name you want to give the new system, as shown in Figure 4-11 on page 94. Click **Apply and Next**.



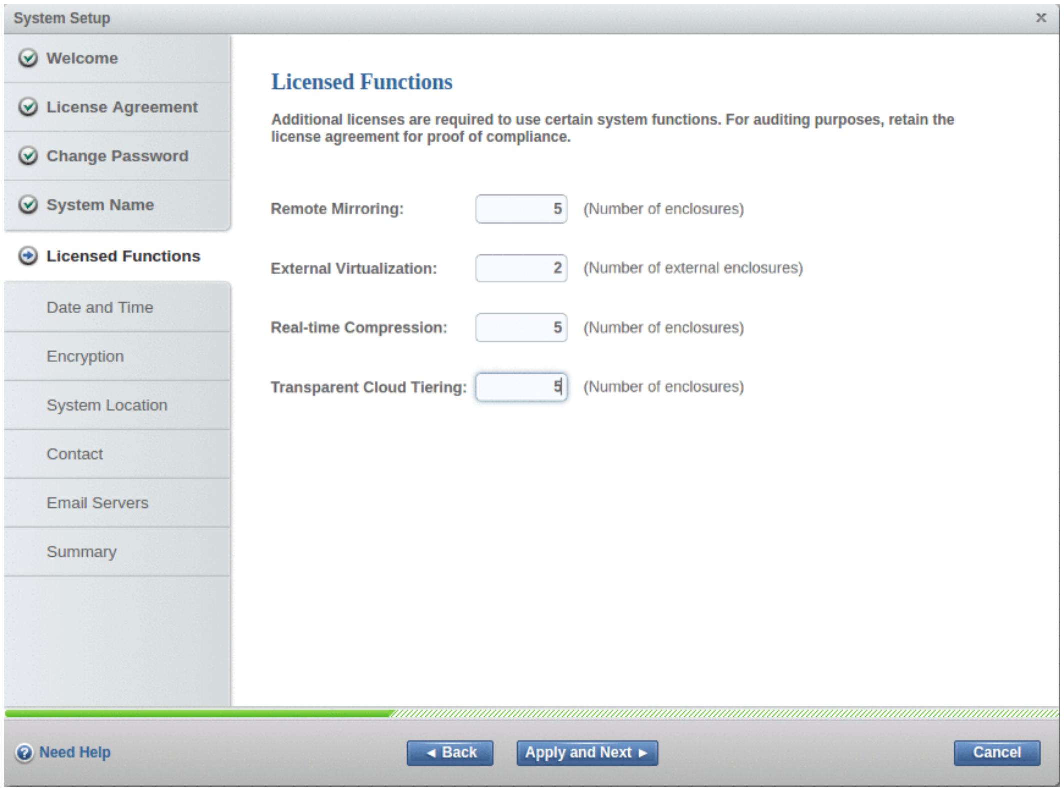
System Name

Enter a name for the system: ITSO_V7000

*Figure 4-11 System setup: setting the system name*

6. Enter the number of enclosures licensed for each function as authorized by your license agreement. Figure 4-12 shows some values as an example only. Click **Apply and Next**.

> **Note:** Encryption uses a different licensing scheme and is activated later in the wizard.



Licensed Functions

Additional licenses are required to use certain system functions. For auditing purposes, retain the license agreement for proof of compliance.

Remote Mirroring: 5 (Number of enclosures)

External Virtualization: 2 (Number of external enclosures)

Real-time Compression: 5 (Number of enclosures)

Transparent Cloud Tiering: 5 (Number of enclosures)

*Figure 4-12 System setup: setting the system licenses*

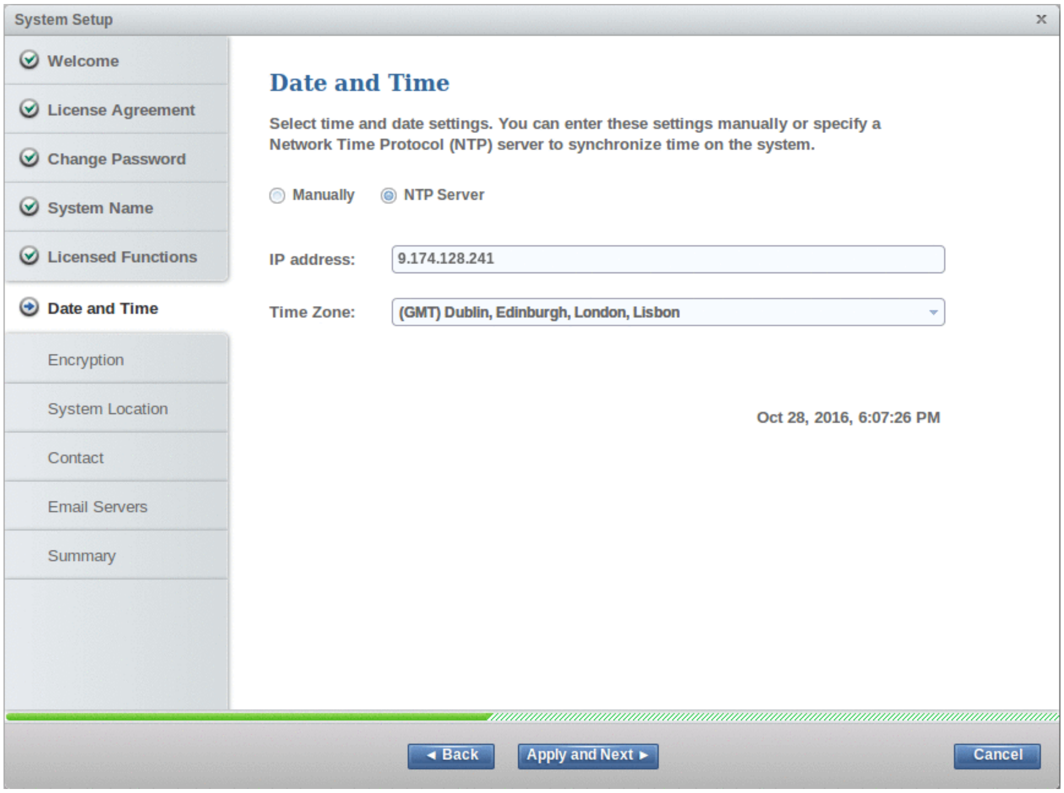
All functions follow an enclosure-based licensing scheme. Read the guidelines below to ensure you are licensing the correct number of enclosures:

– **External Virtualization**. The system does not require a license for its own control and expansion enclosures. However, a license is required for each enclosure of any external storage systems being virtualized. For example, if the system is made up of one control enclosure, one expansion enclosure, and one virtualized storage system that is itself made up of two enclosures, then two licenses are required. The system does not require an external virtualization license for external enclosures that are only being used to provide managed disks for a quorum disk and are not providing any capacity for volumes.

– **Remote Mirroring**, **Real-time Compression**, or **Transparent Cloud Tiering**. The total number of enclosures must include the enclosures on external storage system licensed for virtualization plus the number of control and expansion enclosures that are part of your local system. For example, if the system is made up of one control enclosure, one expansion enclosure, and one virtualized storage system that is itself made up of two enclosures, then four licenses are required. Only Storwize V7000 Gen2 and Storwize V7000 Gen2+ models support transparent cloud tiering.

7. Enter the date and time details. In this example date and time are set using a NTP, as shown in Figure 4-13 on page 96. It is recommended that you use an NTP server so that all of your storage area network and storage devices have a common time stamp for troubleshooting. Click **Apply and Next**.

**Note:** If you choose to manually enter these settings, you cannot select the 24-hour clock at this time. However, you can select the 24-hour clock after you complete the wizard by navigating to **Settings** → **System** and selecting **Date and Time**.



Date and Time

Select time and date settings. You can enter these settings manually or specify a Network Time Protocol (NTP) server to synchronize time on the system.

○ Manually ● NTP Server

IP address: 9.174.128.241

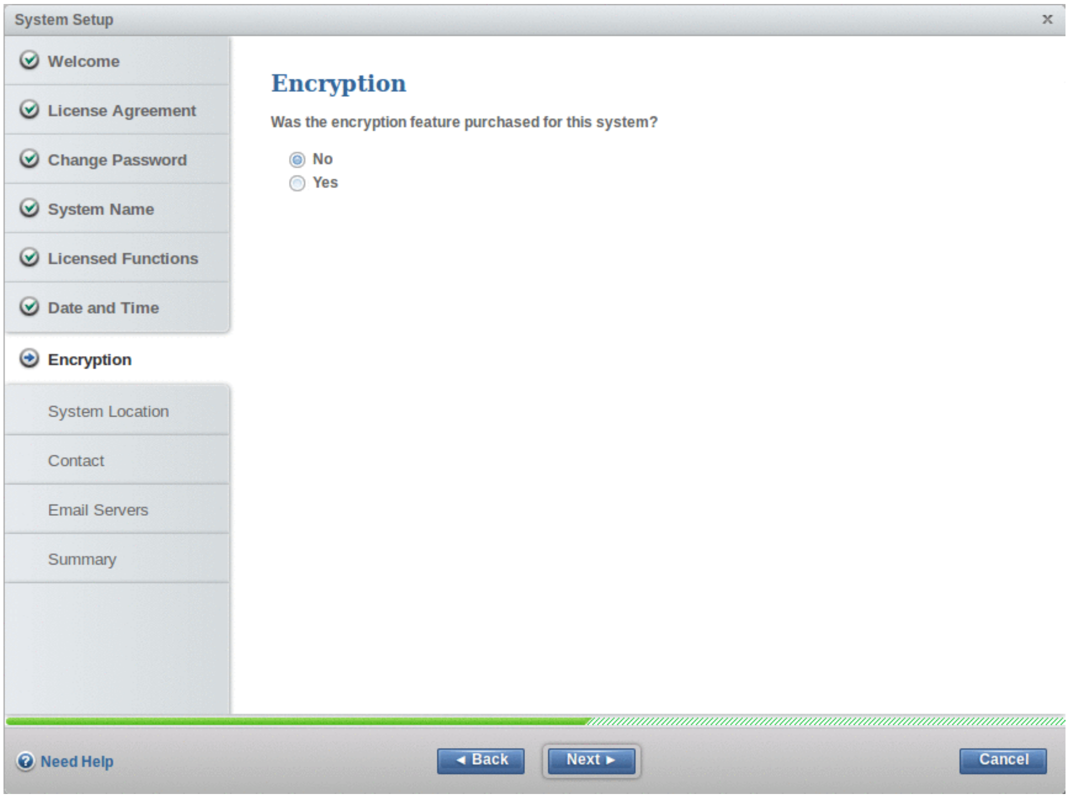
Time Zone: (GMT) Dublin, Edinburgh, London, Lisbon

Oct 28, 2016, 6:07:26 PM

*Figure 4-13 System setup: setting date and time settings*

8. Select whether the encryption feature has been purchased for this system or not. In this example it is assumed encryption has not been purchased, as shown in Figure 4-14 on page 97.

> **Note:** If you have purchased the encryption feature you are prompted to activate your encryption license either manually or automatically. For information about how to activate your encryption license during the system setup wizard see Chapter 12, “Encryption” on page 563.



Encryption

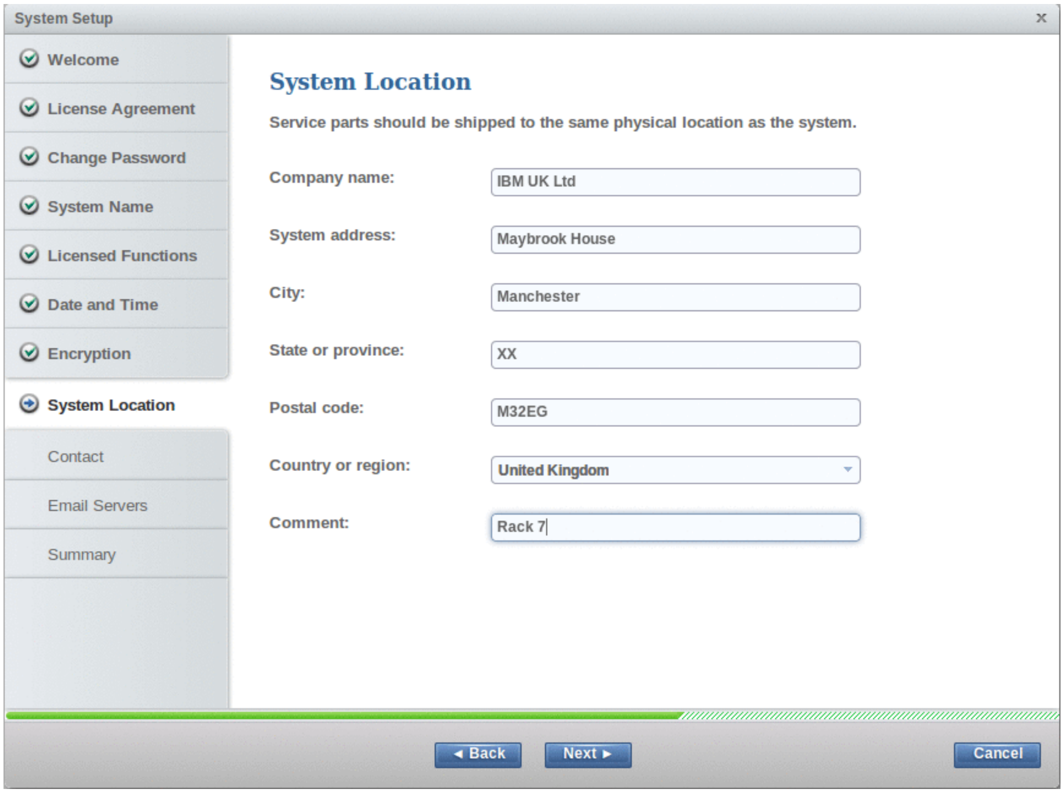
Was the encryption feature purchased for this system?

● No
○ Yes

*Figure 4-14 System setup: encryption*

9. Enter the system location details. Figure 4-15 shows some details as an example only. Click **Next**.

> **Note:** If your system is not in the US, complete the state or province field with `XX`.



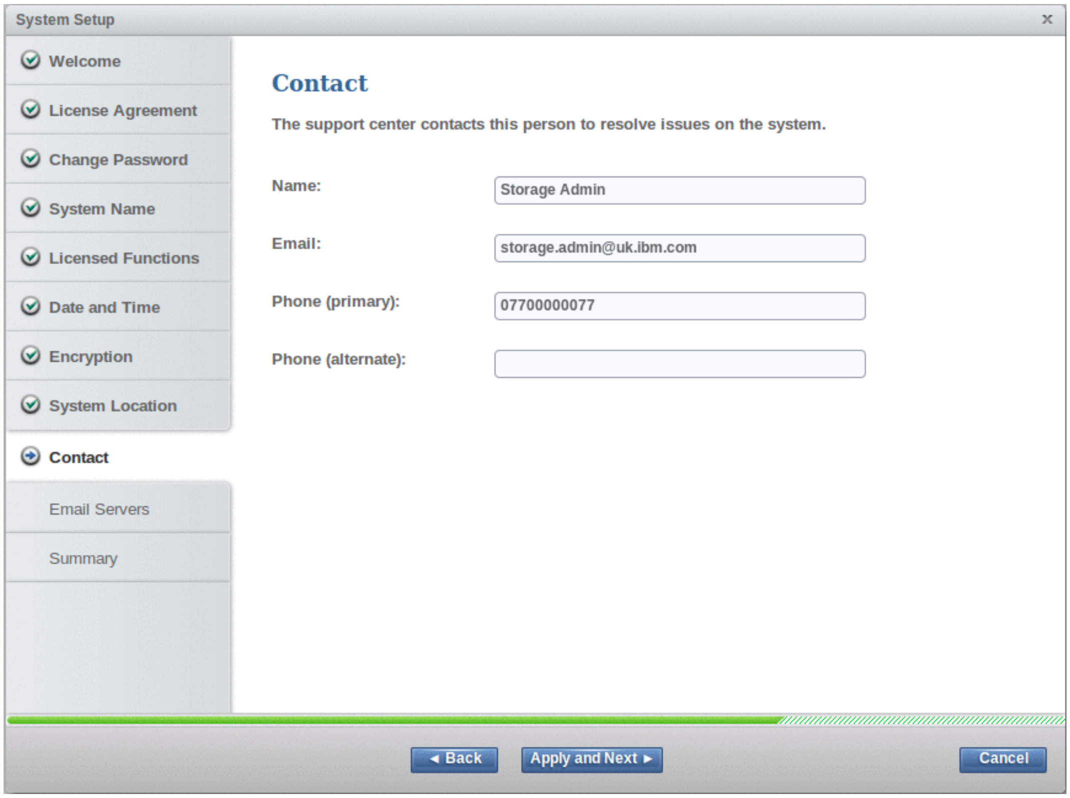
## System Location

Service parts should be shipped to the same physical location as the system.

| | |
|---|---|
| Company name: | IBM UK Ltd |
| System address: | Maybrook House |
| City: | Manchester |
| State or province: | XX |
| Postal code: | M32EG |
| Country or region: | United Kingdom |
| Comment: | Rack 7 |

*Figure 4-15 System setup: setting the system location details*

10. Enter the contact details of the person to be contacted to resolve issues on the system. You can choose to enter the details for a 24-hour operations desk. Figure 4-16 shows some details as an example only. Click **Apply and Next**.



Contact

The support center contacts this person to resolve issues on the system.

Name: Storage Admin

Email: storage.admin@uk.ibm.com

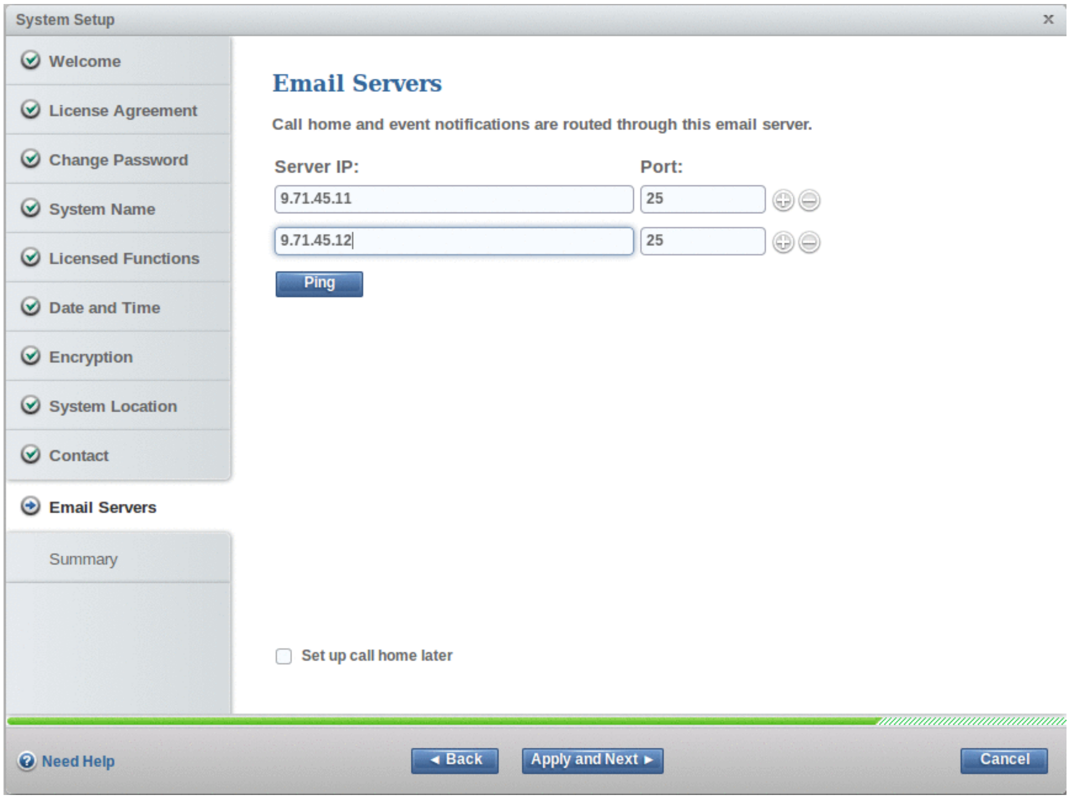
Phone (primary): 07700000077

Phone (alternate):

*Figure 4-16 System setup: setting contact information*

11. Enter the details for the email servers to be used for *call home*. Call home sends email reports to IBM with inventory details and event notifications. This allows IBM to automatically open problem reports and to contact you to verify if replacement parts are required.

    Figure 4-17 shows some details as an example only. You can click **Ping** to verify the email server is accessible over the network. Click **Apply and Next**.



Email Servers

Call home and event notifications are routed through this email server.

Server IP: 9.71.45.11 Port: 25

Server IP: 9.71.45.12 Port: 25

□ Set up call home later

*Figure 4-17 Setting email servers details*

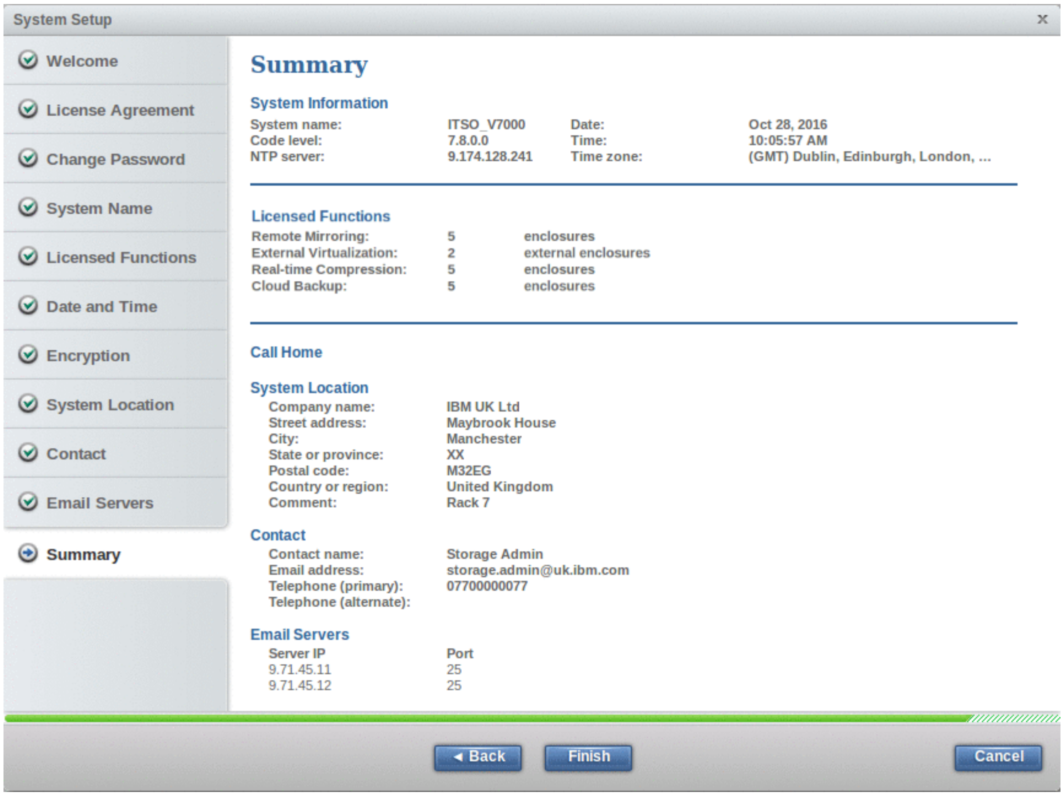
Storwize V7000 can use Simple Network Management Protocol (SNMP) traps, syslog messages, and call home to notify you and the IBM Support Center when significant events are detected. Any combination of these notification methods can be used simultaneously. However, only call home is configured during the system setup wizard. For information about how to configure other notification methods see Chapter 13, “RAS, monitoring, and troubleshooting” on page 605.

> **Note:** When call home is configured the system automatically creates a support contact with one of the following email addresses, depending on country or region of installation:
>
> - US, Canada, Latin America, and Caribbean Islands: `callhome1@de.ibm.com`
> - All other countries or regions: `callhome0@de.ibm.com`

If you do not want to configure call home now, it can be done later by navigating to **Settings** → **Notifications**.

> **Note:** If your system is under warranty or you have a hardware maintenance agreement, it is advisable to configure call home.

12. A summary of all the changes is displayed, as shown in Figure 4-18. Confirm the changes are correct and click **Finish**.



Summary

System Information

| | | | |
|---|---|---|---|
| System name: | ITSO_V7000 | Date: | Oct 28, 2016 |
| Code level: | 7.8.0.0 | Time: | 10:05:57 AM |
| NTP server: | 9.174.128.241 | Time zone: | (GMT) Dublin, Edinburgh, London, ... |

Licensed Functions

| | | |
|---|---|---|
| Remote Mirroring: | 5 | enclosures |
| External Virtualization: | 2 | external enclosures |
| Real-time Compression: | 5 | enclosures |
| Cloud Backup: | 5 | enclosures |

Call Home

System Location

| | |
|---|---|
| Company name: | IBM UK Ltd |
| Street address: | Maybrook House |
| City: | Manchester |
| State or province: | XX |
| Postal code: | M32EG |
| Country or region: | United Kingdom |
| Comment: | Rack 7 |

Contact

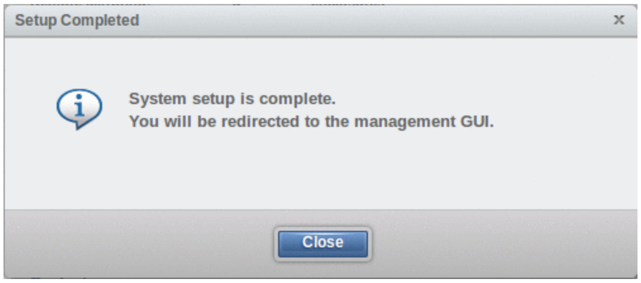
| | |
|---|---|
| Contact name: | Storage Admin |
| Email address: | storage.admin@uk.ibm.com |
| Telephone (primary): | 07700000077 |
| Telephone (alternate): | |

Email Servers

| Server IP | Port |
|---|---|
| 9.71.45.11 | 25 |
| 9.71.45.12 | 25 |

*Figure 4-18 System setup: summary*

13. The message shown in Figure 4-19 opens, confirming the setup is complete. Click **Close**. You are automatically redirected to the System panel.



System setup is complete.
You will be redirected to the management GUI.

*Figure 4-19 System setup: completion*

When the wizard completes your system consists only of the control enclosure that includes the node canister you used to initialize the system. If you have other control and expansion enclosures, you must add them to complete system setup.

The System panel displays the enclosures that are part of the system and allows you to add new enclosures to the system. Figure 4-20 on page 102 shows the System panel for a system with one control enclosure and connected to no other enclosures.

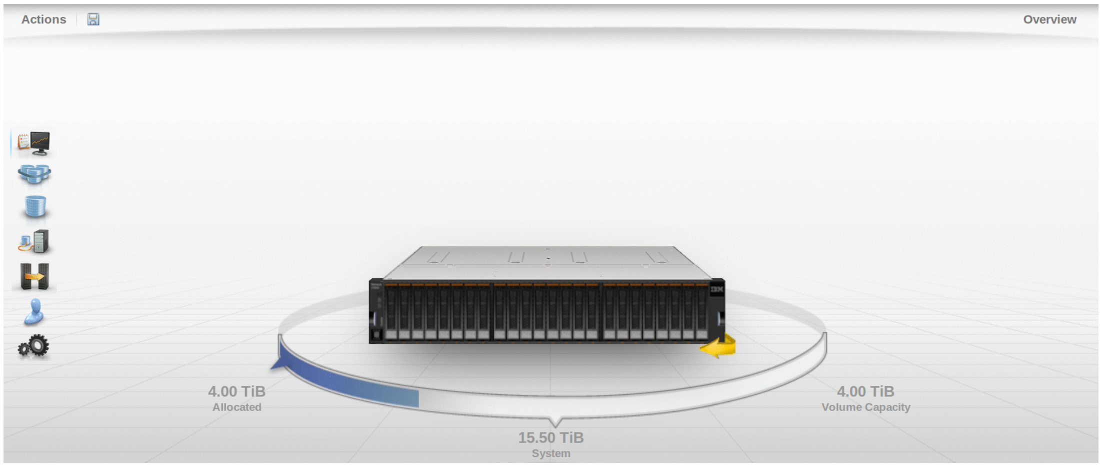


*Figure 4-20 System panel: one control enclosure and no enclosures to add*

Figure 4-21 shows the System panel for a system with one control enclosure and connected to an expansion that is not yet part of the system.

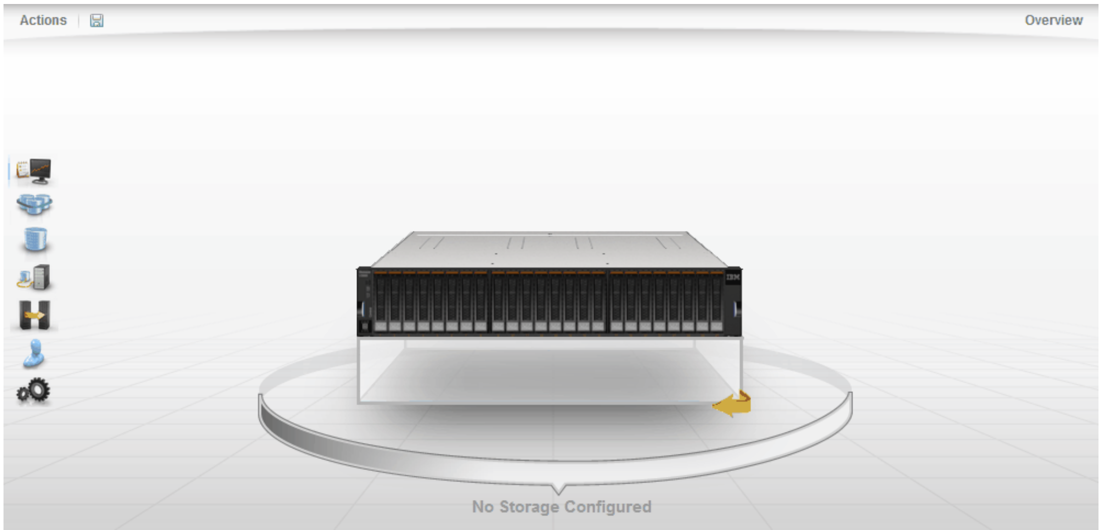


*Figure 4-21 System panel: one control enclosure and one expansion to add*

Figure 4-22 on page 103 shows the System panel for a system with one control enclosure and connected to a second control enclosure that is not yet part of the system.

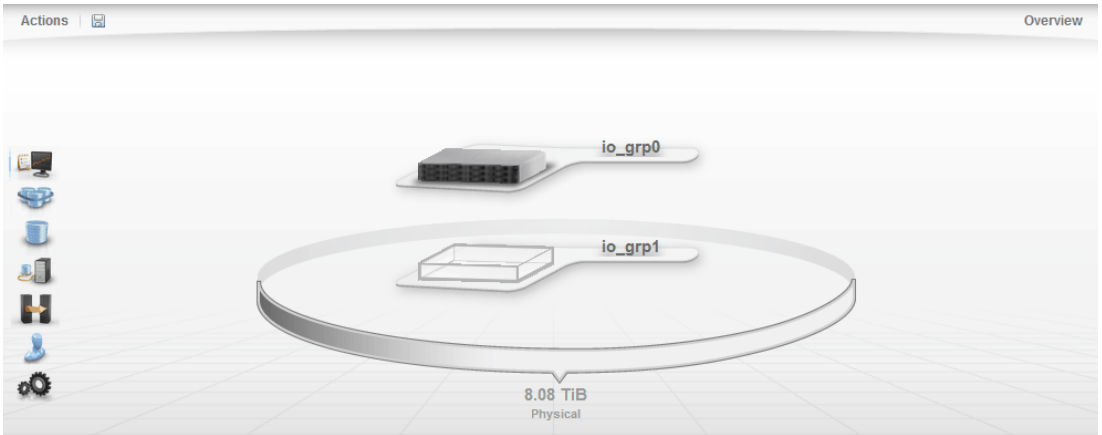


*Figure 4-22 System panel: one control enclosure and one control enclosure to add*

See “Adding a control enclosure” on page 103 and “Adding an expansion enclosure” on page 105 for instructions on how to add the enclosures to complete system setup.

Completing system setup means all mandatory steps of the initial configuration have been executed and you can start configuring your storage. Optionally, you can configure other features, such as user authentication and secure communications.

## 4.3.2 Adding a control enclosure

This procedure is the same whether you are configuring the system for the first time or expanding it afterwards.

Before commencing, ensure the new control enclosure is correctly installed and cabled to the existing system. Ensure the Ethernet and Fibre Channel connectivity has been correctly configured and that the enclosure is powered on.

If all prerequisites have been fulfilled, the Systems panel displays an empty I/O group, as shown in Figure 4-23. If more control enclosures are available, more empty I/O groups are shown. The number of empty I/O groups corresponds to the number of control enclosures in the fabric that have not yet been added to the system.

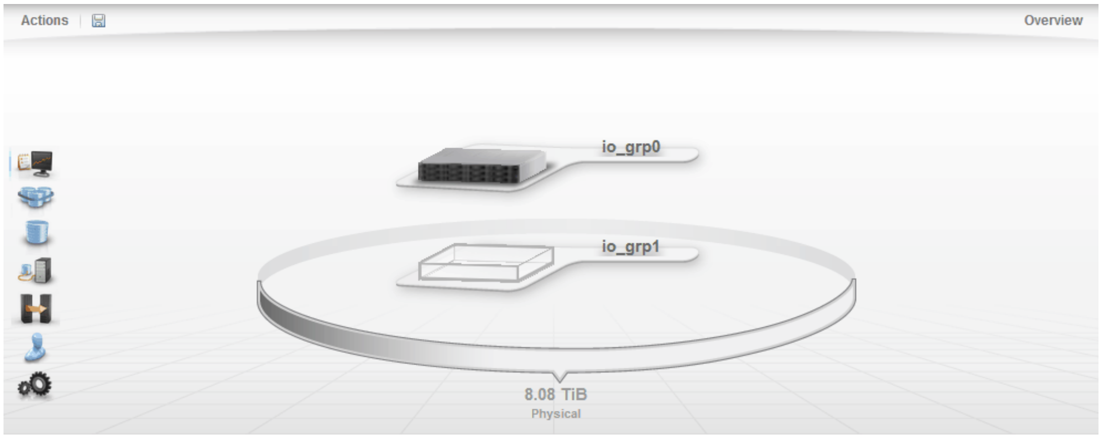


*Figure 4-23 System panel: empty I/O group*

Complete the following steps to add a new control enclosure to the system:

1. Click the first empty I/O group, as shown in Figure 4-24 on page 104. Alternatively, click **Actions** and then **Add Enclosures**.

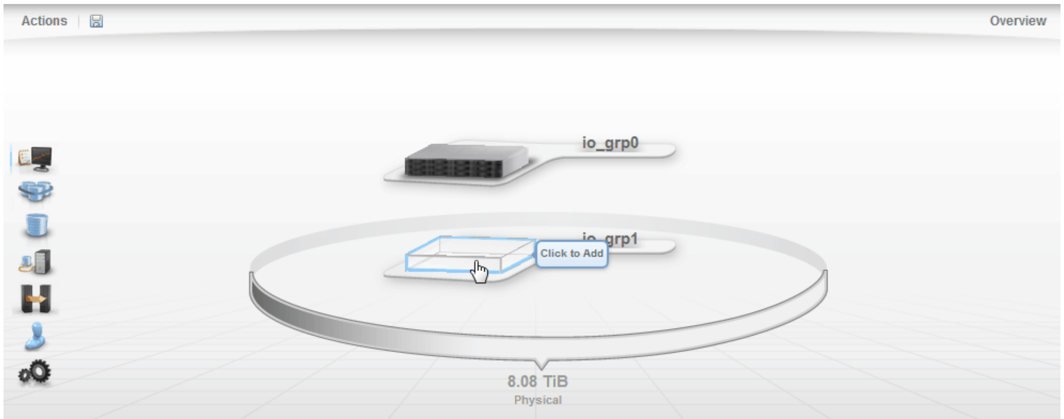
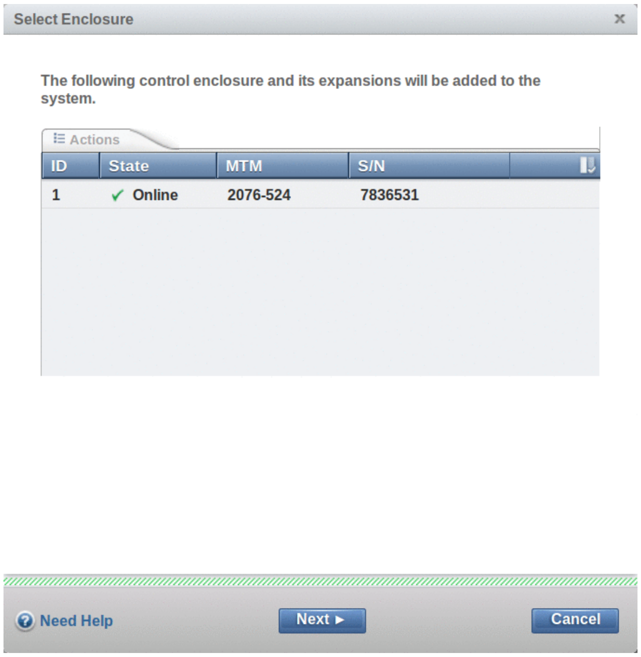


*Figure 4-24 System panel: option to add a control enclosure*

2. The first dialog displays the available control enclosure and its details, as shown in Figure 4-25. You can turn on the identify LED lights on the control enclosure by right-clicking the enclosure and selecting **Identify**. Click **Next** to proceed.

> **Note:** The expansions directly cabled to the new control enclosure are not listed. However, they are added automatically when the control enclosure is added.



The following control enclosure and its expansions will be added to the system.

| ID | State | MTM | S/N |
|---|---|---|---|
| 1 | ✓ Online | 2076-524 | 7836531 |

*Figure 4-25 Adding a control enclosure to the system*

3. Review the summary in the next dialog and press **Finish** to add the control enclosure and all its expansions to the system.

The System panel now displays two complete I/O groups, as shown in Figure 4-26 on page 105.

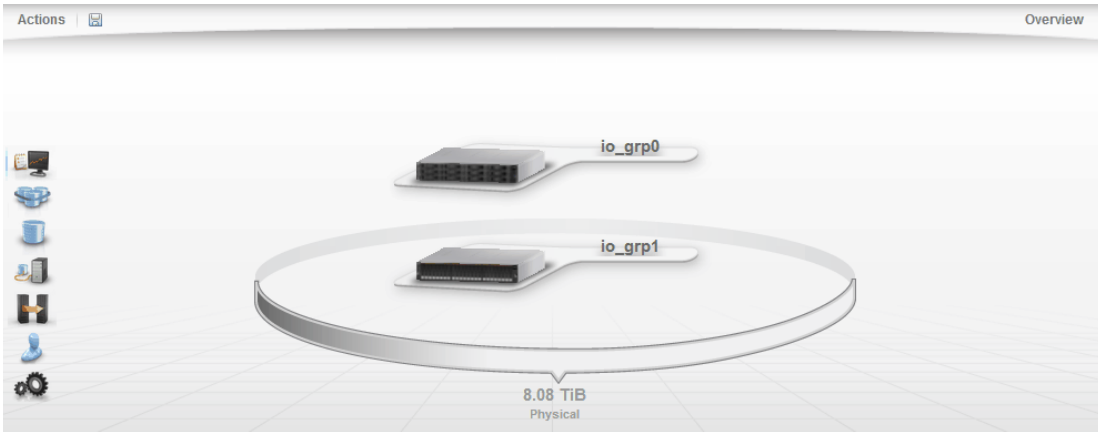


*Figure 4-26 System panel: two control enclosures and no expansions*

### 4.3.3 Adding an expansion enclosure

When a control enclosure is manually added to the system all the expansion enclosures attached to it are added automatically. However, you must manually add expansion enclosures if:

- You completed the system setup wizard and you have expansions attached to the control enclosure you used to initiate the system
- You attached expansions to a control enclosure that is already part of the system

When an expansion enclosure is attached to a system, the System panel displays an empty expansion under the control enclosure to which the expansion is attached, provided no expansions have been added before under the same control enclosure. Figure 4-27 shows an empty expansion enclosure for a system with one control enclosure.

If you are adding expansions under a control enclosure that already has other expansions installed, a plus sign is displayed on the left side of the expansion enclosure.

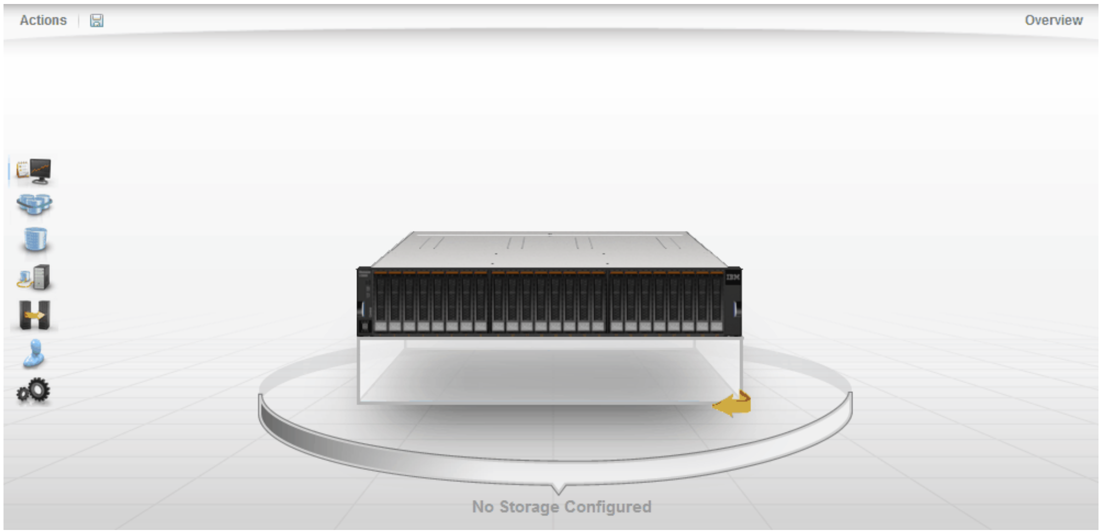


*Figure 4-27 System panel: empty expansion enclosure*

Complete the following steps to add a new expansion enclosure:

1. Click the empty expansion enclosure, as shown in Figure 4-28. Alternatively, click **Actions** and then **Add Enclosures**.

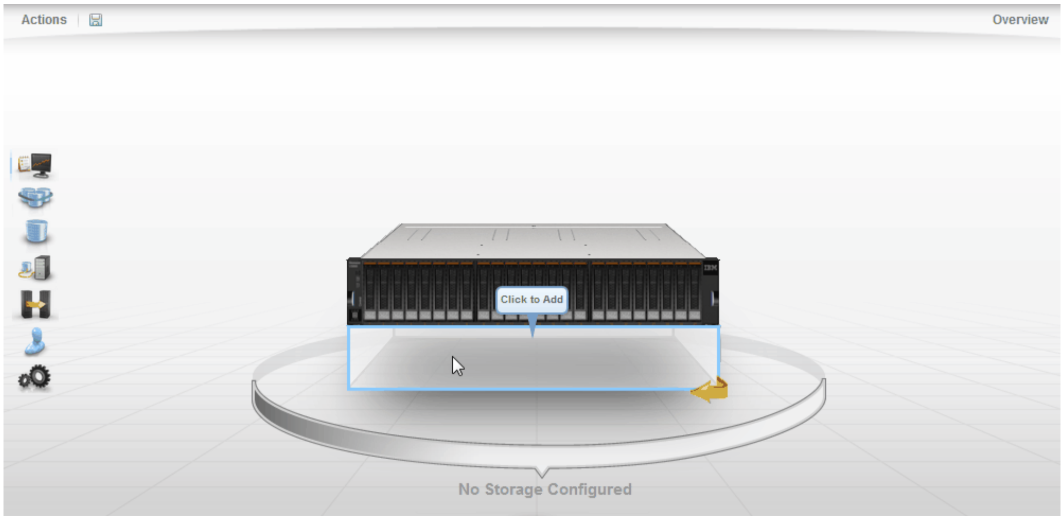
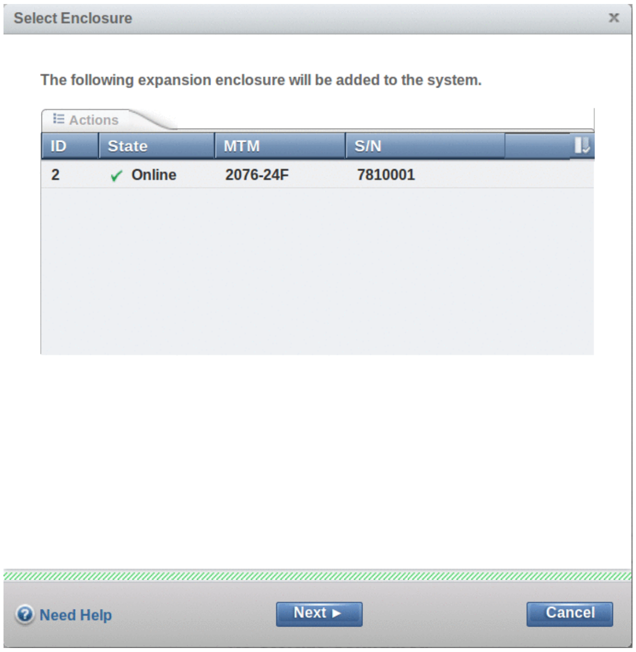


*Figure 4-28 System panel: option to add an expansion enclosure*

2. The first dialog displays the available expansion enclosure and its details, as shown in Figure 4-29 on page 106. Click **Next**.



The following expansion enclosure will be added to the system.

| ID | State | MTM | S/N |
|---|---|---|---|
| 2 | ✓ Online | 2076-24F | 7810001 |

*Figure 4-29 Adding an expansion enclosure to the system*

3. Review the summary in the next dialog and press **Finish** to add the expansion enclosure to the system.

The System panel now displays the new expansion enclosure, as shown in Figure 4-30 on page 107.

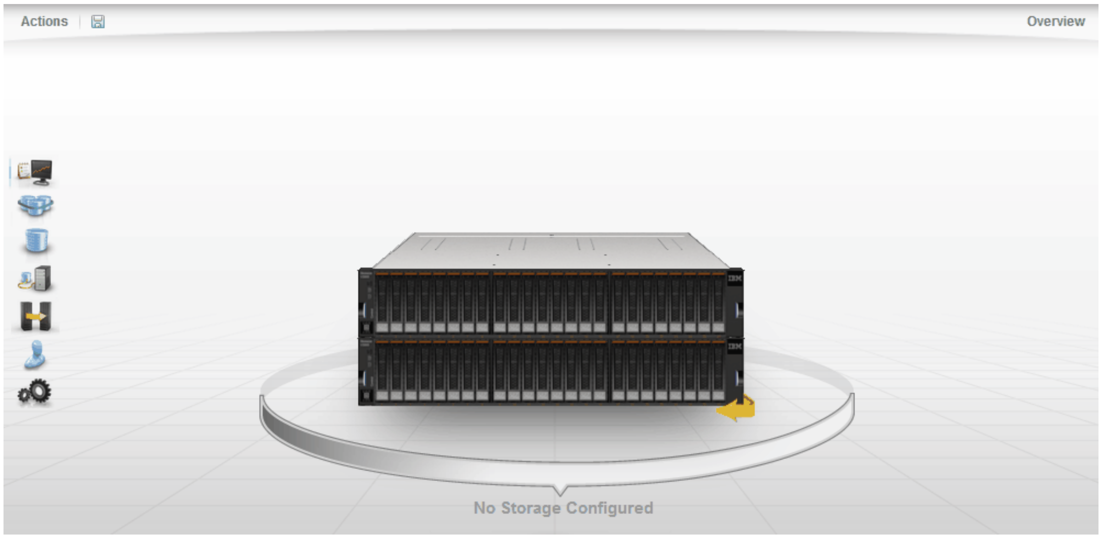


Figure 4-30 System panel: one control enclosure and one expansion installed

If more than one expansion is installed then the expansions are hidden. Figure 4-31 shows the System panel for a system with one control enclosure and two expansions.

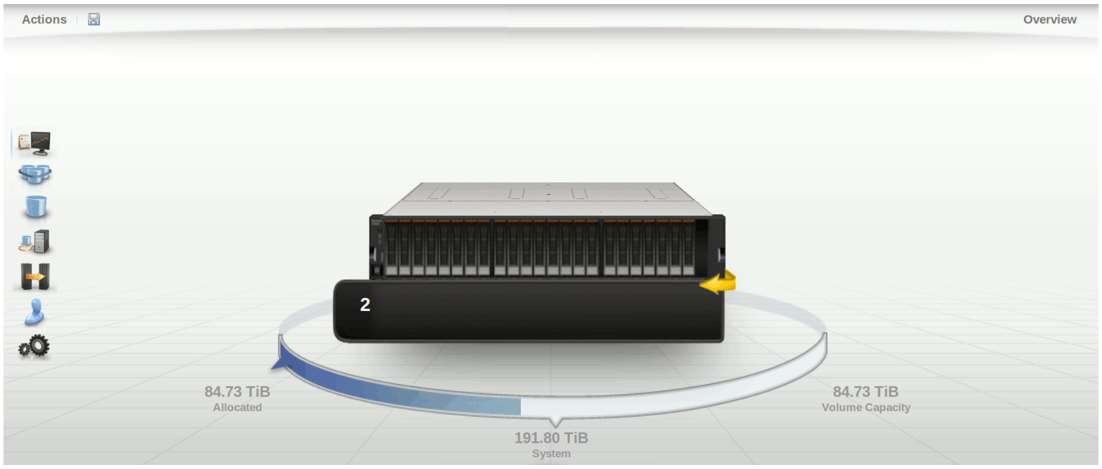


Figure 4-31 System panel: one control enclosure and two expansions installed

Click the element under the control enclosure to see all the expansions, as shown in Figure 4-32 on page 108.

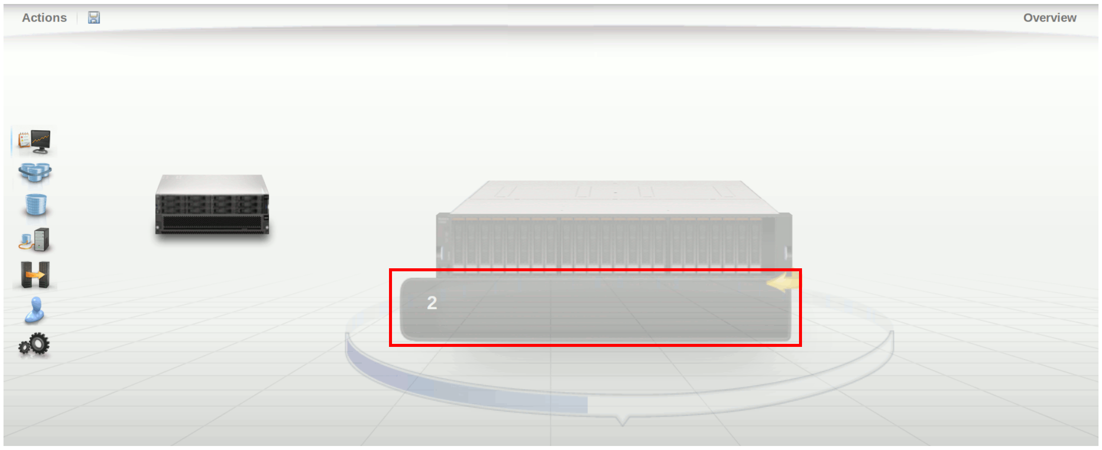


Figure 4-32 System panel: two expansions in detail

## 4.4 Configuring user authentication

There are two methods of user authentication to control access to the GUI and to the CLI:

- *Local authentication* is performed within the Storwize V7000 system. Local GUI authentication is done with user name and password. Local CLI authentication is done either with SSH key or user name and password.
- *Remote authentication* allows users to authenticate to the system using credentials stored on an external authentication service. This means you can use the passwords and user groups defined on the remote service to simplify user management and access, to enforce password policies more efficiently, and to separate user management from storage management.

Locally administered users can coexist with remote authentication.

### 4.4.1 Default superuser account

Every system has a default user called *superuser*. Superuser cannot be deleted or manipulated, except for its password and SSH key. Superuser is a *local* user and cannot be authenticated remotely.

> **Note:** Superuser is the only user allowed to log in to the Service Assistant Tool. It is also the only user allowed to run **`sainfo`** and **`satask`** commands through the CLI.

Superuser is a member of the SecurityAdmin user group, which is the most privileged role within the system.

The password for superuser is set by the user during system setup. The superuser password can be reset to its default value of `passw0rd` using the technician port, for example.

### 4.4.2 Local authentication

A *local user* is managed entirely on the system. A local user belongs to one user group only and it must have a password, or an SSH public key, or both. Each user has a name, which must be unique across all users in one system.

User names can contain up to 256 printable American Standard Code for Information Interchange characters. Forbidden characters are the single quotation mark ('), colon (:), percent symbol (%), asterisk (*), comma (,), and double quotation marks ("). A user name cannot begin or end with a blank space.

Passwords for local users can be up to 64 printable ASCII characters. There are no forbidden characters; however, passwords cannot begin or end with blanks.

Key authentication is attempted first with the password as a fallback. The password and the SSH key are used for CLI or file transfer access. For GUI access, only the password is used.

**Note:** Local users are created for each Storwize V7000 system. If you want to allow access for a user on multiple systems, you must define the user in each system with the same name and the same privileges.

## 4.4.3 Remote authentication

A *remote user* is authenticated on a remote service with either IBM Rational Jazz for Service Management (JazzSM), or Lightweight Directory Access Protocol (LDAP) servers.

**Note:** Remote authentication can be configured to use either JazzSM or LDAP, but not both at the same time.

**Note:** Users authenticated through a LDAP server can log in to the management GUI and the CLI without being configured on the system. However, users authenticated through JazzSM can only log in to the GUI. If CLI access is required, you must create the users on the system with the same password and SSH key set both on the system and on the authentication service. These users must also be enabled for remote authentication. For more information about this topic, see the following website:

https://ibm.biz/BdsKSu

### Configuring remote authentication with LDAP

Storwize V7000 supports three types of LDAP servers:

- IBM Security Directory Server
- Microsoft Active Directory
- OpenLDAP

Users that are authenticated by an LDAP server can log in to the management GUI and the CLI. Unlike remote authentication through JazzSM, users do not need to be configured locally for CLI access. An SSH key is not required for CLI login in this scenario either.

If multiple LDAP servers are available, you can assign multiple LDAP servers to improve availability. Authentication requests are processed by those LDAP servers that are marked as preferred unless the connections fail or a user is not found. Requests are distributed across all preferred servers for load balancing in a round-robin fashion.

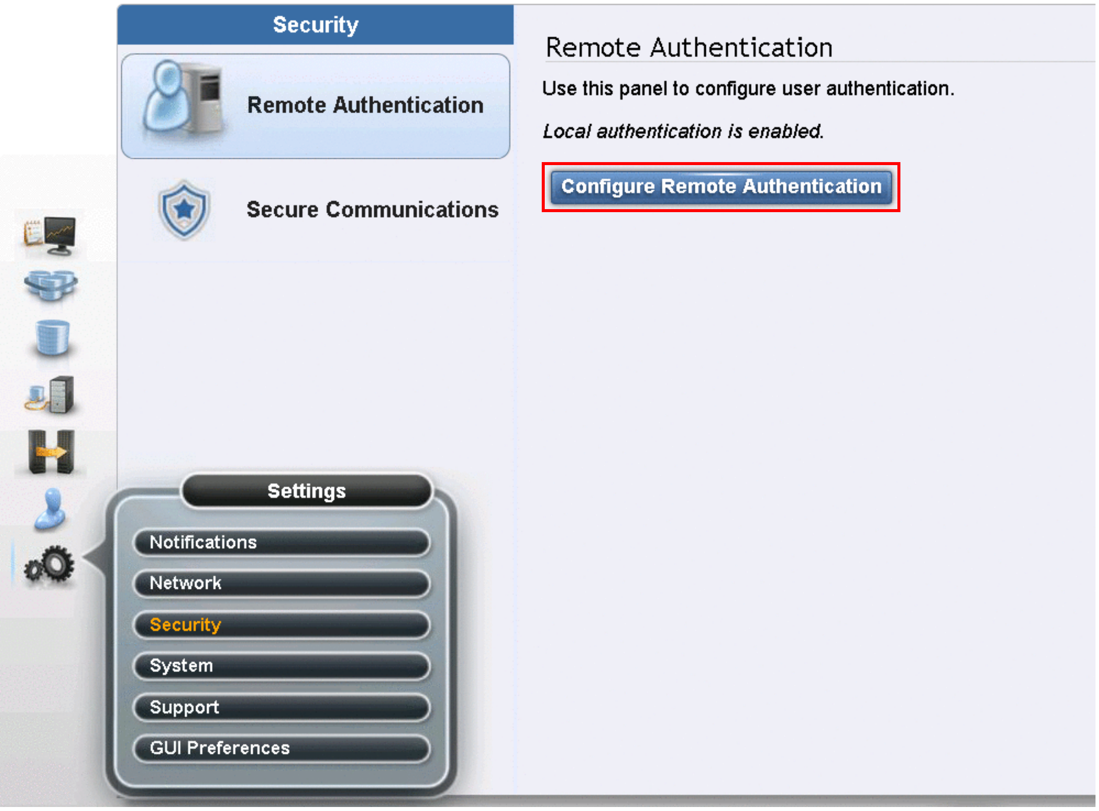
**Note:** All LDAP servers configured within the same system must be of the same type.

If users that are part of a group on the LDAP server are to be authenticated remotely, a user group with an identical name must exist on the system. The user group name is *case-sensitive*. The user group must also be enabled for remote authentication on the system.

A user that is authenticated remotely is granted permissions according to the role that is assigned to the user group of which the user is a member.

To configure remote authentication using LDAP, start by enabling remote authentication. To do so, complete the following steps:

1. Navigate to **Settings** → **Security**, select **Remote Authentication** and then **Configure Remote Authentication**, as shown in Figure 4-33 on page 110.



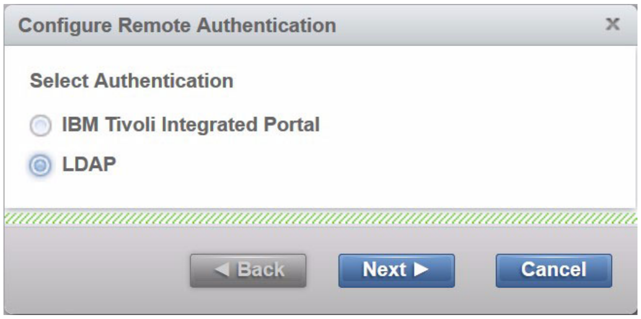
Remote Authentication

Use this panel to configure user authentication.

*Local authentication is enabled.*

Figure 4-33 Configuring remote authentication

2. Select the authentication type. In this example **LDAP** is selected, as shown in Figure 4-34. To configure JazzSM, select **IBM Tivoli Integrated Portal** instead. Click **Next**.


Configure Remote Authentication

Select Authentication

○ IBM Tivoli Integrated Portal

● LDAP

Figure 4-34 Configure remote authentication: authentication type

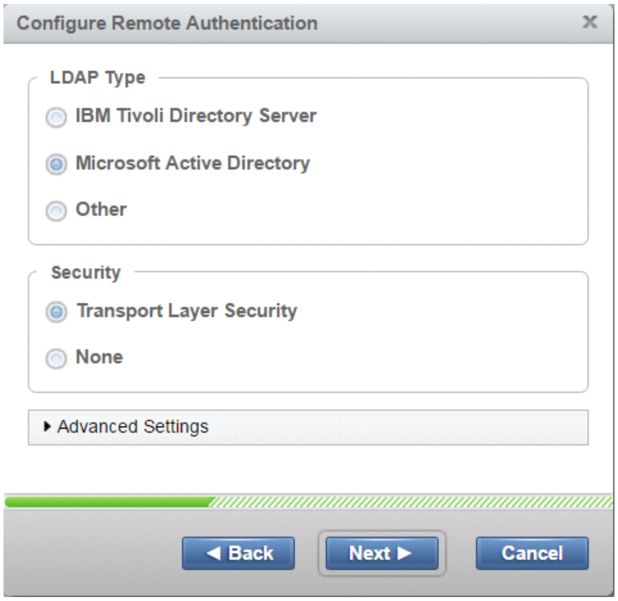
3. Enter the LDAP settings. Note that these settings are not server specific; they are common to every server configured. Extra optional settings are available by clicking **Advanced Settings**. All settings are described below:

   – LDAP type. Choose between:

     - **IBM Tivoli Directory Server** (for IBM Security Directory Server)
     - **Microsoft Active Directory**
     - **Other** (for OpenLDAP)

     In this example we configure a Microsoft Active Directory server, as shown in Figure 4-35 on page 111.

   – Security. Choose between **None** or **Transport Layer Security**. Select the latter to establish secure connections, ensuring that user credentials are encrypted before being transmitted. In this example we choose Transport Layer Security, as shown in Figure 4-35.



Configure Remote Authentication

LDAP Type

○ IBM Tivoli Directory Server

● Microsoft Active Directory

○ Other

Security

● Transport Layer Security

○ None

▸ Advanced Settings

*Figure 4-35 Configure remote authentication: mandatory LDAP settings*

   – Service Credentials. This is an advanced and optional setting. Leave **User Name** and **Password** empty if your LDAP server supports anonymous bind. In this example we enter the credentials of an existing user on the LDAP server with permission to query the LDAP directory. You can enter this information in the format of an email address (for example, `administrator@itso.corp`, as shown in Figure 4-36 on page 112) or as a distinguished name (for example, `cn=Administrator,cn=users,dc=itso,dc=corp`).

   – User Attribute. LDAP attribute used to determine the user name of remote users. The attribute must exist in your LDAP schema and must be unique for each of your users.

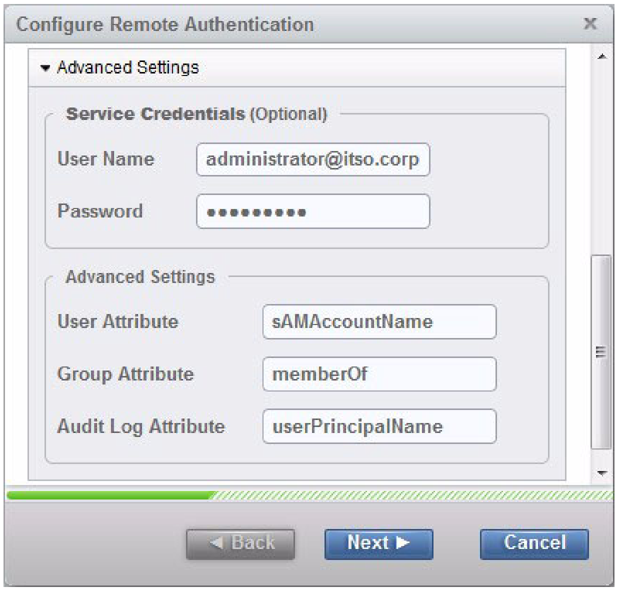
     This is an advanced setting that defaults to `sAMAaccountName` for Microsoft Active Directory and to `uid` for IBM Security Directory Server and OpenLDAP.

   – Group Attribute. LDAP attribute used to determine the user group memberships of remote users. The attribute must contain either the distinguished name of a group or a colon-separated list of group names.

     This is an advanced setting that defaults to `memberOf` for Microsoft Active Directory and OpenLDAP and to `ibm-allGroups` for IBM Security Directory Server. For OpenLDAP implementations, you might need to configure the `memberOf` overlay if it is not in place.

– Audit Log Attribute. LDAP attribute used to determine the identity of remote users. When an LDAP user performs an audited action, this identity is recorded in the audit log.

  This is an advanced setting that defaults to `userPrincipalName` for Microsoft Active Directory and to `uid` for IBM Security Directory Server and OpenLDAP.


Configure Remote Authentication

Advanced Settings

Service Credentials (Optional)

User Name: administrator@itso.corp

Password: ●●●●●●●●●

Advanced Settings

User Attribute: sAMAccountName

Group Attribute: memberOf

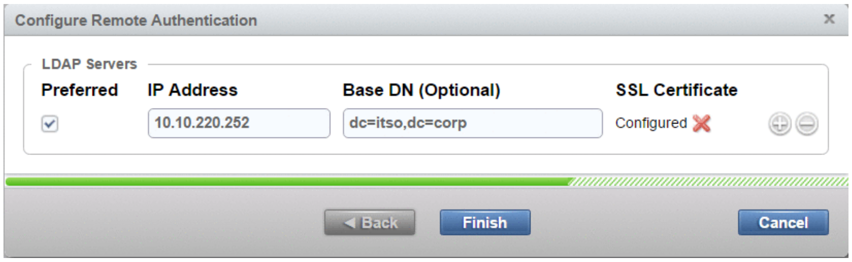
Audit Log Attribute: userPrincipalName

*Figure 4-36 Configure remote authentication: advanced LDAP settings*

4. Enter the server settings for one or more LDAP servers, as shown in Figure 4-37. To add more servers click the plus (+) icon. All settings are described below:

- Preferred. Authentication requests are processed by the preferred servers unless the connections fail or a user is not found. Requests are distributed across all preferred servers for load balancing. Check **Preferred** to set the server as a preferred server.
- IP Address. IP address of the server.
- Base DN. Distinguished name to use as a starting point for searching for users on the server (for example, `dc=itso,dc=corp`).
- SSL Certificate. Secure Sockets Layer certificate used to securely connect to the LDAP server. This is required only if you chose to use Transport Layer Security as a security method earlier.

Click **Finish** to save the settings.


Configure Remote Authentication

LDAP Servers

| Preferred | IP Address | Base DN (Optional) | SSL Certificate |
|---|---|---|---|
| ☑ | 10.10.220.252 | dc=itso,dc=corp | Configured ✗ |

*Figure 4-37 Configure remote authentication: creating an LDAP server*

Now that remote authentication is enabled, the remote user groups must be configured. You can use the default built-in user groups for remote authentication. However, remember that the name of the default user groups cannot be changed. If the LDAP server already contains a group that you want to use, the name of the group must be changed on the server side to match the default name. Any user group, whether default or self-defined, must be enabled for remote authentication.

Complete the following steps to create a new user group with remote authentication enabled:

1. Navigate to **Access** → **Users** and select **Create User Group**, as shown in Figure 4-38.

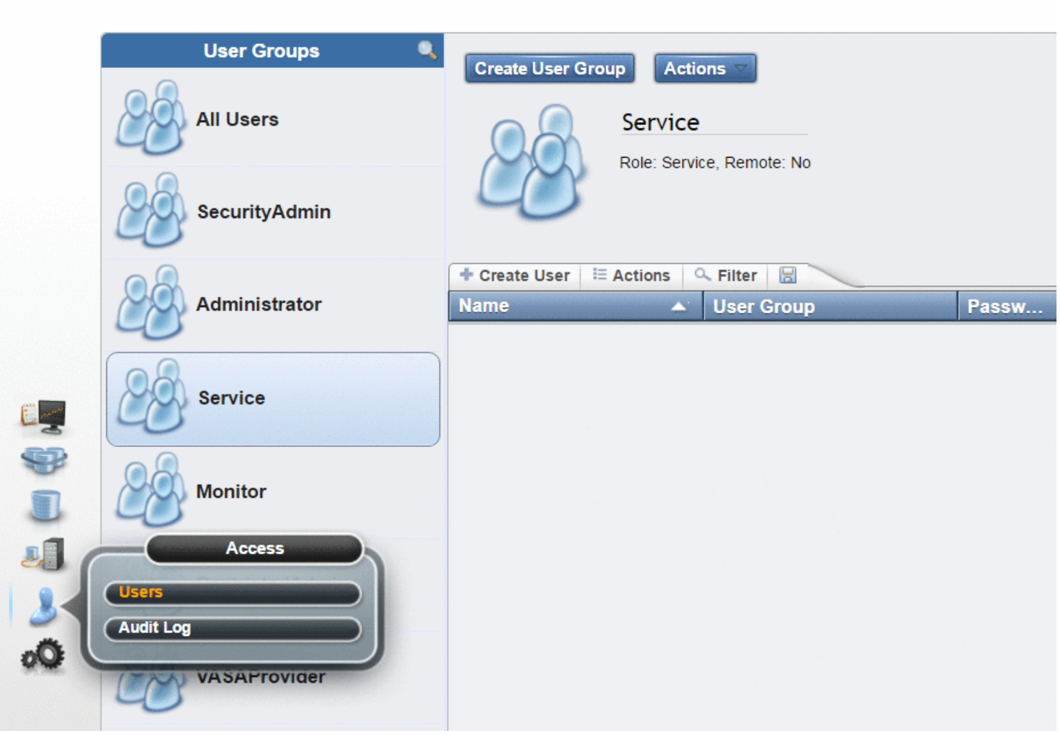
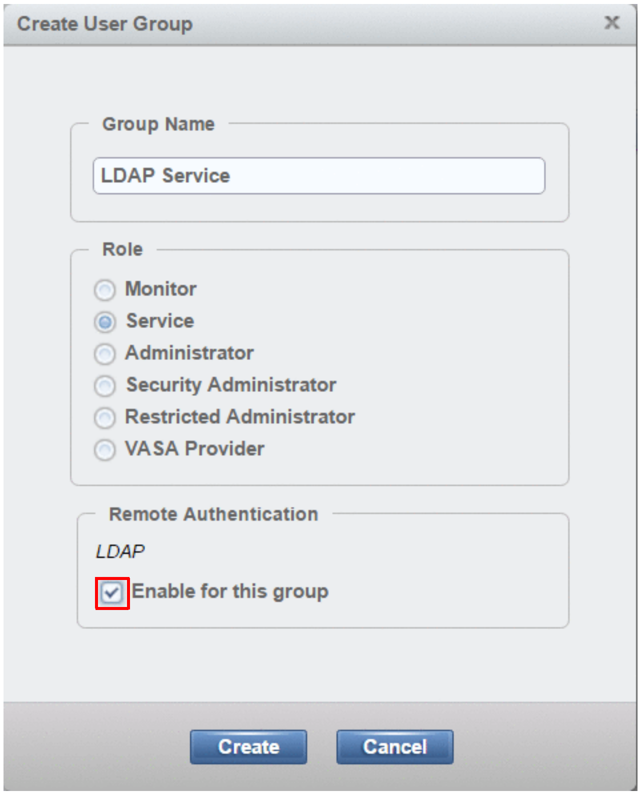


*Figure 4-38 Option to create a new user group*

2. Enter the details for the new group. Check **Enable for this group** to enable remote authentication, as shown in Figure 4-39. Click **Create**.

> **Note:** This option is not available if LDAP is not enabled yet.



Group Name: LDAP Service

Role

○ Monitor
● Service
○ Administrator
○ Security Administrator
○ Restricted Administrator
○ VASA Provider

Remote Authentication

*LDAP*

☑ Enable for this group

*Figure 4-39 Creating a user group with remote authentication enabled*

Complete the following steps to enable remote authentication for a default role:

1. Navigate to **Access** → **Users**. Select the user group you want to modify, click **Actions**, and then **Properties**. In this example the Service user group is chosen, as shown in Figure 4-40.

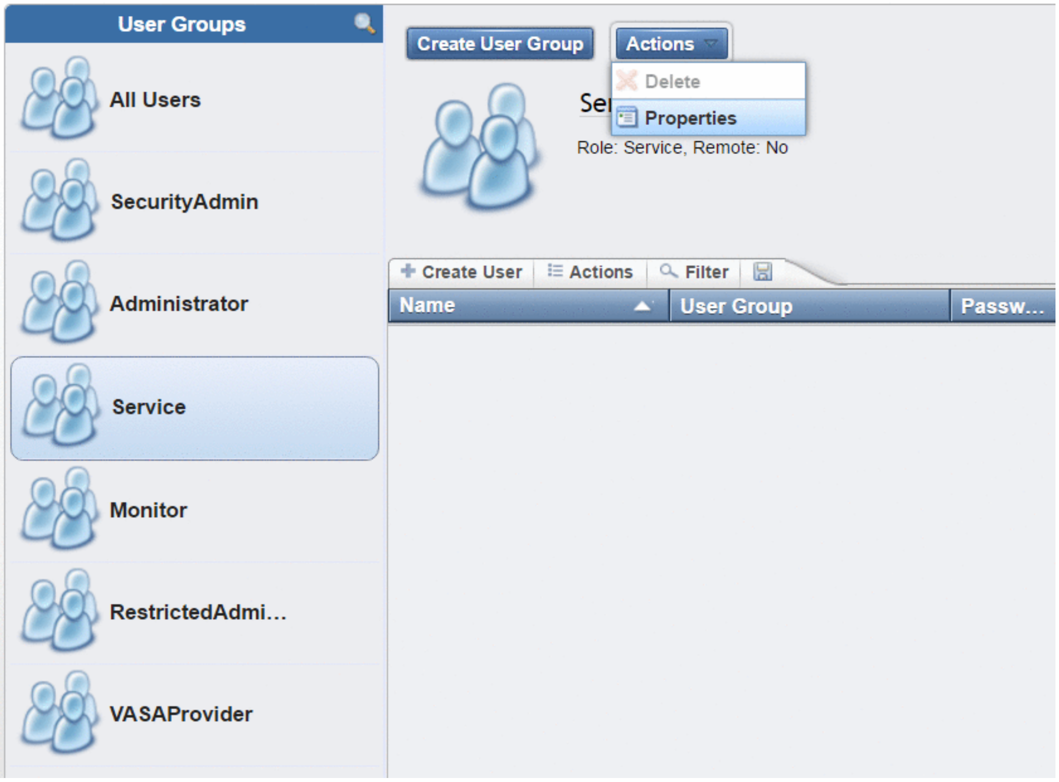
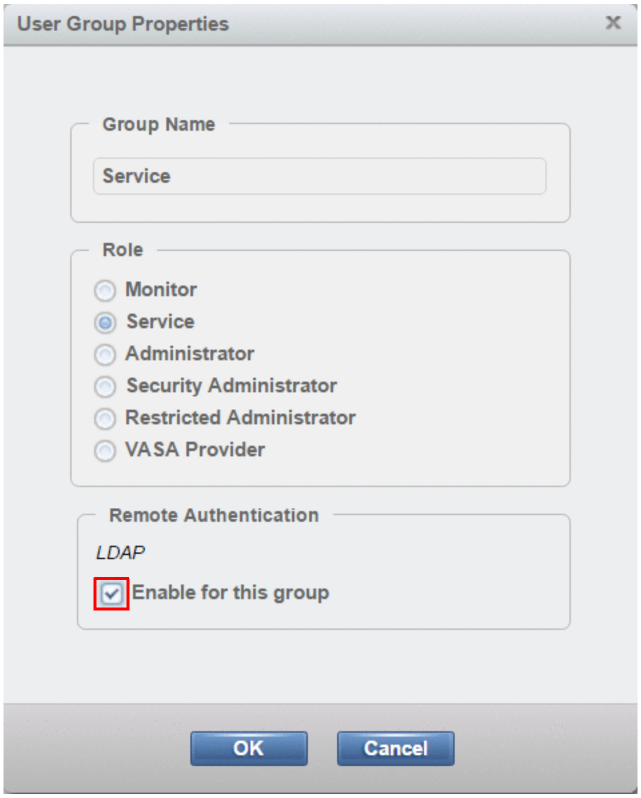


*Figure 4-40 Changing the properties of a user group*

2. Check **Enable for this group**, as shown in Figure 4-41.

**Note:** This option is not available if LDAP is not enabled yet.



User Group Properties

Group Name: Service

Role
- ○ Monitor
- ● Service
- ○ Administrator
- ○ Security Administrator
- ○ Restricted Administrator
- ○ VASA Provider

Remote Authentication

*LDAP*

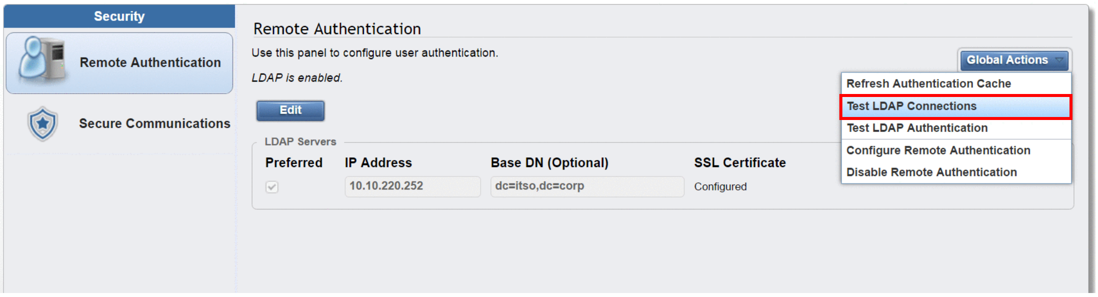
☑ Enable for this group

*Figure 4-41 Enabling remote authentication for a default group*

When you have at least one user group enabled for remote authentication, make sure that the LDAP server is configured correctly by verifying the following conditions are true:

- The name of the user group on the LDAP server matches the one you just modified or created.
- Each user that you want to authenticate remotely is a member of the appropriate user group for the intended system role.

The system is now ready to authenticate users using the LDAP server. To ensure that everything works correctly, navigate to **Settings** → **Security**, select **Global Actions** and then **Test LDAP Connections**, as shown in Figure 4-42.



Remote Authentication

Use this panel to configure user authentication.

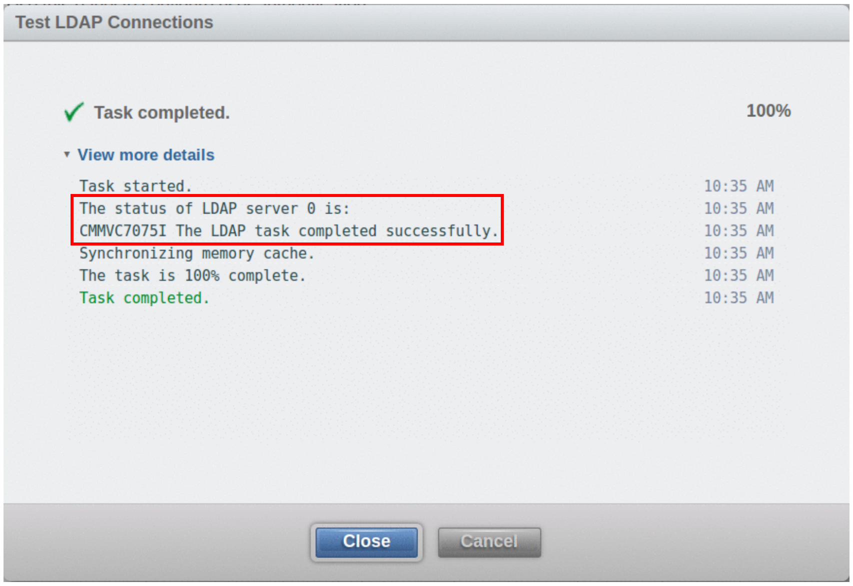
*LDAP is enabled.*

LDAP Servers

| Preferred | IP Address | Base DN (Optional) | SSL Certificate |
|---|---|---|---|
| ☑ | 10.10.220.252 | dc=itso,dc=corp | Configured |

*Figure 4-42 Option to test LDAP connections*

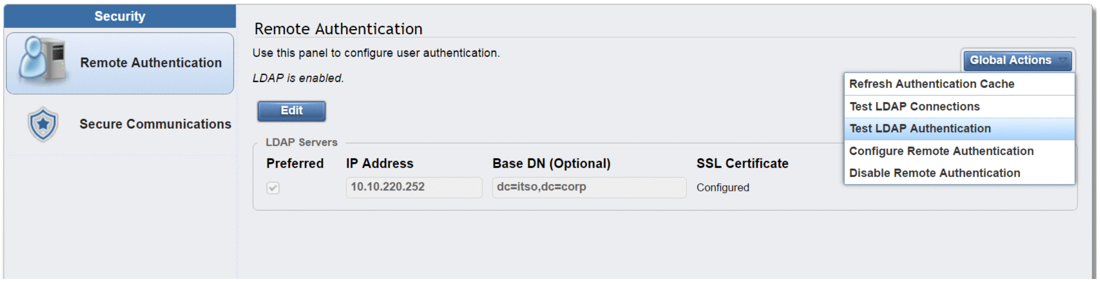
Figure 4-43 shows the result of a successful connection test. If the connection is not successful, an error is logged in the event log.



```
Task started.                                          10:35 AM
The status of LDAP server 0 is:                        10:35 AM
CMMVC7075I The LDAP task completed successfully.       10:35 AM
Synchronizing memory cache.                            10:35 AM
The task is 100% complete.                             10:35 AM
Task completed.                                        10:35 AM
```

*Figure 4-43 Successful LDAP connection test*

There is also the option to test a real user authentication attempt. Navigate to **Settings** → **Security**, select **Global Actions** and then **Test LDAP Authentication**, as shown in Figure 4-44.



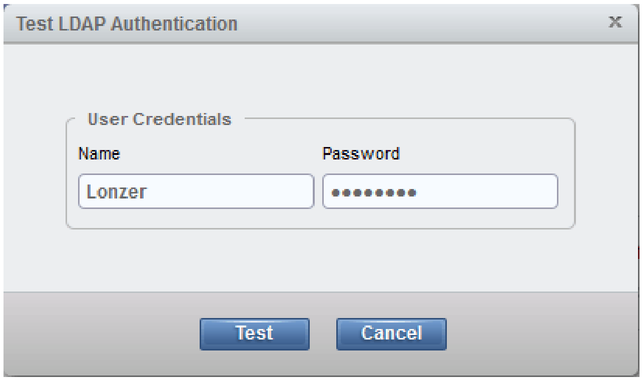
Use this panel to configure user authentication.

*LDAP is enabled.*

| Preferred | IP Address | Base DN (Optional) | SSL Certificate |
|---|---|---|---|
| ☑ | 10.10.220.252 | dc=itso,dc=corp | Configured |

*Figure 4-44 Option to test LDAP authentication*

Enter the user credentials of a user defined on the LDAP server, as shown in Figure 4-45. Click **Test**.



Name: Lonzer

Password: ••••••••

*Figure 4-45 LDAP authentication test*

Once again, the message `CMMVC70751 The LDAP task completed successfully` is shown after a successful test.

Both the connection test and the authentication test must complete successfully to ensure that LDAP authentication works correctly. Assuming both tests succeed, users can log in to the GUI and CLI using their network credentials.

A user can log in with their short name (that is, without the domain component) or with the fully qualified user name inca the form of an email address.

### 4.4.4 User groups and roles

User groups are used for local and remote authentication. Each user group is associated with a single *role*. The role for a user group cannot be changed, but user groups (with one of the defined roles) can be created.

The rights of a user who belongs to a specific user group are defined by the role that is assigned to the user group. It is the role that defines what a user can or cannot do on the system.

Storwize V7000 provides six user groups and seven roles by default, as shown in Table 4-1. The VasaProvider role is not associated with a default user group.

**Note:** The VasaProvider role is used to allow VMware to interact with the system when implementing Virtual Volumes. It is advised not to use this role for users not controlled by VMware.

*Table 4-1 Default user groups and roles*

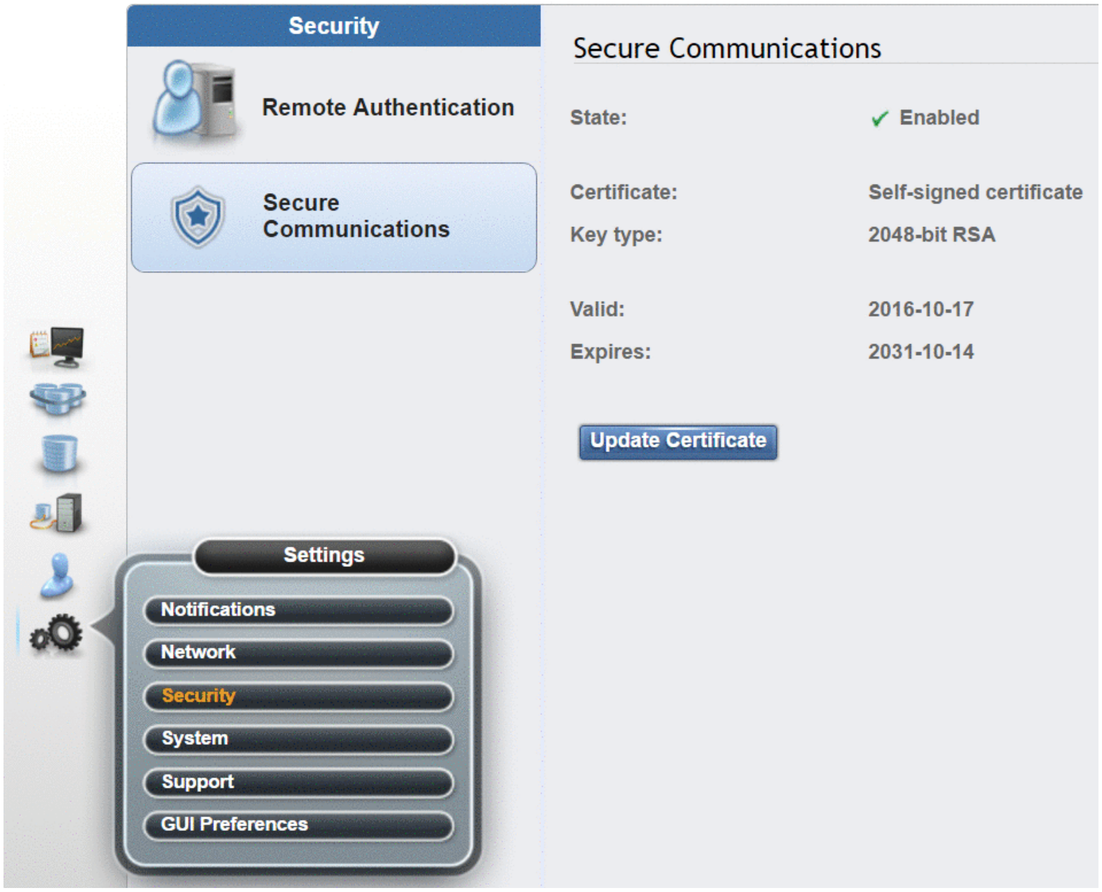
| User group | Role |
|---|---|
| SecurityAdmin | SecurityAdmin |
| Administrator | Administrator |
| CopyOperator | CopyOperator |
| Service | Service |
| Monitor | Monitor |
| RestrictedAdmin | RestrictedAdmin |
| - | VasaProvider |

## 4.5 Configuring secure communications

During system initialization a *self-signed* SSL certificate is automatically generated by the system to encrypt communications between the browser and the system. Self-signed certificates generate web browser security warnings and might not comply with organizational security guidelines.

*Signed* SSL certificates are issued by a third-party certificate authority. A browser maintains a list of trusted certificate authorities, identified by their *root* certificate. The root certificate must be included in this list in order for the signed certificate to be trusted; if it is not, the browser presents security warnings.

To see the details of your current system certificate navigate to **Settings** → **Security** and select **Secure Communications**, as shown in Figure 4-46 on page 119.



Secure Communications

State: ✓ Enabled

Certificate: Self-signed certificate

Key type: 2048-bit RSA

Valid: 2016-10-17

Expires: 2031-10-14

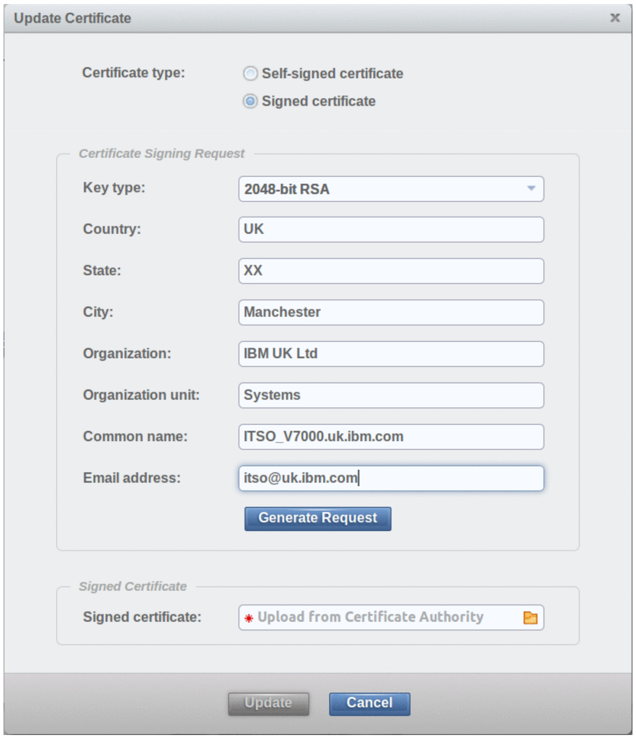
*Figure 4-46 Accessing the Secure Communications panel*

Storwize V7000 allows you to generate a new self-signed certificate or configure a signed certificate.

## 4.5.1 Configuring a signed certificate

Complete the following steps to configure a signed certificate:

1. Select **Update Certificate** on the Secure Communications panel.
2. Select **Signed certificate** and enter the details for the new certificate signing request. All fields are mandatory except for the email address. Figure 4-47 on page 120 shows some values as an example.



Certificate type: ○ Self-signed certificate ● Signed certificate

Certificate Signing Request

| | |
|---|---|
| Key type: | 2048-bit RSA |
| Country: | UK |
| State: | XX |
| City: | Manchester |
| Organization: | IBM UK Ltd |
| Organization unit: | Systems |
| Common name: | ITSO_V7000.uk.ibm.com |
| Email address: | itso@uk.ibm.com |

Signed Certificate

Signed certificate: * Upload from Certificate Authority

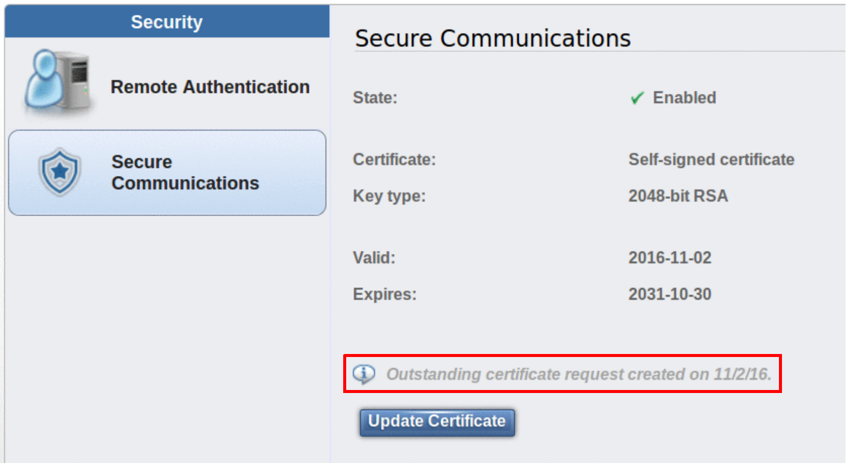
*Figure 4-47 Generating a certificate request*

> **Attention:** Before generating a request, ensure that your current browser does not have restrictions on the type of keys used for certificates. Some browsers limit the use of specific key-types for security and compatibility issues.

Click **Generate Request**.

3. Save the generated request file. The Secure Communications panel now mentions that there is an outstanding certificate request, as shown in Figure 4-48. This is the case until the associated signed certificate is installed.

> **Attention:** If you need to update a field in the certificate request you can generate a new request. However, do *not* generate a new request after sending the original one to the certificate authority. Generating a new request overrides the original one and the signed certificate associated with the original *cannot* be installed.



Secure Communications

State: ✓ Enabled

Certificate: Self-signed certificate

Key type: 2048-bit RSA

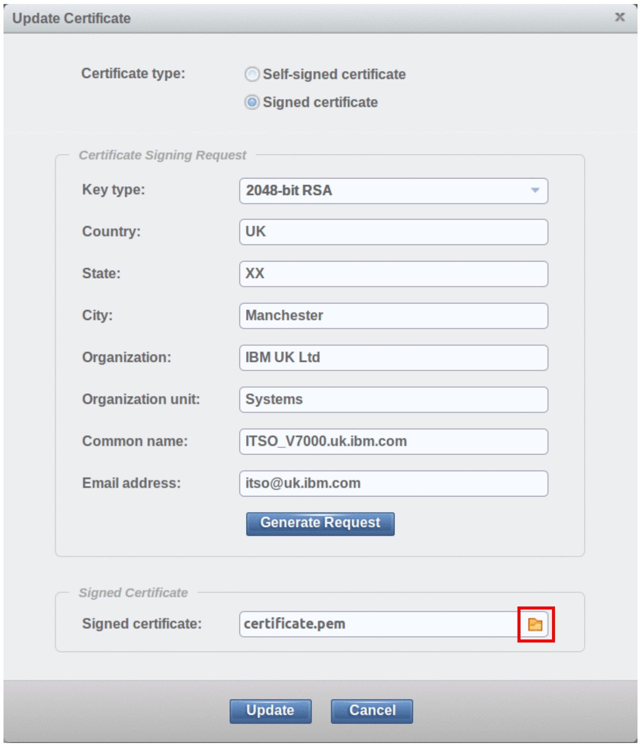
Valid: 2016-11-02

Expires: 2031-10-30

Outstanding certificate request created on 11/2/16.

*Figure 4-48 Outstanding certificate request*

4. Submit the request to the certificate authority to receive a signed certificate.
5. When you receive the signed certificate, select **Update Certificate** on the Secure Communications panel once again.
6. Click the folder icon to upload the signed certificate, as shown in Figure 4-49. Click **Update**.



Certificate type: ○ Self-signed certificate ● Signed certificate

Certificate Signing Request

Key type: 2048-bit RSA

Country: UK

State: XX

City: Manchester

Organization: IBM UK Ltd

Organization unit: Systems

Common name: ITSO_V7000.uk.ibm.com

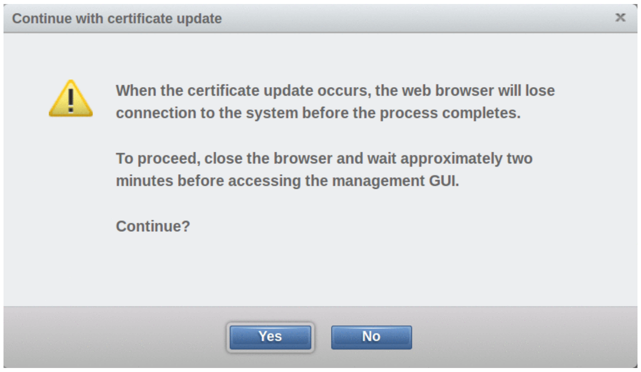
Email address: itso@uk.ibm.com

Signed Certificate

Signed certificate: certificate.pem

*Figure 4-49 Installing a signed certificate*

7. You are prompted to confirm the action, as shown in Figure 4-50. Click **Yes** to proceed. The signed certificate is installed.



Continue with certificate update

When the certificate update occurs, the web browser will lose connection to the system before the process completes.

To proceed, close the browser and wait approximately two minutes before accessing the management GUI.

Continue?

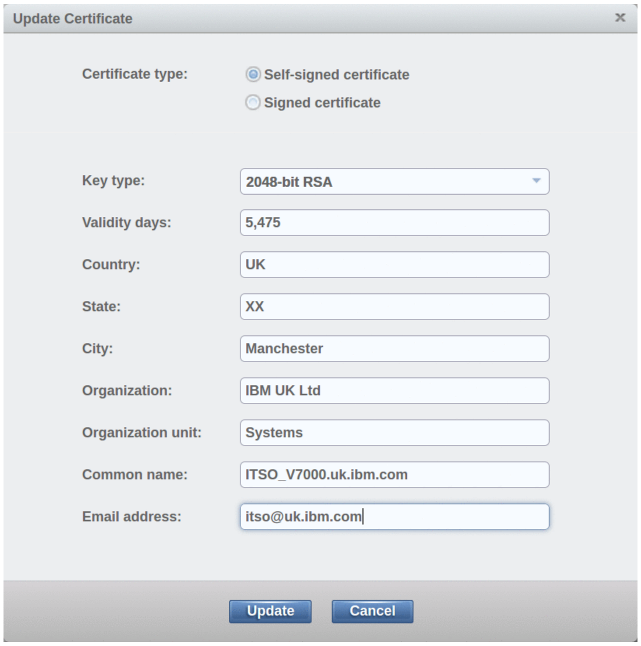
*Figure 4-50 Certificate update warning*

## 4.5.2 Generating a self-signed certificate

Complete the following steps to generate a self-signed certificate:

1. Select **Update Certificate** on the Secure Communications panel.
2. Select **Self-signed certificate** and enter the details for the new certificate. Key type and validity days are the only mandatory fields. Figure 4-51 shows some values as an example.

> **Attention:** Before creating a new self-signed certificate, ensure that your current browser does not have restrictions on the type of keys used for certificates. Some browsers limit the use of specific key-types for security and compatibility issues.



Update Certificate

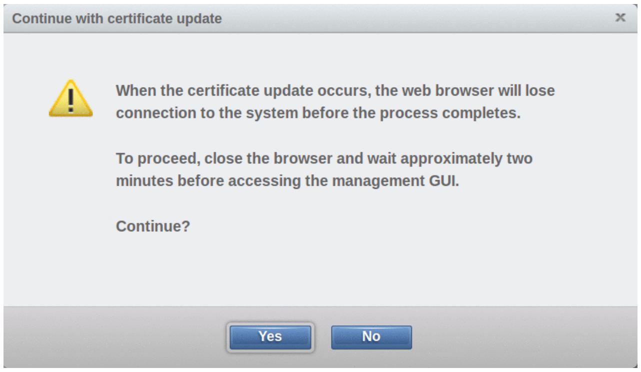
Certificate type: ● Self-signed certificate ○ Signed certificate

|  |  |
|---|---|
| Key type: | 2048-bit RSA |
| Validity days: | 5,475 |
| Country: | UK |
| State: | XX |
| City: | Manchester |
| Organization: | IBM UK Ltd |
| Organization unit: | Systems |
| Common name: | ITSO_V7000.uk.ibm.com |
| Email address: | itso@uk.ibm.com |

*Figure 4-51 Generating a new self-signed certificate*

Click **Update**.

3. You are prompted to confirm the action, as shown in Figure 4-52 on page 123. Click **Yes** to proceed. The self-signed is generated immediately.



Continue with certificate update

When the certificate update occurs, the web browser will lose connection to the system before the process completes.

To proceed, close the browser and wait approximately two minutes before accessing the management GUI.

Continue?

*Figure 4-52 Certificate update warning*

# 5

# Graphical User Interface

This chapter describes an overview of the IBM Spectrum Virtualize graphical user interface (GUI). The management GUI is a tool enabled and provided by IBM Spectrum Virtualize that helps you to monitor, manage, and configure your system.

This chapter explains the basic view and configuration procedures that are required to get your IBM Spectrum Virtualize environment running as quickly as possible by using the GUI.

This chapter does not describe advanced troubleshooting or problem determination or some of the complex operations (compression, encryption), because they are explained later in this publication.

This chapter includes the following topics:

- Normal operations using the GUI
- Introduction to the GUI
- Viewing Cluster settings using the GUI
- System overview
- Setting GUI preferences

## 5.1 Normal operations using the GUI

In this section, we describe useful tasks using the GUI, that help administrators to manage, monitor and configure the IBM Storwize V7000 as quick as possible. For illustration, we configured the IBM Storwize V7000 in a standard topology.

As mentioned, the GUI is a built-in software component within the IBM Spectrum Virtualize software.

Multiple users can be logged in to the GUI at any time. However, no locking mechanism exists, so be aware that if two users change the same object at the same time, the last action that is entered from the GUI is the one that takes effect.

> **Important:** Data entries that are made through the GUI are case-sensitive.

## 5.2 Introduction to the GUI

As shown in Figure 5-1, the IBM Storwize V7000 GUI System pane is an important user interface. Throughout this chapter, we refer to this interface as the IBM Storwize V7000 or just System pane.

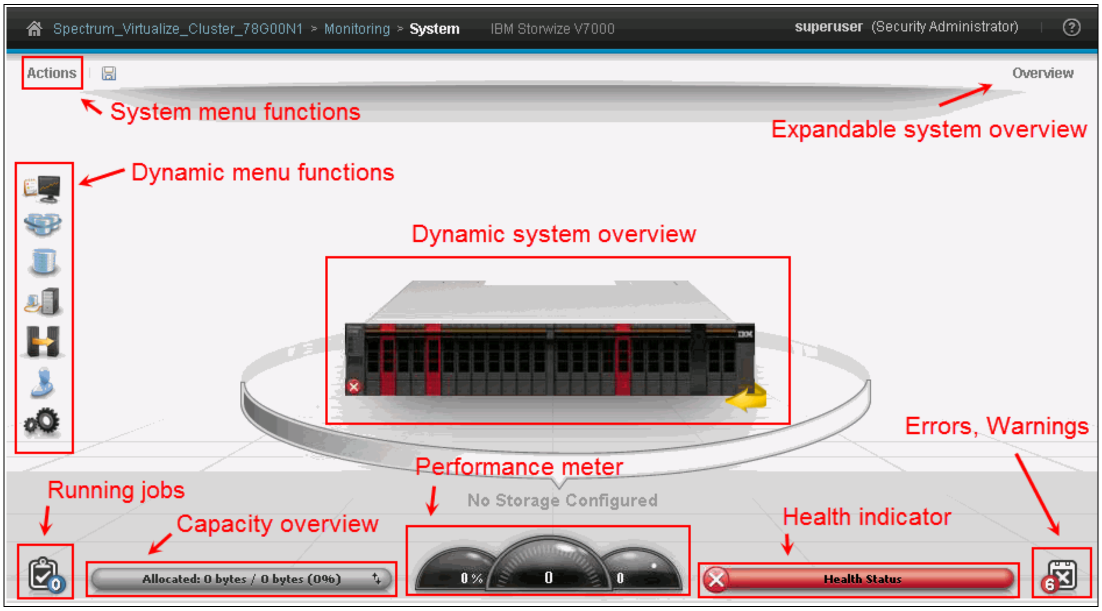


*Figure 5-1 IBM Storwize V7000 GUI*

### 5.2.1 Dynamic menu

From any page in IBM Spectrum Virtualize, you can always access the dynamic menu as it is always uncovered.

The IBM Spectrum Virtualize GUI dynamic menu is on the left side of the GUI window. To browse by using this menu, hover the mouse pointer over the various icons and choose a page that you want to display, as shown in Figure 5-2.

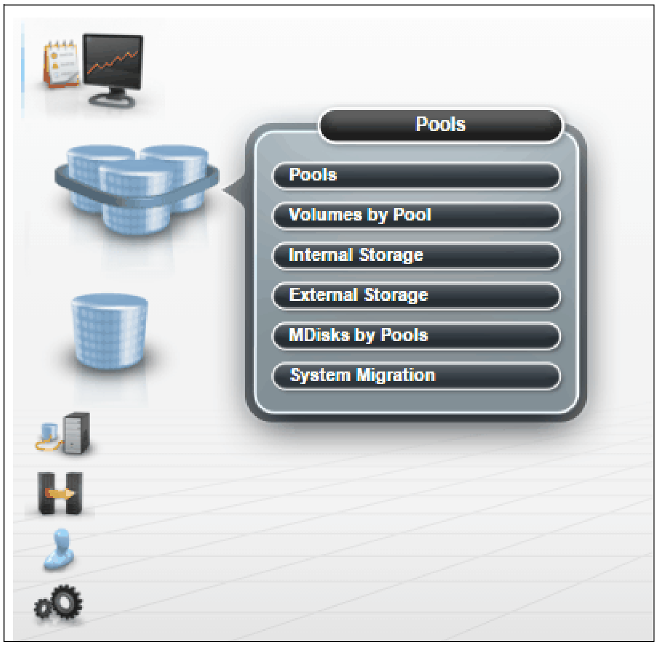


*Figure 5-2 The dynamic menu*

The IBM Spectrum Virtualize dynamic menu consists of multiple panes independently on the underlying hardware (IBM SAN Volume Controller, IBM Storwize family). These panes group common configuration and administration objects and present individual administrative objects to the IBM Spectrum Virtualize GUI users, as shown in Figure 5-3.

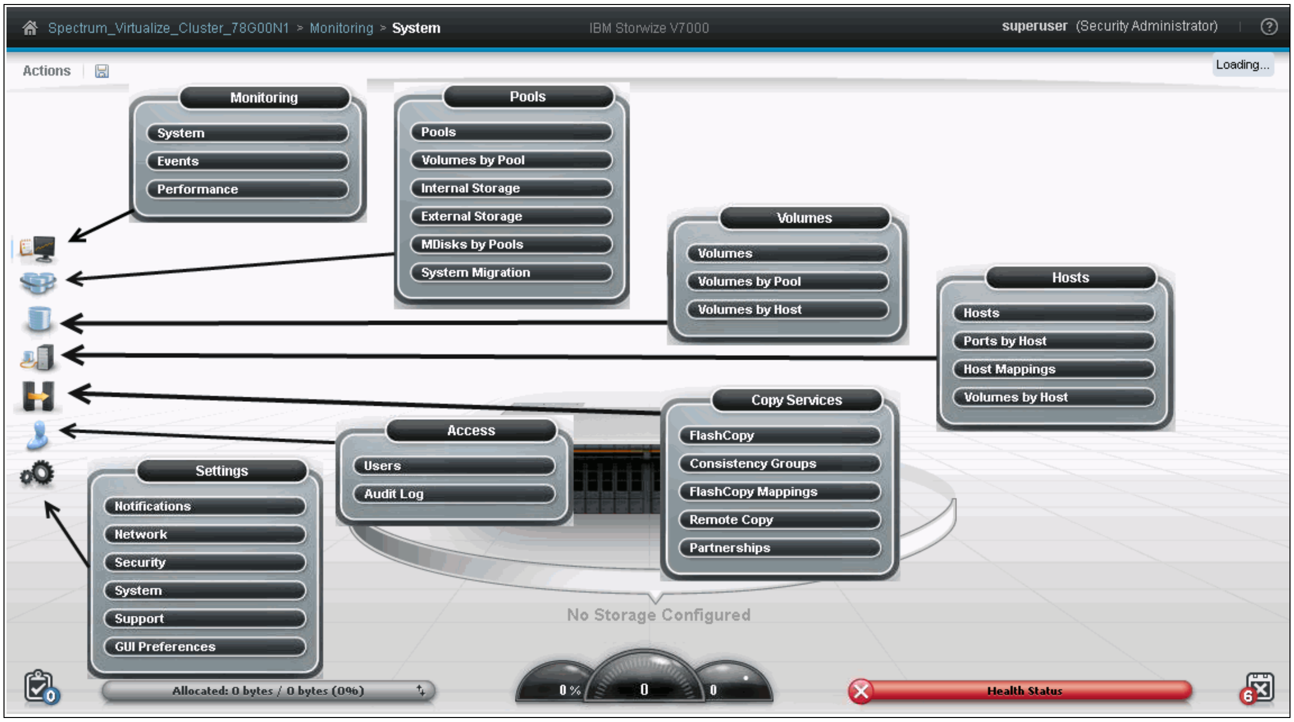
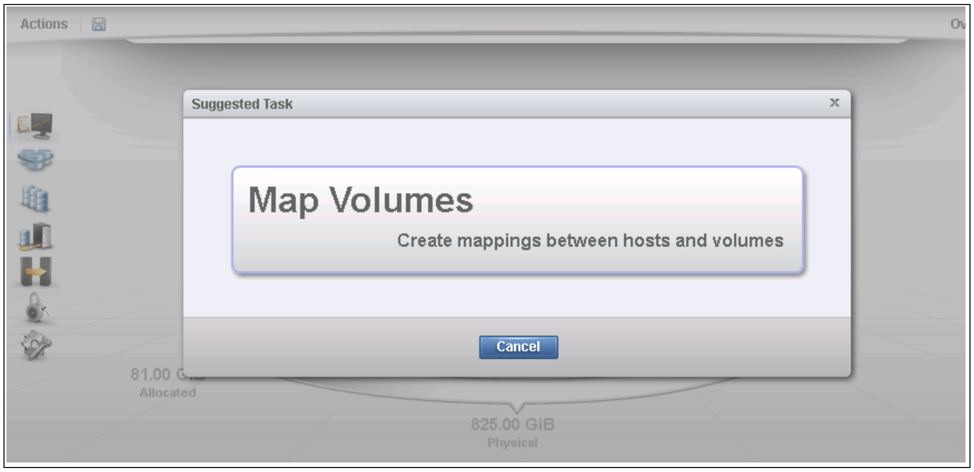


*Figure 5-3 IBM Storwize V7000 GUI*

## 5.2.2 Suggested tasks

After a successful login, IBM Spectrum Virtualize opens a pop-up window with suggested tasks notifying administrators that several key IBM Storwize V7000 functions are not yet configured. If necessary, this window can be closed and the tasks can be performed at any time. Figure 5-4 shows the suggested tasks in the System pane.



Suggested Task

Map Volumes

Create mappings between hosts and volumes

*Figure 5-4 Suggested tasks*

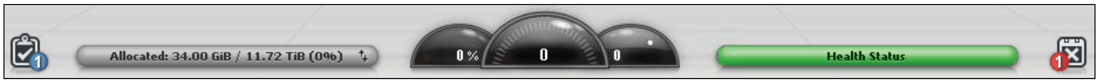
In this case, the GUI warns you that so far no volume is mapped to the host, or that no host is defined yet. You can directly perform the task from this window or cancel it and run the procedure later at any convenient time. Other suggested tasks that typically appear after the initial system configuration are to create a volume and configure a storage pool, for example.

The dynamic IBM Spectrum Virtualize menu contains the following panes (Figure 5-3 on page 127):

- Monitoring
- Pools
- Volumes
- Hosts
- Copy Services
- Access
- Settings

## 5.2.3 Notification status area

A control pane is available in the bottom part of the window. This pane is divided into five status areas to provide information about your system. These persistent state notification widgets are reduced, by default, as shown in Figure 5-5. The notification area contains two status bar indicators (capacity and system health), a performance meter, a running tasks indicator, and an alerts indicator.



Allocated: 34.00 GiB / 11.72 TiB (0%)

*Figure 5-5 Notification area*

### Health status and alerts indication

The rightmost area of the control pane provides information about the health status of the system and alerts logged in your system, as shown in Figure 5-6.

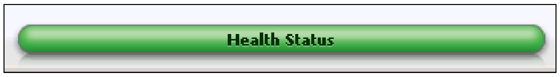
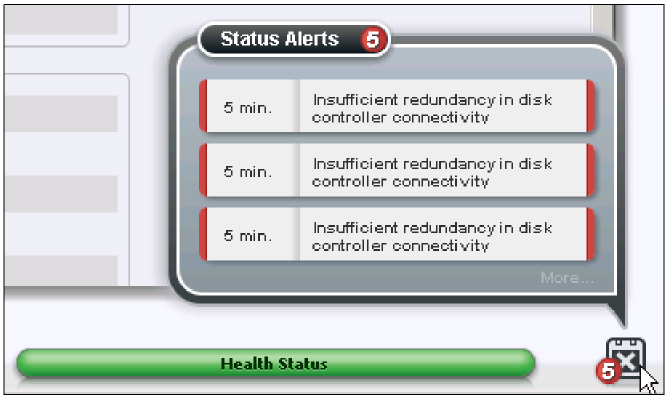


*Figure 5-6 Health Status area*

If non-critical issues exist for your system nodes, external storage controllers, or remote partnerships, a new status area opens next to the Health Status widget, as shown in Figure 5-7.



5 min. Insufficient redundancy in disk controller connectivity

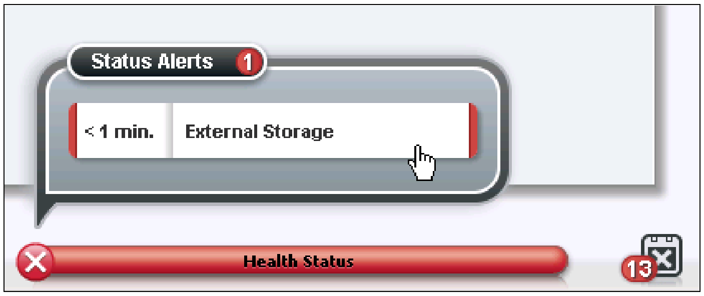
5 min. Insufficient redundancy in disk controller connectivity

5 min. Insufficient redundancy in disk controller connectivity

*Figure 5-7 Controller path status alert*

You can fix the error by clicking **Status Alerts** to direct you to the Events pane fix procedures.

If a critical system connectivity error exists, the Health Status bar turns red and alerts the system administrator for immediate action, as shown in Figure 5-8.

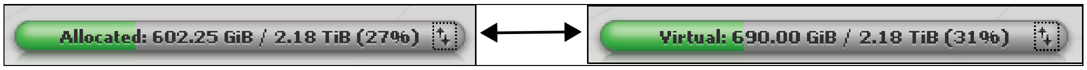


< 1 min. External Storage

*Figure 5-8 External storage connectivity loss*

### Storage allocation indicator

The leftmost indicator shows information about the overall physical capacity (the initial amount of storage that was allocated). This indicator also shows the virtual capacity (thin-provisioned storage). The virtual volume size is dynamically changed as data grows or shrinks, but you still see a fixed capacity. Click the indicator to switch between physical and virtual capacity, as shown in Figure 5-9.



Allocated: 602.25 GiB / 2.18 TiB (27%)

Virtual: 690.00 GiB / 2.18 TiB (31%)

*Figure 5-9 Storage allocation area*

The following information is displayed in this storage allocation indicator window. To view all of the information, you must use the up and down arrow keys:

- Allocated capacity
- Virtual capacity
- Compression ratio

> **Important:** Starting with IBM Spectrum Virtualize V7.4, the capacity units use the binary prefixes that are defined by the International Electrotechnical Commission (IEC). The prefixes represent a multiplication by 1024 with symbols KiB (kibibyte), MiB (mebibyte), GiB (gibibyte), TiB (tebibyte), and PiB (pebibyte).

## Running tasks indicator

The left corner of area provides information about the running tasks in the system, as shown in Figure 5-10.

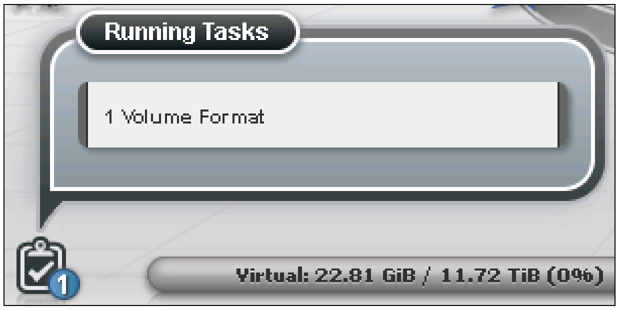
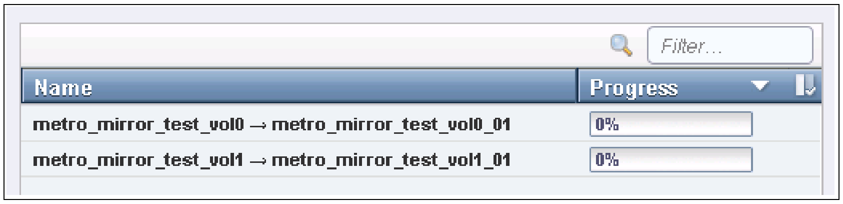


*Figure 5-10 Long-running tasks area*

The following information is displayed in this window:

- Volume migration
- Managed disk (MDisk) removal
- Image mode migration
- Extent migration
- IBM FlashCopy
- Metro Mirror
- Global Mirror
- Volume formatting
- Space-efficient copy repair
- Volume copy verification
- Volume copy synchronization
- Estimated time for the task completion

By clicking within the square (as shown in Figure 5-10), this area provides detailed information about running and recently completed tasks, as shown in Figure 5-11.



| Name | Progress |
|---|---|
| metro_mirror_test_vol0 → metro_mirror_test_vol0_01 | 0% |
| metro_mirror_test_vol1 → metro_mirror_test_vol1_01 | 0% |

*Figure 5-11 Details about running tasks*

### Performance meter

In the middle of the notification area there is a Performance meter consisting of three measured read and write parameters:

- Bandwidth
- Input/output operations per second (IOPS)
- Latency

See Figure 5-12 for details.

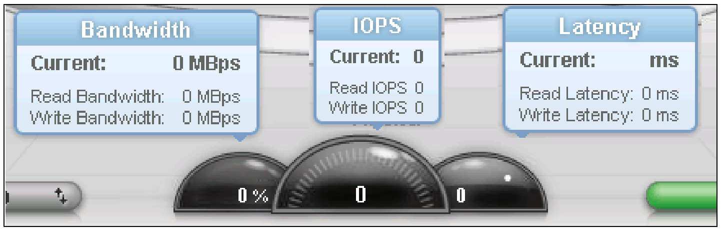


*Figure 5-12 Performance meter*

## 5.3 Overview window

Starting with IBM Spectrum Virtualize V7.4, the welcome window of the GUI changed from the well-known former Overview pane to the new System pane, as shown in Figure 5-13. Clicking **Overview** (Figure 5-13) in the upper-right corner of the System pane opens a modified Overview pane with options that are similar to previous versions of the software.

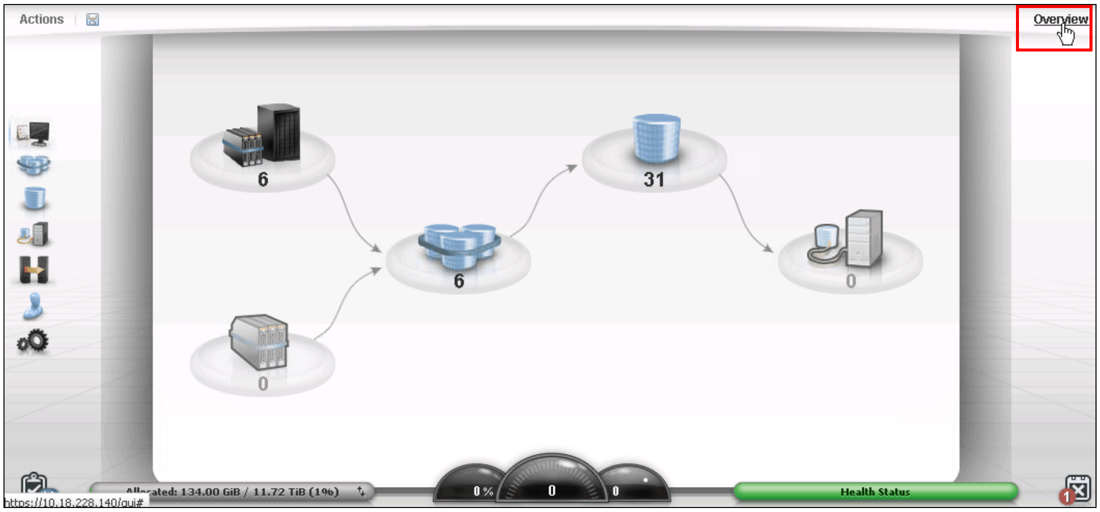

Actions
Overview
6
31
6
0
0
Allocated: 134.00 GiB / 11.72 TiB (1%)
0 %
0
0
Health Status
https://10.18.228.140/gui#
1

*Figure 5-13 Opening the Overview pane*

The following content of the chapter helps you to understand the structure of the pane and how to navigate to various system components to manage them more efficiently and quickly.

### 5.3.1 Content view organization

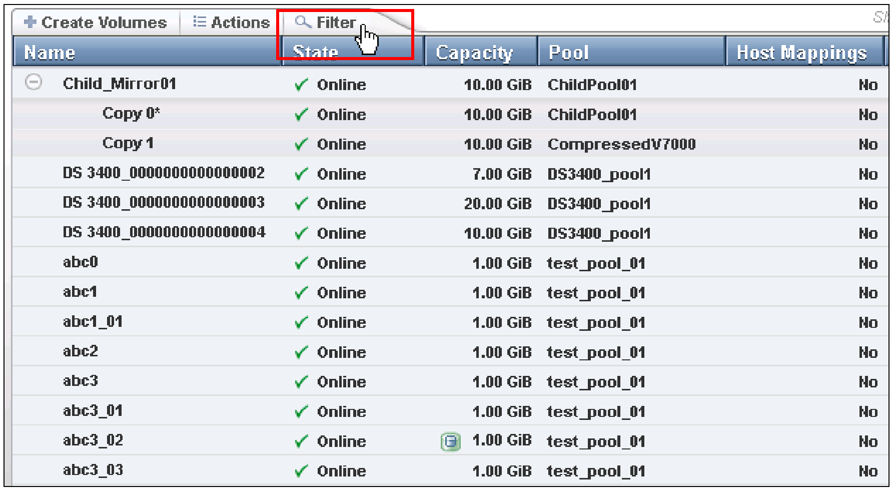
The following sections describe several view options within the GUI in which you can filter (to minimize the amount of data that is shown on the window), sort, and reorganize the content on the window.

## Table filtering

On most pages, a Filter option (magnifying glass icon) is available on the upper-left side of the window. Use this option if the list of object entries is too long.

Complete the following steps to use search filtering:

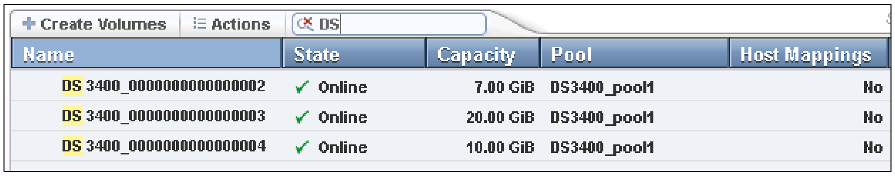
1. Click **Filter** on the upper-left side of the window, as shown in Figure 5-14, to open the search box.



| Name | State | Capacity | Pool | Host Mappings |
|---|---|---|---|---|
| Child_Mirror01 | ✓ Online | 10.00 GiB | ChildPool01 | No |
| Copy 0* | ✓ Online | 10.00 GiB | ChildPool01 | No |
| Copy 1 | ✓ Online | 10.00 GiB | CompressedV7000 | No |
| DS 3400_0000000000000002 | ✓ Online | 7.00 GiB | DS3400_pool1 | No |
| DS 3400_0000000000000003 | ✓ Online | 20.00 GiB | DS3400_pool1 | No |
| DS 3400_0000000000000004 | ✓ Online | 10.00 GiB | DS3400_pool1 | No |
| abc0 | ✓ Online | 1.00 GiB | test_pool_01 | No |
| abc1 | ✓ Online | 1.00 GiB | test_pool_01 | No |
| abc1_01 | ✓ Online | 1.00 GiB | test_pool_01 | No |
| abc2 | ✓ Online | 1.00 GiB | test_pool_01 | No |
| abc3 | ✓ Online | 1.00 GiB | test_pool_01 | No |
| abc3_01 | ✓ Online | 1.00 GiB | test_pool_01 | No |
| abc3_02 | ✓ Online | 1.00 GiB | test_pool_01 | No |
| abc3_03 | ✓ Online | 1.00 GiB | test_pool_01 | No |

*Figure 5-14 Show filter search box*

2. Enter the text string that you want to filter and press Enter.
3. By using this function, you can filter your table that is based on column names. In our example, a volume list is displayed that contains the names that include `DS` somewhere in the name. `DS` is highlighted in amber, as shown in Figure 5-15. The search option is not case-sensitive.



| Name | State | Capacity | Pool | Host Mappings |
|---|---|---|---|---|
| DS 3400_0000000000000002 | ✓ Online | 7.00 GiB | DS3400_pool1 | No |
| DS 3400_0000000000000003 | ✓ Online | 20.00 GiB | DS3400_pool1 | No |
| DS 3400_0000000000000004 | ✓ Online | 10.00 GiB | DS3400_pool1 | No |

*Figure 5-15 Show filtered rows*

4. Remove this filtered view by clicking the reset filter icon, as shown in Figure 5-16.

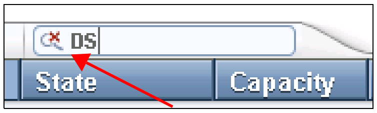


*Figure 5-16 Reset the filtered view*

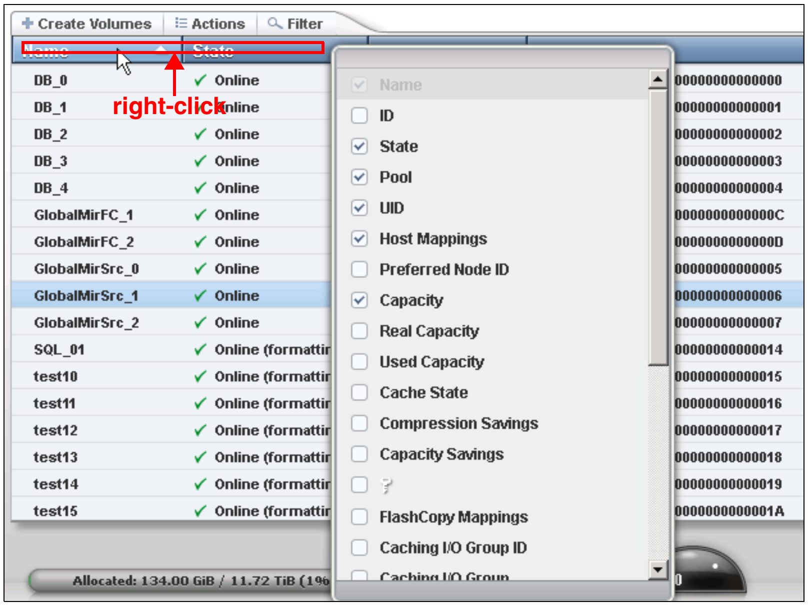
> **Filtering:** This filtering option is available in most menu options of the GUI.

## Table information

In the table view, you can add or remove the information in the tables on most pages.

For example, on the Volumes page, complete the following steps to add a column to our table:

1. Right-click any column headers of the table or select the icon in the left corner of the table header. A list of all of the available columns displays, as shown in Figure 5-17.



Create Volumes | Actions | Filter

right-click

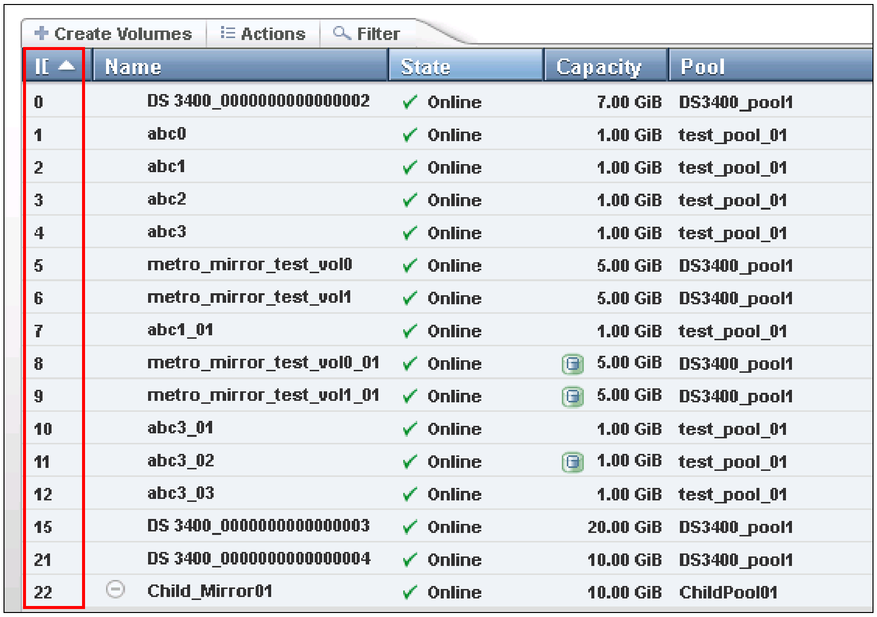
| Name | State | |
|---|---|---|
| DB_0 | ✓ Online | 00000000000000 |
| DB_1 | ✓ Online | 00000000000001 |
| DB_2 | ✓ Online | 00000000000002 |
| DB_3 | ✓ Online | 00000000000003 |
| DB_4 | ✓ Online | 00000000000004 |
| GlobalMirFC_1 | ✓ Online | 0000000000000C |
| GlobalMirFC_2 | ✓ Online | 0000000000000D |
| GlobalMirSrc_0 | ✓ Online | 00000000000005 |
| GlobalMirSrc_1 | ✓ Online | 00000000000006 |
| GlobalMirSrc_2 | ✓ Online | 00000000000007 |
| SQL_01 | ✓ Online (formattir | 00000000000014 |
| test10 | ✓ Online (formattir | 00000000000015 |
| test11 | ✓ Online (formattir | 00000000000016 |
| test12 | ✓ Online (formattir | 00000000000017 |
| test13 | ✓ Online (formattir | 00000000000018 |
| test14 | ✓ Online (formattir | 00000000000019 |
| test15 | ✓ Online (formattir | 0000000000001A |

☑ Name
☐ ID
☑ State
☑ Pool
☑ UID
☑ Host Mappings
☐ Preferred Node ID
☑ Capacity
☐ Real Capacity
☐ Used Capacity
☐ Cache State
☐ Compression Savings
☐ Capacity Savings
☐ ?
☐ FlashCopy Mappings
☐ Caching I/O Group ID
☐ Caching I/O Group

Allocated: 134.00 GiB / 11.72 TiB (1%

Figure 5-17 Add or remove details in a table

2. Select the column that you want to add (or remove) from this table. In our example, we added the volume ID column and sorted the content by ID, as shown on the left in Figure 5-18.



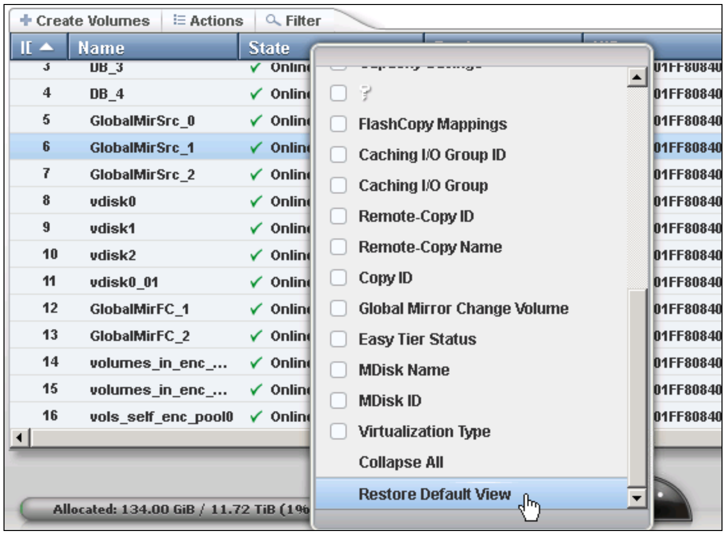
Create Volumes | Actions | Filter

| II | Name | State | Capacity | Pool |
|---|---|---|---|---|
| 0 | DS 3400_0000000000000002 | ✓ Online | 7.00 GiB | DS3400_pool1 |
| 1 | abc0 | ✓ Online | 1.00 GiB | test_pool_01 |
| 2 | abc1 | ✓ Online | 1.00 GiB | test_pool_01 |
| 3 | abc2 | ✓ Online | 1.00 GiB | test_pool_01 |
| 4 | abc3 | ✓ Online | 1.00 GiB | test_pool_01 |
| 5 | metro_mirror_test_vol0 | ✓ Online | 5.00 GiB | DS3400_pool1 |
| 6 | metro_mirror_test_vol1 | ✓ Online | 5.00 GiB | DS3400_pool1 |
| 7 | abc1_01 | ✓ Online | 1.00 GiB | test_pool_01 |
| 8 | metro_mirror_test_vol0_01 | ✓ Online | 5.00 GiB | DS3400_pool1 |
| 9 | metro_mirror_test_vol1_01 | ✓ Online | 5.00 GiB | DS3400_pool1 |
| 10 | abc3_01 | ✓ Online | 1.00 GiB | test_pool_01 |
| 11 | abc3_02 | ✓ Online | 1.00 GiB | test_pool_01 |
| 12 | abc3_03 | ✓ Online | 1.00 GiB | test_pool_01 |
| 15 | DS 3400_0000000000000003 | ✓ Online | 20.00 GiB | DS3400_pool1 |
| 21 | DS 3400_0000000000000004 | ✓ Online | 10.00 GiB | DS3400_pool1 |
| 22 | ⊖ Child_Mirror01 | ✓ Online | 10.00 GiB | ChildPool01 |

Figure 5-18 Table with an added ID column

3. You can repeat this process several times to create custom tables to meet your requirements.

4. You can always return to the default table view by selecting **Restore Default View** in the column selection menu, as shown in Figure 5-19.



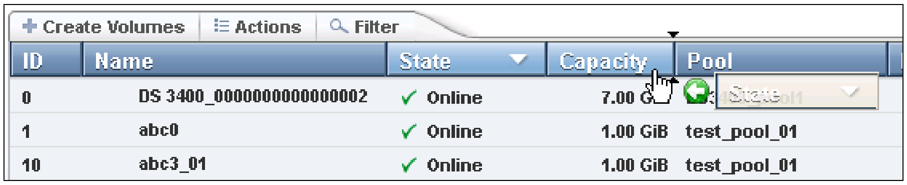
| II | Name | State | |
|---|---|---|---|
| 3 | DB_3 | ✓ Onlin | 01FF80840 |
| 4 | DB_4 | ✓ Onlin | 01FF80840 |
| 5 | GlobalMirSrc_0 | ✓ Onlin | 01FF80840 |
| 6 | GlobalMirSrc_1 | ✓ Onlin | 01FF80840 |
| 7 | GlobalMirSrc_2 | ✓ Onlin | 01FF80840 |
| 8 | vdisk0 | ✓ Onlin | 01FF80840 |
| 9 | vdisk1 | ✓ Onlin | 01FF80840 |
| 10 | vdisk2 | ✓ Onlin | 01FF80840 |
| 11 | vdisk0_01 | ✓ Onlin | 01FF80840 |
| 12 | GlobalMirFC_1 | ✓ Onlin | 01FF80840 |
| 13 | GlobalMirFC_2 | ✓ Onlin | 01FF80840 |
| 14 | volumes_in_enc_... | ✓ Onlin | 01FF80840 |
| 15 | volumes_in_enc_... | ✓ Onlin | 01FF80840 |
| 16 | vols_self_enc_pool0 | ✓ Onlin | 01FF80840 |

*Figure 5-19 Restore default table view*

**Sorting:** By clicking a column, you can sort a table that is based on that column in ascending or descending order.

### Reorganizing columns in tables

You can move columns by left-clicking and moving the column right or left, as shown in Figure 5-20. We are attempting to move the State column after the Capacity column.



| ID | Name | State | Capacity | Pool |
|---|---|---|---|---|
| 0 | DS 3400_0000000000000002 | ✓ Online | 7.00 G | |
| 1 | abc0 | ✓ Online | 1.00 GiB | test_pool_01 |
| 10 | abc3_01 | ✓ Online | 1.00 GiB | test_pool_01 |

*Figure 5-20 Reorganizing the table columns*

## 5.3.2 Help

To access online help, move the mouse pointer over the question mark (**?**) icon in the upper-right corner of any pane and select the context-based help topic, as shown in Figure 5-21 on page 135. Depending on the pane you are working with, the help displays its context item.

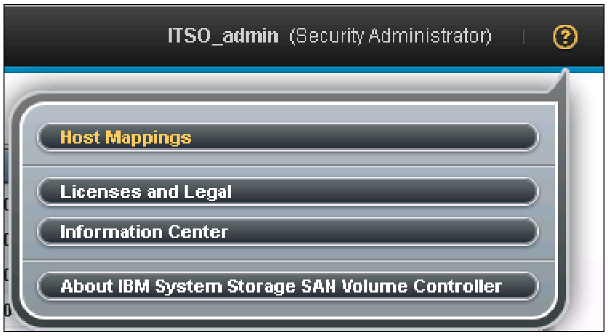

ITSO_admin (Security Administrator)
Host Mappings
Licenses and Legal
Information Center
About IBM System Storage SAN Volume Controller

*Figure 5-21 Help link*

By clicking **Information Center**, you are directed to the public IBM Knowledge Center, which provides all of the information about the IBM Storwize systems.

# 5.4 Viewing Cluster settings using the GUI

This section describes the various configuration and administrative tasks allowed to be performed at the IBM Storwize V7000 clustered level.

## 5.4.1 System status information

From the System pane, complete the following steps to display the system and node information:

1. On the IBM Storwize V7000 System pane, move the mouse pointer over the Monitoring selection and click **System**.

   The System Status pane opens, as shown in Figure 5-22.

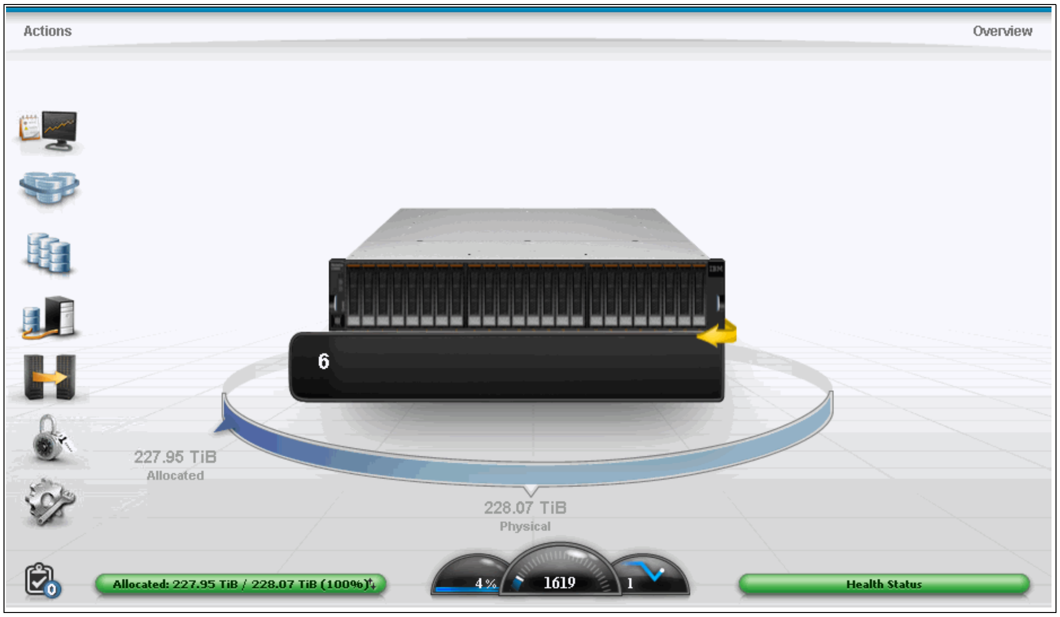


*Figure 5-22 System status pane*

2. On the System status pane, you can view the global storage usage, as shown in Figure 5-22 on page 135. By using this method, you can monitor the physical capacity and the allocated capacity of your system.
3. You can change between the Allocation view and the Thin-provisioning view to see the capacity usage and space savings of the Real-time Compression feature, as shown in Figure 5-23.

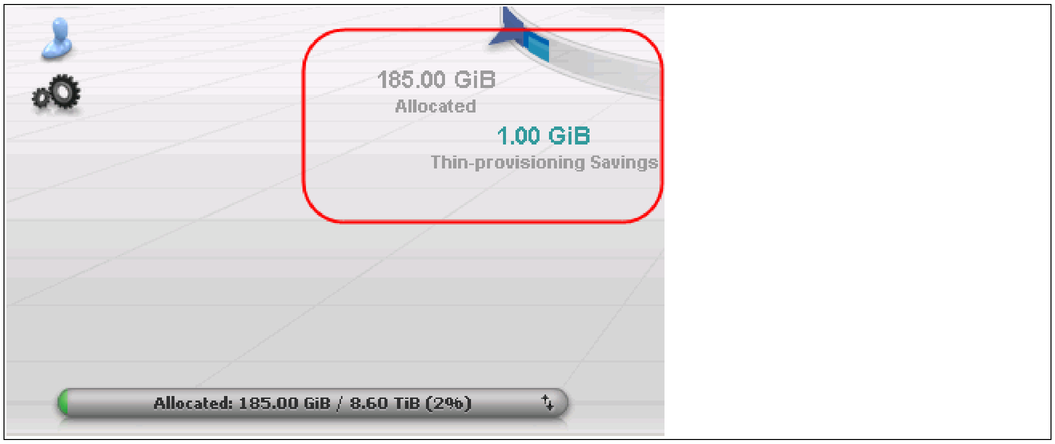

185.00 GiB
Allocated
1.00 GiB
Thin-provisioning Savings
Allocated: 185.00 GiB / 8.60 TiB (2%)

*Figure 5-23 Physical capacity information: Compression view*

## 5.4.2 View IBM Storwize V7000 system properties

Complete the following steps to view the IBM Storwize V7000 system properties:

1. To see more information about the system, select the control enclosure and right-click. The options are shown in Figure 5-24.

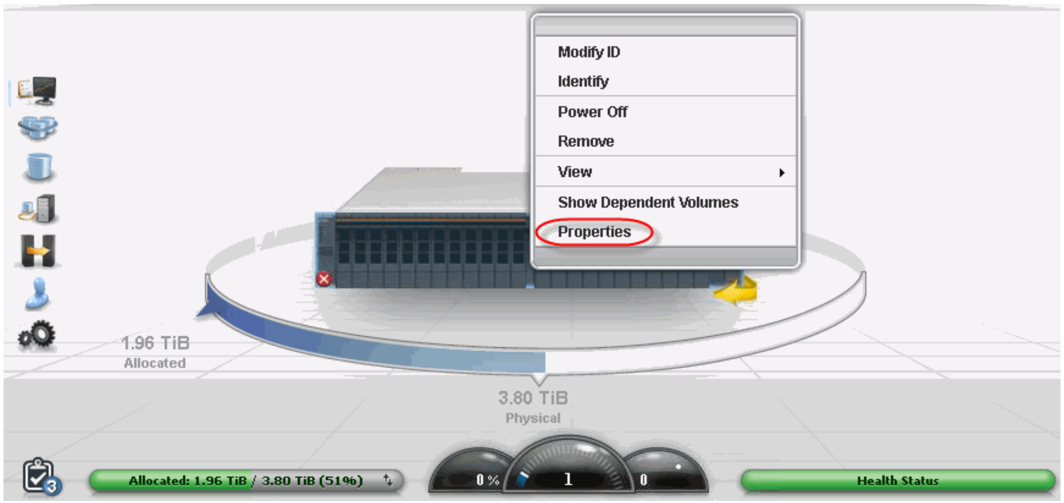
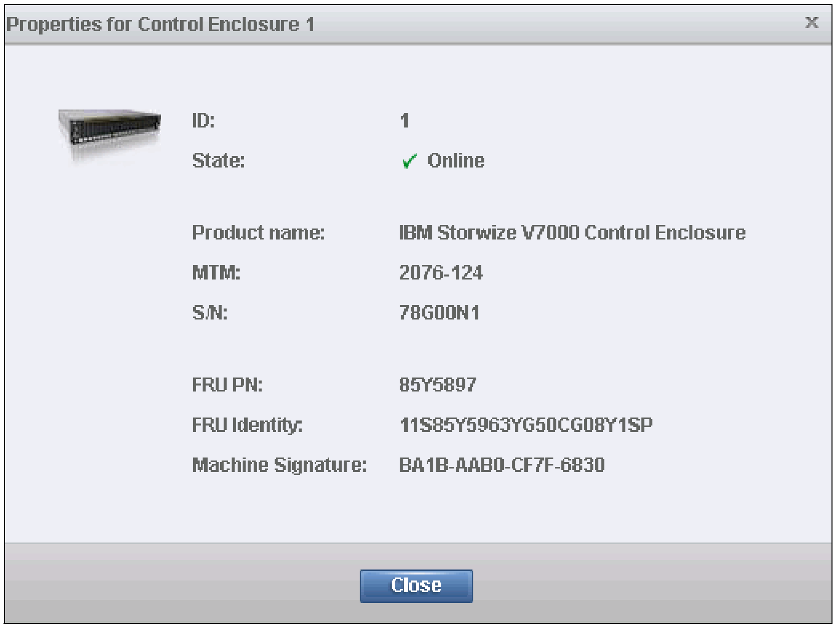


*Figure 5-24 System Properties*

2. You can see the following information:
   - Modify ID
   - Identify
   - Power Off
   - Remove
   - View

  - Show Dependent Volumes
  - Properties

3. Click **Properties** to see the state and the system information such as Serial Number and Machine Signature as shown in Figure 5-25.



Properties for Control Enclosure 1

ID: 1

State: ✓ Online

Product name: IBM Storwize V7000 Control Enclosure

MTM: 2076-124

S/N: 78G00N1

FRU PN: 85Y5897

FRU Identity: 11S85Y5963YG50CG08Y1SP

Machine Signature: BA1B-AAB0-CF7F-6830

*Figure 5-25 Properties of control enclosure*

## 5.4.3 Renaming the IBM Storwize V7000 system

All objects in the system have names that are user-defined or system-generated. Choose a meaningful name when you create an object. If you do not choose a name for the object, the system generates a name for you. A well-chosen name serves not only as a label for an object, but also as a tool for tracking and managing the object. Choosing a meaningful name is important if you decide to use configuration backup and restore.

### Naming rules

When you choose a name for an object, the following rules apply:

- Names must begin with a letter.

> **Important**: Do not start names by using an underscore (_) character even though it is possible. The use of the underscore as the first character of a name is a reserved naming convention that is used by the system configuration restore process.

- The first character cannot be numeric.
- The name can be a maximum of 63 characters with the following exceptions: The name can be a maximum of 15 characters for Remote Copy relationships and groups. The `lsfabric` command displays long object names that are truncated to 15 characters for nodes and systems. V5.1.0 systems display truncated volume names when they are partnered with a version V6.1.0 or later system that has volumes with long object names (`lsrcrelationshipcandidate` or `lsrcrelationship` commands).
- Valid characters are uppercase letters (A - Z), lowercase letters (a - z), digits (0 - 9), the underscore (_) character, a period (.), a hyphen (-), and a space.

- Names must not begin or end with a space.
- Object names must be unique within the object type. For example, you can have a volume named ABC and an MDisk called ABC, but you cannot have two volumes called ABC.
- The default object name is valid (object prefix with an integer).
- Objects can be renamed to their current names.

To rename the system from the System pane, complete the following steps:

1. Click **Actions** in the upper-left corner of the System pane, as shown in Figure 5-26.

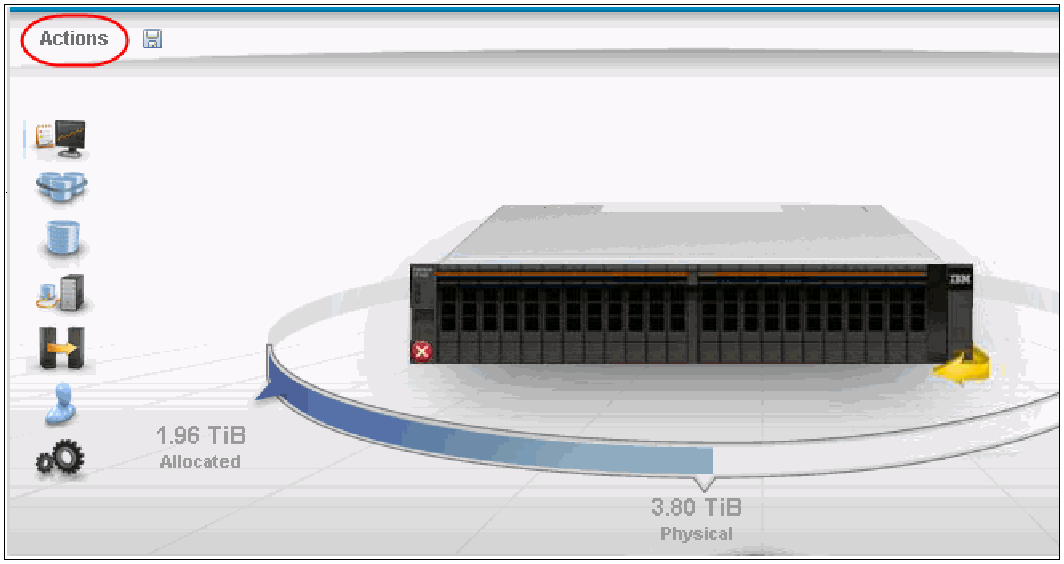


*Figure 5-26 Actions on the System pane*

2. From the pane, select **Rename System**, as shown in Figure 5-27.

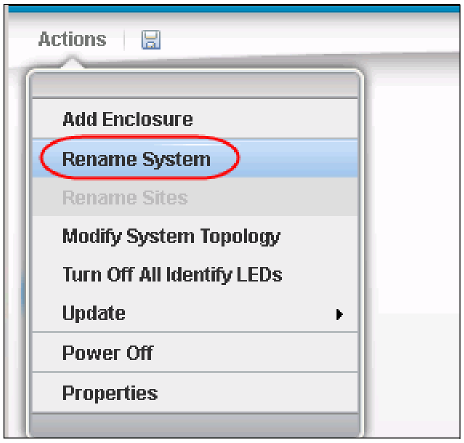
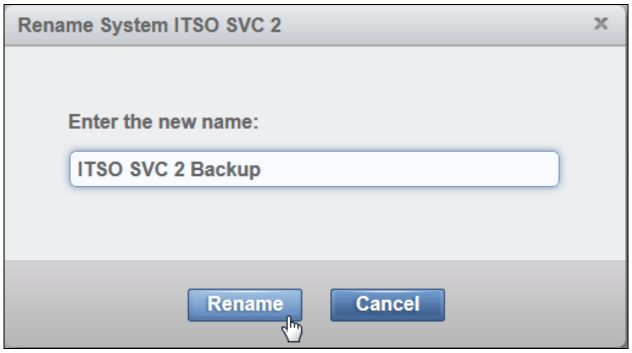


*Figure 5-27 Select Rename System*

3. The pane opens (Figure 5-28). Specify a new name for the system and click **Rename**.



Rename System ITSO SVC 2

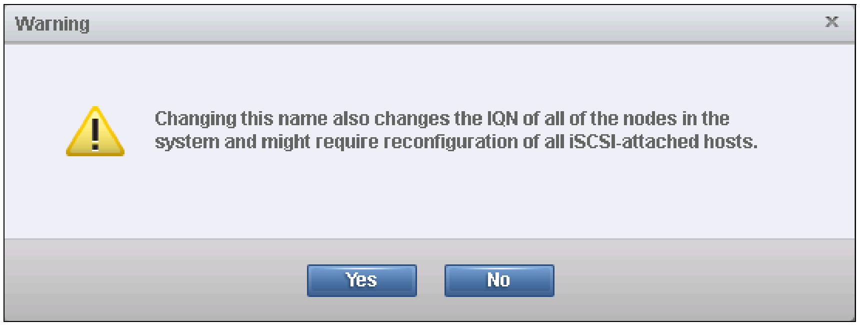
Enter the new name:

ITSO SVC 2 Backup

*Figure 5-28 Rename the system*

**System name:** You can use the letters A - Z and a - z, the numbers 0 - 9, and the underscore (_) character. The clustered system name can be 1 - 63 characters.

4. The Warning window opens, as shown in Figure 5-29. If you are using the iSCSI protocol, changing the system name or the iSCSI Qualified Name (IQN) also changes the IQN of all of the nodes in the system. Changing the system name or the IQN might require the reconfiguration of all iSCSI-attached hosts. This reconfiguration might be required because the IQN for each node is generated by using the system and node names.



Warning

Changing this name also changes the IQN of all of the nodes in the system and might require reconfiguration of all iSCSI-attached hosts.

*Figure 5-29 System rename warning*

5. If you are certain to change the system name, then click **Yes**.

## 5.4.4 Rename a node canister

To rename a node canister, follow these steps:

1. Click in the arrow to rotate the control enclosure as shown in Figure 5-30 on page 140.

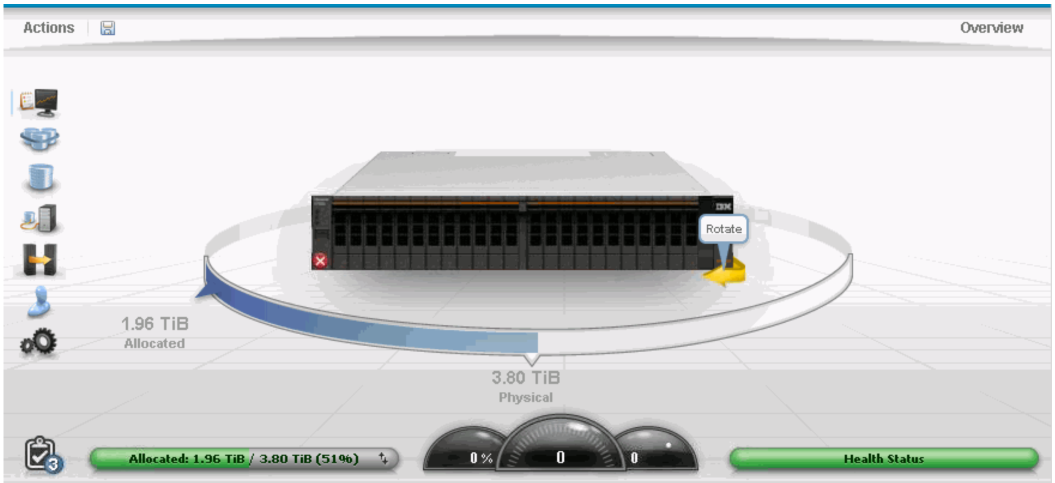


*Figure 5-30 Rotate the control enclosure*

2. Select the node canister you with to rename and click **Rename**.

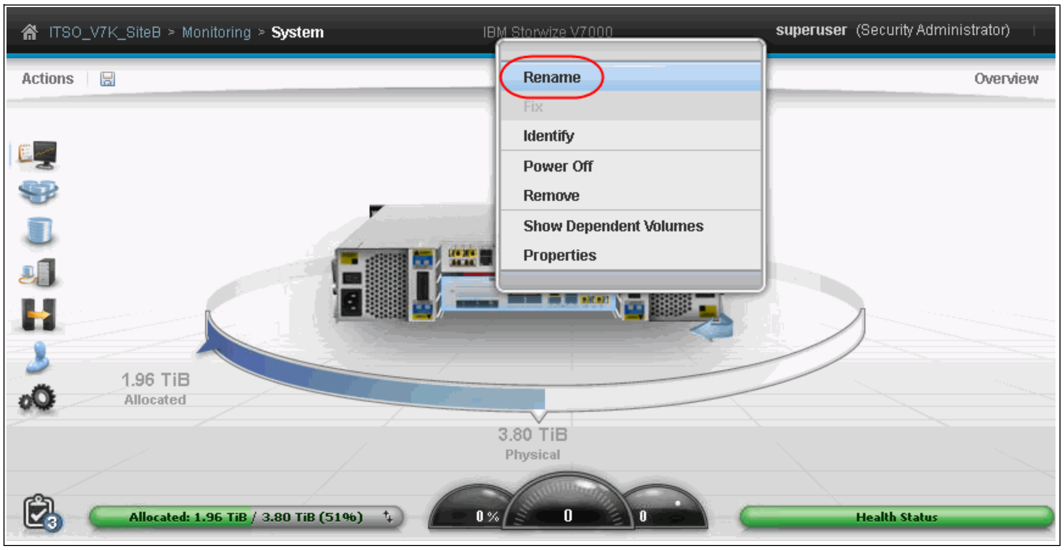
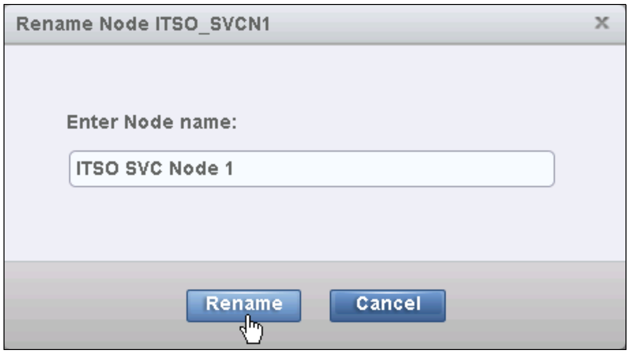


*Figure 5-31 Rename node canister*

3. Enter the new name of the node and click Rename (Figure 5-32).



Enter Node name: ITSO SVC Node 1

*Figure 5-32 Enter the new name of the node*

> **Note:** Changing the node canister name or the iSCSI Qualified Name (IQN) also changes might require the reconfiguration of all iSCSI-attached hosts.

## 5.5 System overview

The System option on the Monitoring menu provides a general overview about the system, including the depiction of all devices in a rack and the allocated or physical storage capacity. When thin-provisioned volumes are enabled, the virtual capacity is also shown by hovering the mouse cursor over the capacity indicator as shown in Figure 5-33.

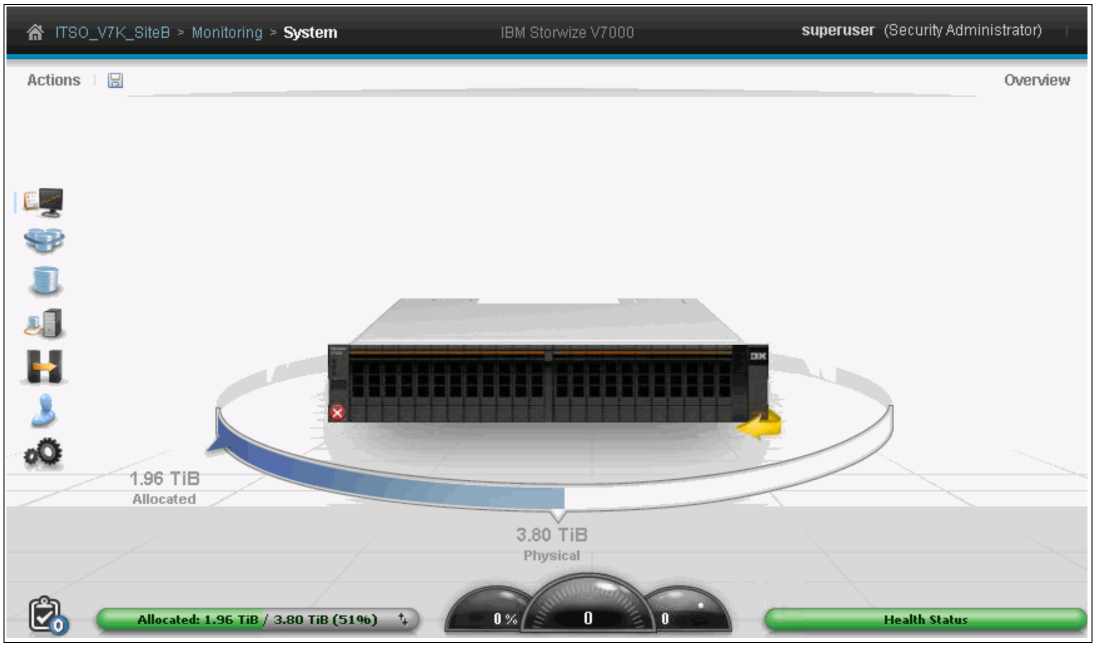
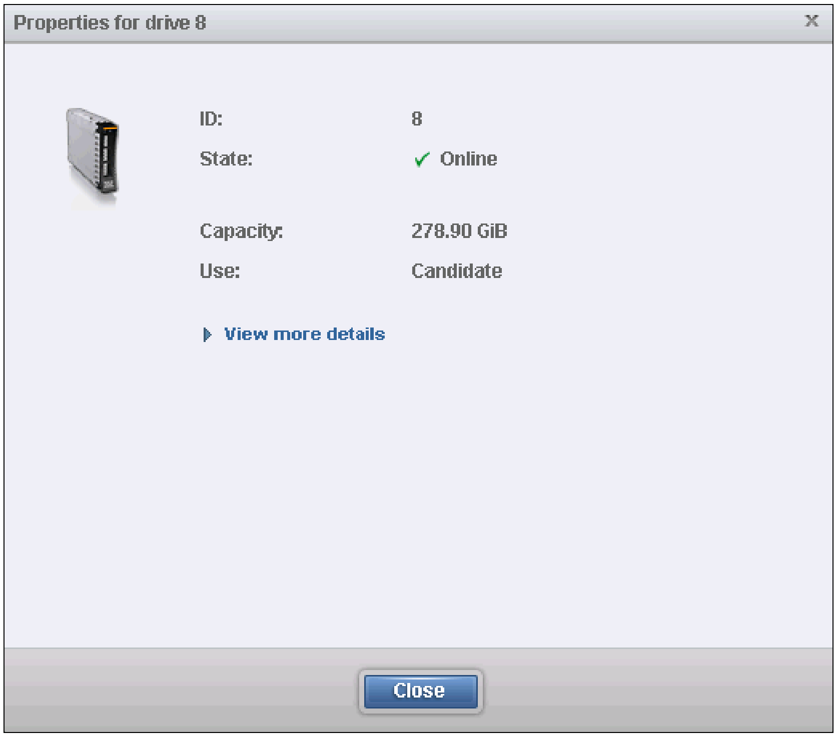

ITSO_V7K_SiteB > Monitoring > System
IBM Storwize V7000
superuser (Security Administrator)
Actions
Overview
1.96 TiB
Allocated
3.80 TiB
Physical
Allocated: 1.96 TiB / 3.80 TiB (51%)
0 %
0
0
Health Status

*Figure 5-33 System overview that shows capacity*

When you click a specific component of the node canister, a pop-up window indicates the details of the component, such as the disk drives in the unit.

By right-clicking and selecting **Properties**, you see detailed technical parameters, such as id, state (online or offline), drive capacity, and the drive use as shown in Figure 5-34 on page 142.


Properties for drive 8

ID: 8

State: ✓ Online

Capacity: 278.90 GiB

Use: Candidate

*Figure 5-34 Drive information*

In an environment with multiple IBM Storwize clusters, you can easily direct the onsite personnel or technician to the correct device by enabling the identification LED on the front pane. Click **Identify** in the pop-up window that is shown in Figure 5-35.

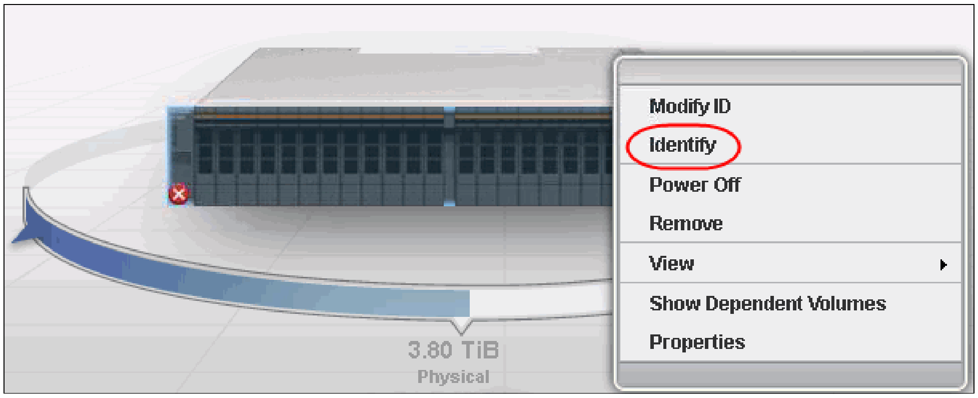
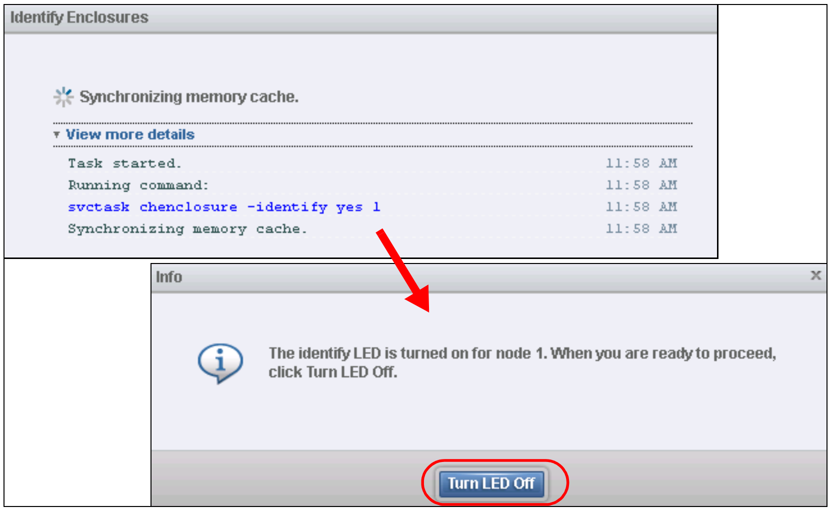


*Figure 5-35 Identification LED*

Then, wait for confirmation from the technician that the device in the data center was correctly identified.

After the confirmation, click **Turn LED Off** (Figure 5-36 on page 143).



```
Task started.                                          11:58 AM
Running command:                                       11:58 AM
svctask chenclosure -identify yes 1                    11:58 AM
Synchronizing memory cache.                            11:58 AM
```

The identify LED is turned on for node 1. When you are ready to proceed, click Turn LED Off.

*Figure 5-36 Using the identification LED*

Alternatively, you can use the command-line interface (CLI) to get the same results. Type the following commands in this sequence:

1. Type **`svctask chenclosure -identify yes 1`** (or just type **`chenclosure -identify yes 1`**).
2. Type **`svctask chenclosure -identify no 1`** (or just type **`chenclosure -identify no 1`**).

Each system that is shown in the Dynamic system view in the middle of a System pane can be rotated by 180° to see its rear side. Click the rotation arrow in the lower-right corner of the device, as illustrated in Figure 5-37.

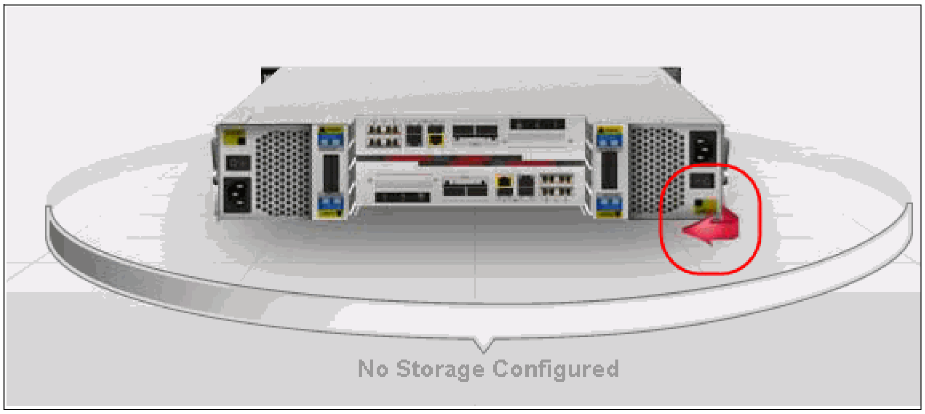


*Figure 5-37 Rotating the enclosure*

## 5.6 Monitoring menu

Hover the cursor over the Monitoring function icon to open the Monitoring menu (Figure 5-38). The Monitoring menu offers these navigation directions:

- System: This option opens the general overview of the IBM Storwize V7000 system, including the depiction of all devices in a rack and the storage capacity. For more information, see 5.5, “System overview” on page 141.
- Events: This option tracks all informational, warning, and error messages that occurred in the system. You can apply various filters to sort the messages according to your needs or

export the messages to an external comma-separated values (CSV) file. For more information, see 5.6.2, "Events" on page 145.

- Performance: This option reports the general system statistics that relate to the processor (CPU) utilization, host and internal interfaces, volumes, and MDisks. The GUI allows you to switch between megabytes per second (MBps) or IOPS. For more information, see 5.6.3, "Performance" on page 146.

With IBM Spectrum Virtualize V7.4 or later, the option that was formerly called System Details is integrated into the device overview on the general System pane, which is available after logging in or when clicking the option System from the Monitoring menu. For more information, see "Overview window" on page 131.

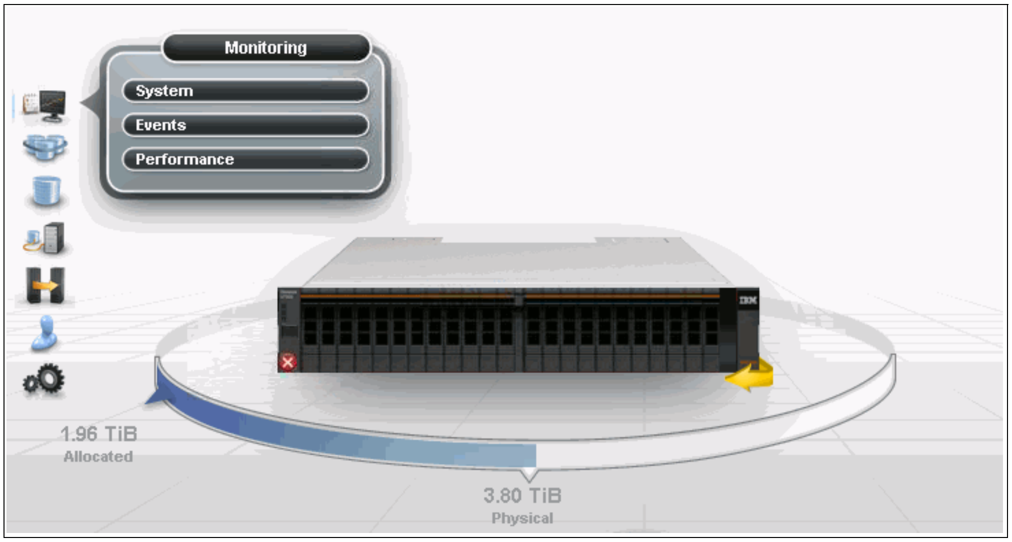


*Figure 5-38 Accessing the Monitoring menu*

In the following section, we describe each option on the Monitoring menu.

## 5.6.1 System details

The System Details option was removed from the Monitoring menu in IBM Spectrum Virtualize V7.3. However, its modified information is still available directly from the System pane. It provides an extended level of information about the parameters and technical details that relate to the system, including the integration of each element into an overall system configuration.

Right-click the control enclosure that you want and click **Properties** (Figure 5-39 on page 145) to obtain detailed information.

By using this menu, you can also power off the machine (without an option for remote start), remove the control enclosure from the system, or, for example, list all of the volumes that are associated with the system.

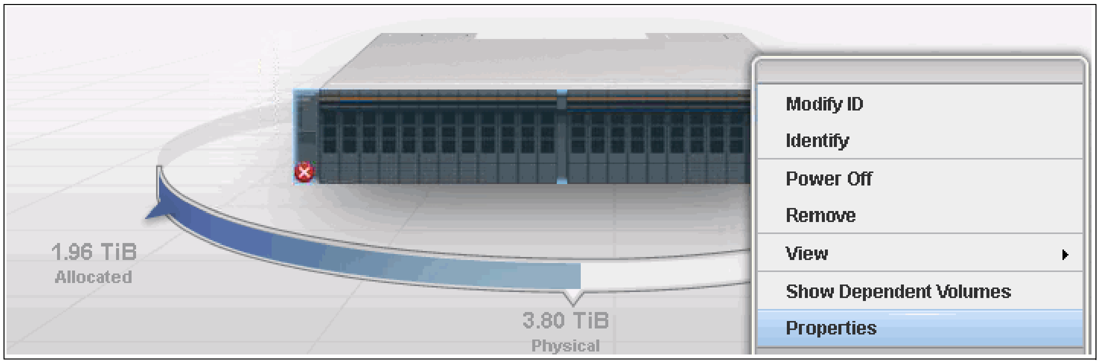
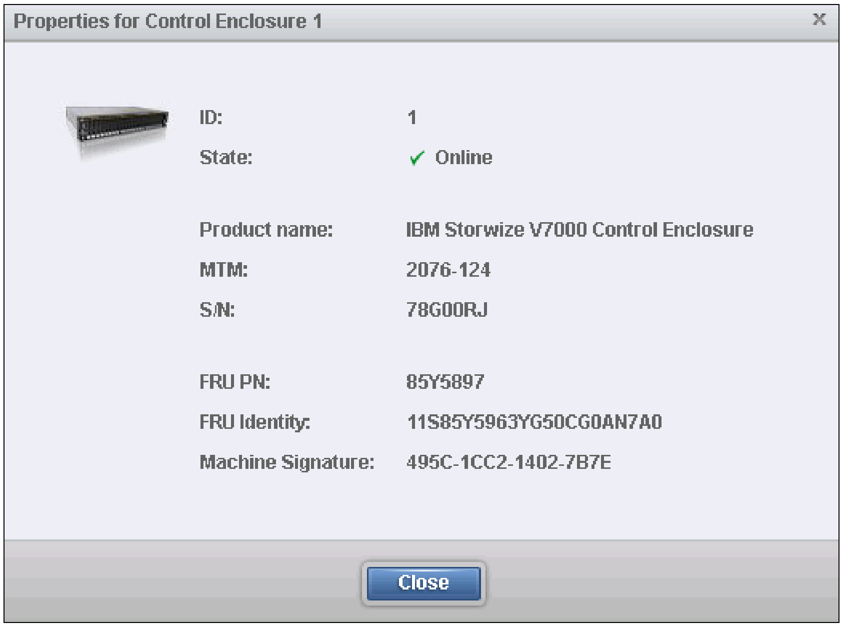


*Figure 5-39 System details*

The output is shown in Figure 5-40.


Properties for Control Enclosure 1

ID: 1

State: ✓ Online

Product name: IBM Storwize V7000 Control Enclosure

MTM: 2076-124

S/N: 78G00RJ

FRU PN: 85Y5897

FRU Identity: 11S85Y5963YG50CG0AN7A0

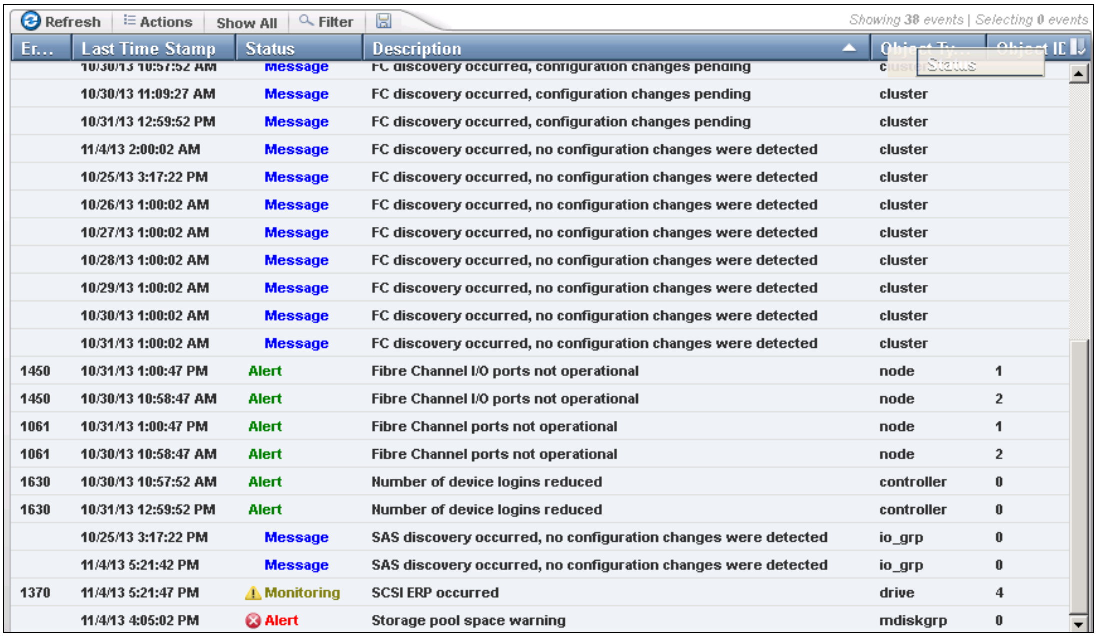
Machine Signature: 495C-1CC2-1402-7B7E

*Figure 5-40 Enclosure technical details*

The Properties option also provides information about the serial number and other machine information.

## 5.6.2 Events

The Events option, which you select from the Monitoring menu, tracks all informational, warning, and error messages that occur in the system. You can apply various filters to sort them, or export them to an external CSV file. A CSV file can be created from the information that is shown here. Figure 5-41 provides an example of records in the system Event log.



| Er... | Last Time Stamp | Status | Description | Object T... | Object ID |
|---|---|---|---|---|---|
| | 10/30/13 10:57:52 AM | Message | FC discovery occurred, configuration changes pending | cluster | |
| | 10/30/13 11:09:27 AM | Message | FC discovery occurred, configuration changes pending | cluster | |
| | 10/31/13 12:59:52 PM | Message | FC discovery occurred, configuration changes pending | cluster | |
| | 11/4/13 2:00:02 AM | Message | FC discovery occurred, no configuration changes were detected | cluster | |
| | 10/25/13 3:17:22 PM | Message | FC discovery occurred, no configuration changes were detected | cluster | |
| | 10/26/13 1:00:02 AM | Message | FC discovery occurred, no configuration changes were detected | cluster | |
| | 10/27/13 1:00:02 AM | Message | FC discovery occurred, no configuration changes were detected | cluster | |
| | 10/28/13 1:00:02 AM | Message | FC discovery occurred, no configuration changes were detected | cluster | |
| | 10/29/13 1:00:02 AM | Message | FC discovery occurred, no configuration changes were detected | cluster | |
| | 10/30/13 1:00:02 AM | Message | FC discovery occurred, no configuration changes were detected | cluster | |
| | 10/31/13 1:00:02 AM | Message | FC discovery occurred, no configuration changes were detected | cluster | |
| 1450 | 10/31/13 1:00:47 PM | Alert | Fibre Channel I/O ports not operational | node | 1 |
| 1450 | 10/30/13 10:58:47 AM | Alert | Fibre Channel I/O ports not operational | node | 2 |
| 1061 | 10/31/13 1:00:47 PM | Alert | Fibre Channel ports not operational | node | 1 |
| 1061 | 10/30/13 10:58:47 AM | Alert | Fibre Channel ports not operational | node | 2 |
| 1630 | 10/30/13 10:57:52 AM | Alert | Number of device logins reduced | controller | 0 |
| 1630 | 10/31/13 12:59:52 PM | Alert | Number of device logins reduced | controller | 0 |
| | 10/25/13 3:17:22 PM | Message | SAS discovery occurred, no configuration changes were detected | io_grp | 0 |
| | 11/4/13 5:21:42 PM | Message | SAS discovery occurred, no configuration changes were detected | io_grp | 0 |
| 1370 | 11/4/13 5:21:47 PM | ⚠ Monitoring | SCSI ERP occurred | drive | 4 |
| | 11/4/13 4:05:02 PM | ⊗ Alert | Storage pool space warning | mdiskgrp | 0 |

*Figure 5-41 Event log list*

## 5.6.3 Performance

The Performance pane reports the general system statistics that relate to processor (CPU) utilization, host and internal interfaces, volumes, and MDisks. You can switch between MBps or IOPS, or even drill down in the statistics to the node level. This capability might be useful when you compare the performance of each control canister in the system if problems exist after a node failover occurs. See Figure 5-42.

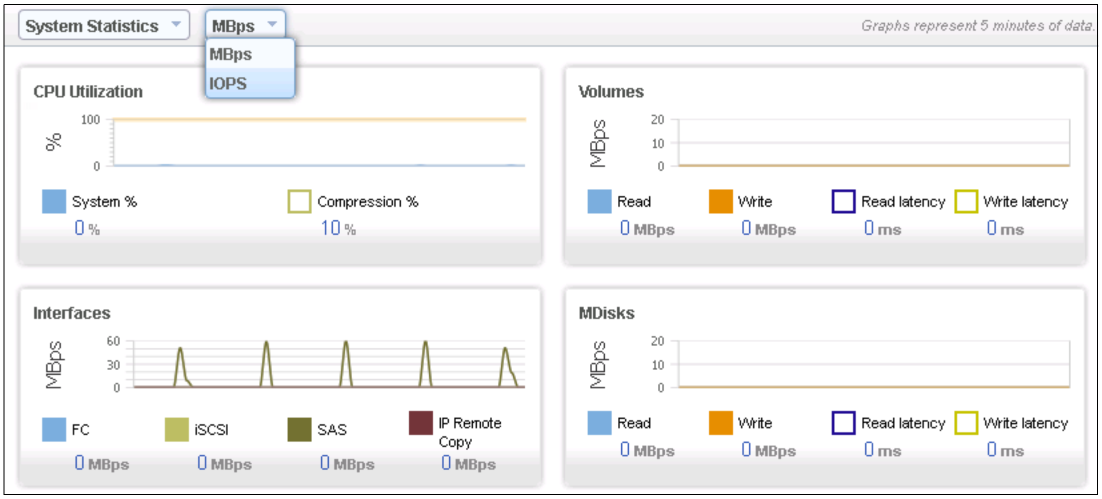


*Figure 5-42 Performance statistics of the IBM Storwize V7000*

The performance statistics in the GUI show, by default, the latest five minutes of data. To see details of each sample, click the graph and select the time stamp, as shown in Figure 5-43.

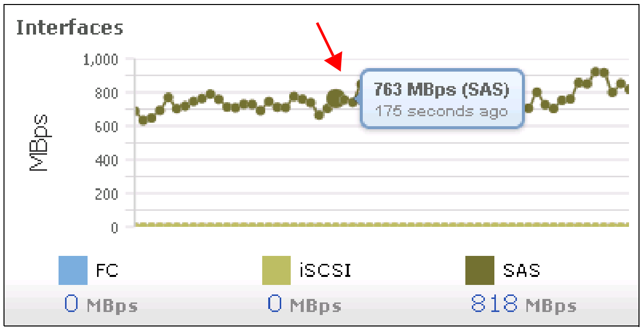


*Figure 5-43 Sample details*

The charts that are shown in Figure 5-43 represent five minutes of the data stream. For in-depth storage monitoring and performance statistics with historical data about your IBM Storwize system, use IBM Spectrum Control (enabled by former IBM Tivoli Storage Productivity Center for Disk and IBM Virtual Storage Center).

# 5.7 Network menu

This section describes how to view the network properties of the IBM Storwize V7000 system. The network information can be obtained by choosing **Network** as shown in Figure 5-44:

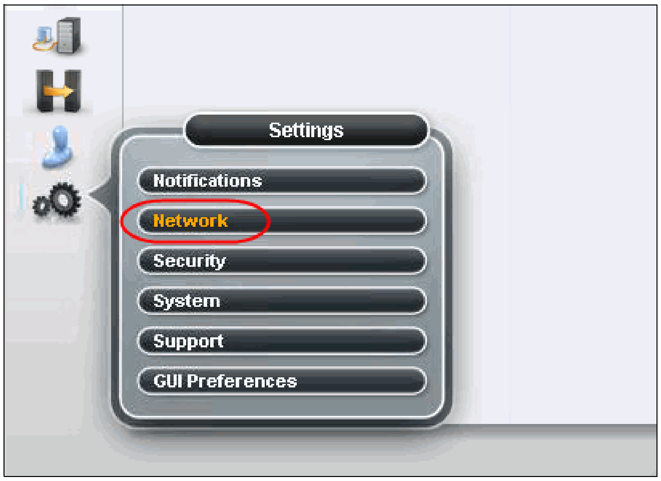


*Figure 5-44 Accessing network information*

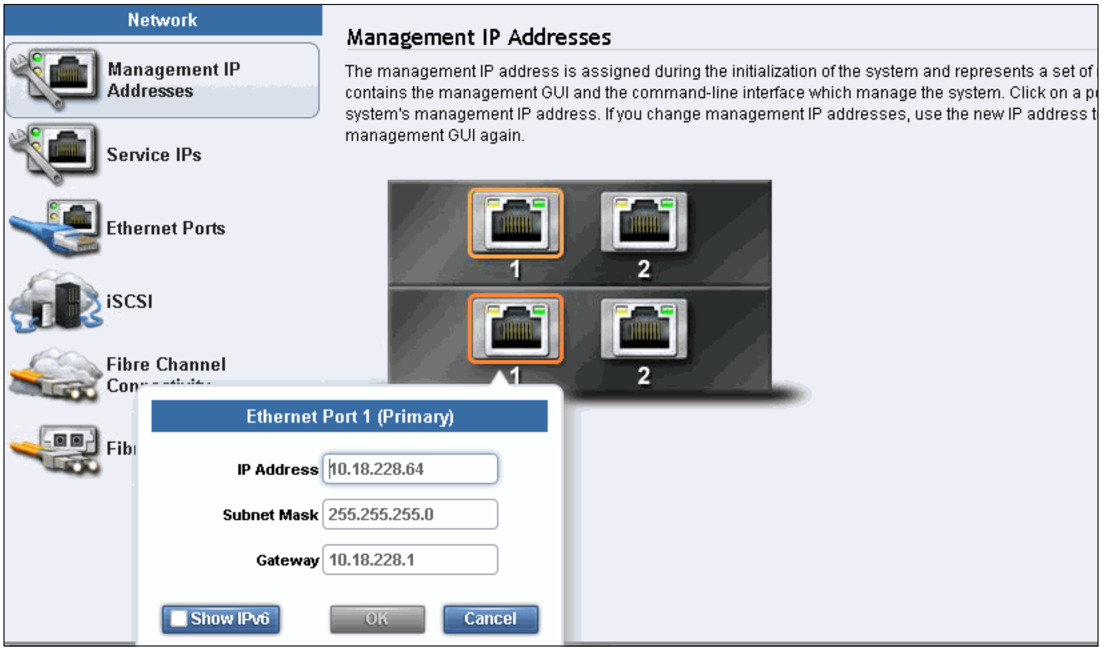
## 5.7.1 Configuring the network

The procedure to set up and configure IBM Storwize V7000 network interfaces is described in Chapter 4, “Initial configuration” on page 85.

## 5.7.2 Management IP addresses

To view the management IP addresses of the IBM Spectrum Virtualize, move your mouse cursor over **Settings** → **Network** as shown in Figure 5-45, and click Management IP Addresses pane.

The GUI shows the management IP address by moving the mouse cursor over the network ports as shown Figure 5-45.



Management IP Addresses

The management IP address is assigned during the initialization of the system and represents a set of contains the management GUI and the command-line interface which manage the system. Click on a p system's management IP address. If you change management IP addresses, use the new IP address t management GUI again.

IP Address 10.18.228.64

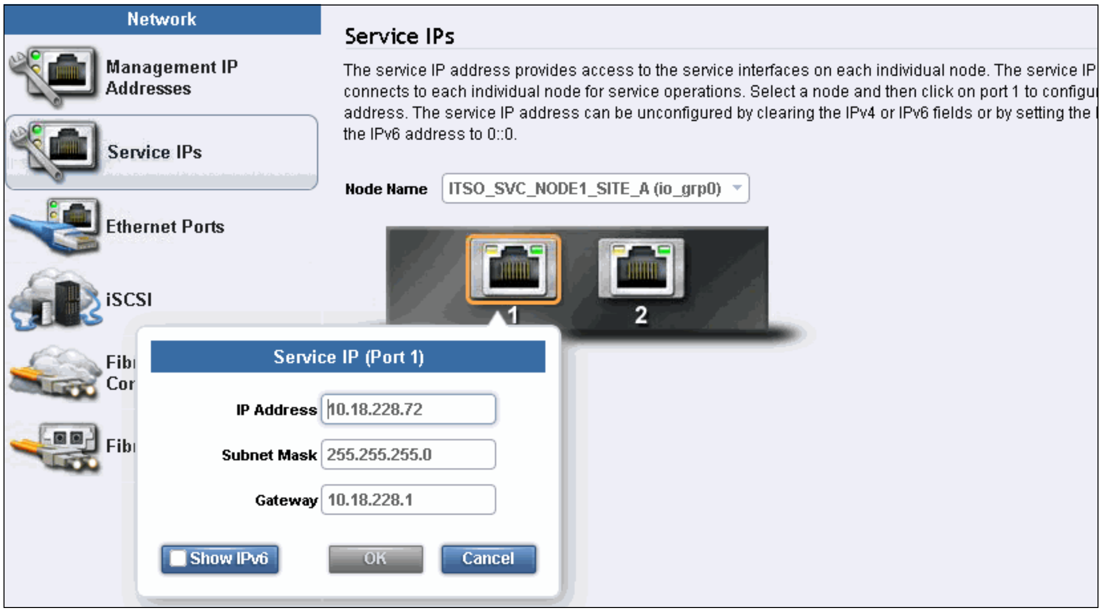
Subnet Mask 255.255.255.0

Gateway 10.18.228.1

*Figure 5-45 Viewing the management IP addresses*

## 5.7.3 Service IP Information

To view the Service IP information of your IBM Spectrum Virtualize, move your mouse cursor over **Settings** → **Network** as shown in Figure 5-44 on page 147, and click over the Service IP Address in the left panel option to view the properties as shown in Figure 5-46.



Service IPs

The service IP address provides access to the service interfaces on each individual node. The service IP connects to each individual node for service operations. Select a node and then click on port 1 to configu address. The service IP address can be unconfigured by clearing the IPv4 or IPv6 fields or by setting the the IPv6 address to 0::0.

Node Name ITSO_SVC_NODE1_SITE_A (io_grp0)

IP Address 10.18.228.72

Subnet Mask 255.255.255.0

Gateway 10.18.228.1

*Figure 5-46 Viewing service IP address*

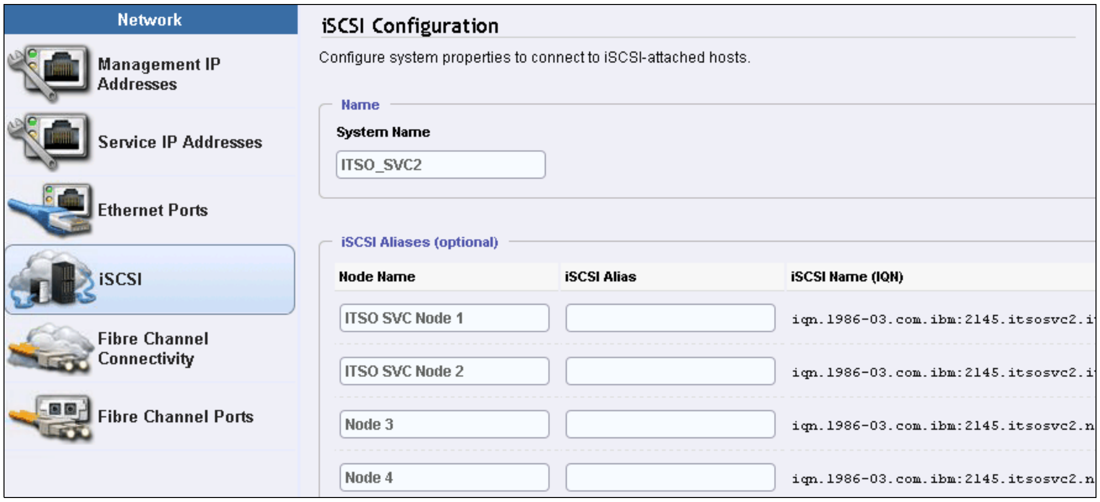
The service IP address is commonly used to provide access to the network interfaces on each individual node of the control enclosure.

Instead of reaching the Management IP address, the service IP address directly connects to each individual node canister for service operations for example.

You may select a node canister of the control enclosure from the drop-down list and then click on any of the ports shown in the GUI. The service IP address can be configure to support IPv4 or IPv6.

## 5.7.4 iSCSI information

From the iSCSI pane in the Settings menu, you can display and configure parameters for the system to connect to iSCSI-attached hosts, as shown in Figure 5-47.



iSCSI Configuration

Configure system properties to connect to iSCSI-attached hosts.

Name

System Name: ITSO_SVC2

iSCSI Aliases (optional)

| Node Name | iSCSI Alias | iSCSI Name (IQN) |
|---|---|---|
| ITSO SVC Node 1 | | iqn.1986-03.com.ibm:2145.itsosvc2.i |
| ITSO SVC Node 2 | | iqn.1986-03.com.ibm:2145.itsosvc2.i |
| Node 3 | | iqn.1986-03.com.ibm:2145.itsosvc2.n |
| Node 4 | | iqn.1986-03.com.ibm:2145.itsosvc2.n |

*Figure 5-47 iSCSI Configuration*

The following parameters can be updated:

- System Name

  It is important to set the system name correctly because it is part of the iSCSI qualified name (IQN) for the node.

  > **Important:** If you change the name of the system after iSCSI is configured, you might need to reconfigure the iSCSI hosts.

  To change the system name, click the system name and specify the new name.

  > **System name**: You can use the letters A - Z and a - z, the numbers 0 - 9, and the underscore (_) character. The name can be 1 - 63 characters.

- iSCSI Aliases (Optional)

  An *iSCSI alias* is a user-defined name that identifies the node to the host. Complete the following steps to change an iSCSI alias:

  a. Click an iSCSI alias.
  b. Specify a name for it.

  Each node has a unique iSCSI name that is associated with two IP addresses. After the host starts the iSCSI connection to a target node, this IQN from the target node is visible in the iSCSI configuration tool on the host.

- iSNS and CHAP

  You can specify the IP address for the iSCSI Storage Name Service (iSNS). Host systems use the iSNS server to manage iSCSI targets and for iSCSI discovery.

  You can also enable Challenge Handshake Authentication Protocol (CHAP) to authenticate the system and iSCSI-attached hosts with the specified shared secret.

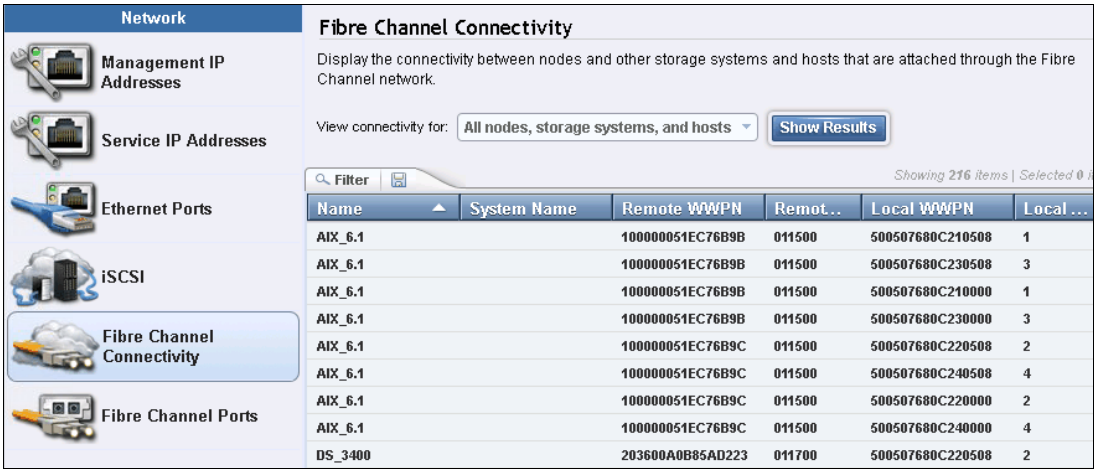
  The CHAP secret is the authentication method that is used to restrict access for other iSCSI hosts to use the same connection. You can set the CHAP for the whole system under the system properties or for each host definition. The CHAP must be identical on the server and the system/host definition. You can create an iSCSI host definition without the use of a CHAP.

## 5.7.5 Fibre Channel information

As shown in Figure 5-48 on page 150, you can use the Fibre Channel Connectivity pane to display the FC connectivity between nodes and other storage systems and hosts that attach through the FC network. You can filter by selecting one of the following fields:

- All nodes, storage systems, and hosts
- Systems
- Nodes
- Storage systems
- Hosts

View the Fibre Channel Connectivity, as shown in Figure 5-48.



Fibre Channel Connectivity

Display the connectivity between nodes and other storage systems and hosts that are attached through the Fibre Channel network.

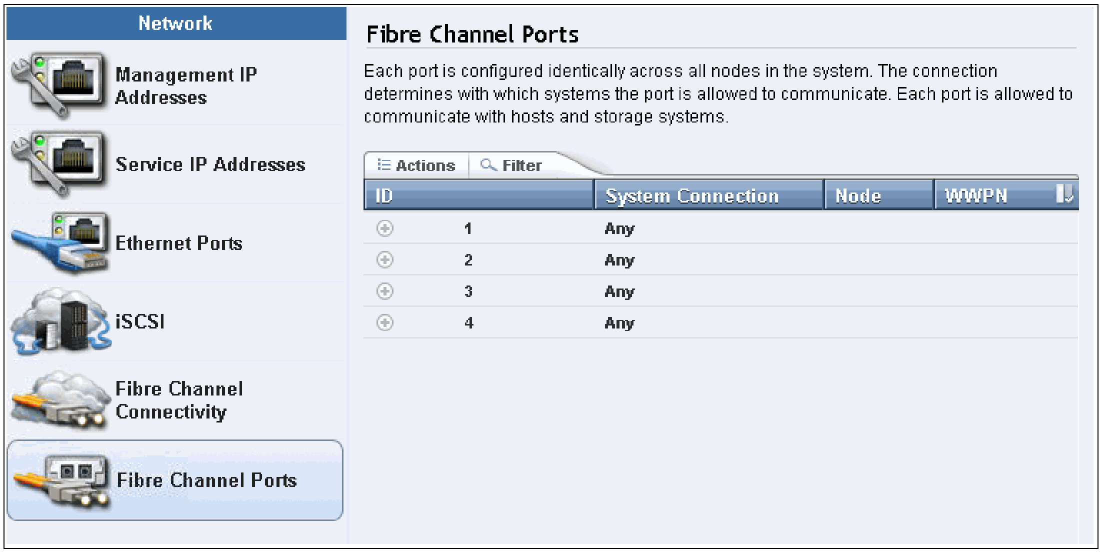
View connectivity for: All nodes, storage systems, and hosts

Showing 216 items | Selected 0 i

| Name | System Name | Remote WWPN | Remot... | Local WWPN | Local ... |
|---|---|---|---|---|---|
| AIX_6.1 | | 100000051EC76B9B | 011500 | 500507680C210508 | 1 |
| AIX_6.1 | | 100000051EC76B9B | 011500 | 500507680C230508 | 3 |
| AIX_6.1 | | 100000051EC76B9B | 011500 | 500507680C210000 | 1 |
| AIX_6.1 | | 100000051EC76B9B | 011500 | 500507680C230000 | 3 |
| AIX_6.1 | | 100000051EC76B9C | 011500 | 500507680C220508 | 2 |
| AIX_6.1 | | 100000051EC76B9C | 011500 | 500507680C240508 | 4 |
| AIX_6.1 | | 100000051EC76B9C | 011500 | 500507680C220000 | 2 |
| AIX_6.1 | | 100000051EC76B9C | 011500 | 500507680C240000 | 4 |
| DS_3400 | | 203600A0B85AD223 | 011700 | 500507680C220508 | 2 |

*Figure 5-48 Fibre Channel*

In the Fibre Channel Ports pane, you can use this view to display how the Fibre Channel port is configured across all control node canisters in the system. This view helps, for example, to determine with which other clusters the port is allowed to communicate (Figure 5-49).



Fibre Channel Ports

Each port is configured identically across all nodes in the system. The connection determines with which systems the port is allowed to communicate. Each port is allowed to communicate with hosts and storage systems.

| ID | System Connection | Node | WWPN |
|---|---|---|---|
| 1 | Any | | |
| 2 | Any | | |
| 3 | Any | | |
| 4 | Any | | |

*Figure 5-49 Viewing Fibre Channel Port properties*

## 5.8 Notifications menu

IBM Storwize V7000 can use Simple Network Management Protocol (SNMP) traps, syslog messages, and Call Home email to notify you and the IBM Support Center when significant events are detected. Any combination of these notification methods can be used simultaneously.

Notifications are normally sent immediately after an event is raised. However, events can occur because of service actions that are performed. If a recommended service action is active, notifications about these events are sent only if the events are still unfixed when the service action completes.

### 5.8.1 Email notifications

The Call Home feature transmits operational and event-related data to you and IBM through a Simple Mail Transfer Protocol (SMTP) server connection in the form of an event notification email. When configured, this function alerts IBM service personnel about hardware failures and potentially serious configuration or environmental issues.

Complete the following steps to view email event notifications:

1. From the main System pane, move the mouse pointer over the **Settings** selection in the dynamic menu and click **Notifications**, as shown in Figure 5-50:

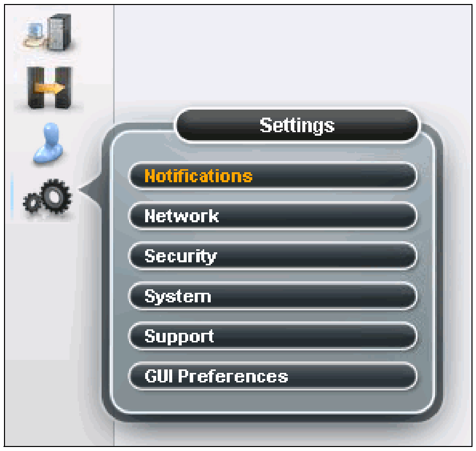
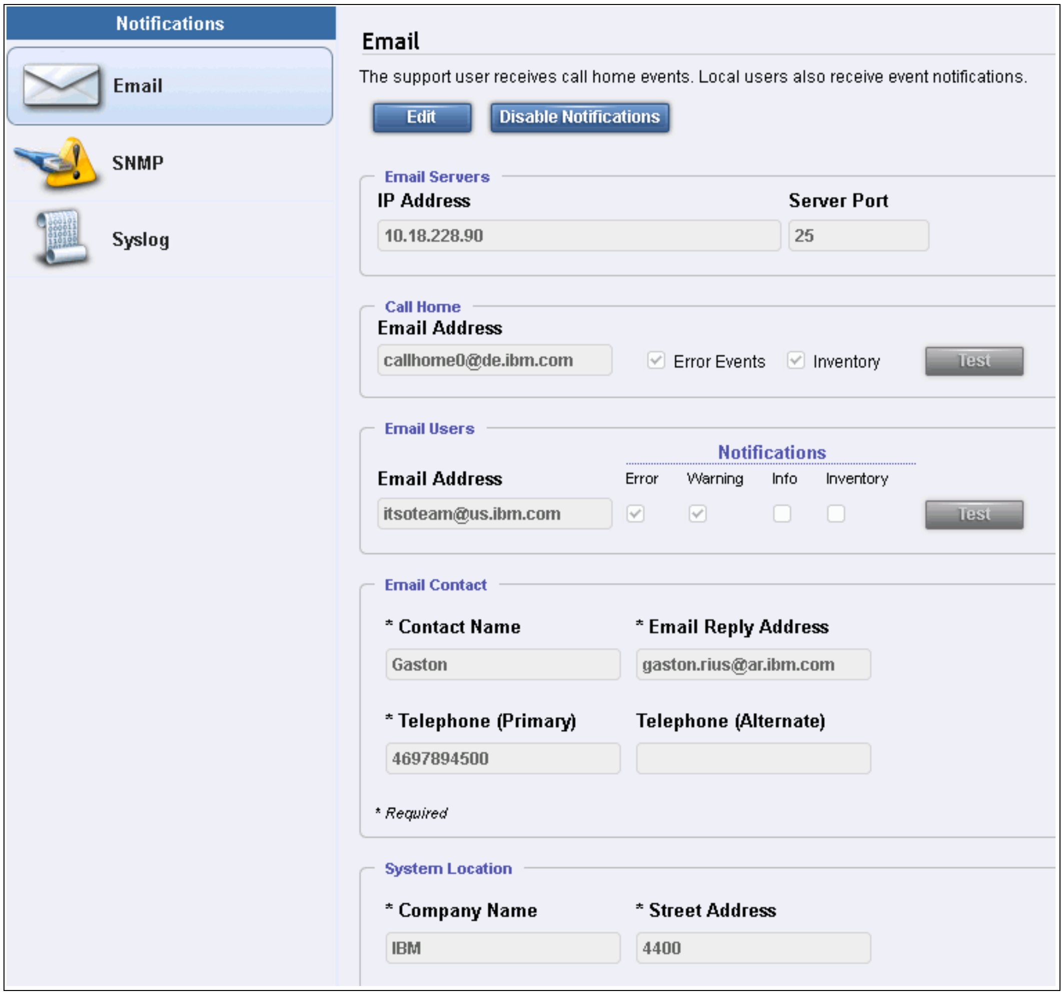


*Figure 5-50*

2. In the left column, select **Email** as shown in Figure 5-51 on page 152



Email

The support user receives call home events. Local users also receive event notifications.

Edit | Disable Notifications

Email Servers

| IP Address | Server Port |
|---|---|
| 10.18.228.90 | 25 |

Call Home

| Email Address | | | |
|---|---|---|---|
| callhome0@de.ibm.com | ☑ Error Events | ☑ Inventory | Test |

Email Users

| Email Address | Error | Warning | Info | Inventory | |
|---|---|---|---|---|---|
| itsoteam@us.ibm.com | ☑ | ☑ | □ | □ | Test |

Email Contact

| * Contact Name | * Email Reply Address |
|---|---|
| Gaston | gaston.rius@ar.ibm.com |

| * Telephone (Primary) | Telephone (Alternate) |
|---|---|
| 4697894500 | |

* *Required*

System Location

| * Company Name | * Street Address |
|---|---|
| IBM | 4400 |

*Figure 5-51 Viewing call home event information*

3. From this view, there is useful information about email notification and call-home information, such as:

- The IP of the email server (SMTP Server) and Port;
- The Call-home email address;
- The e-mail of the user(s) who is set to receive the email notification(s);
- The contact information of the person in the organization responsible for the system;
- System location.

### 5.8.2 SNMP notifications

Simple Network Management Protocol (SNMP) is a standard protocol for managing networks and exchanging messages. The system can send SNMP messages that notify personnel about an event. You can use an SNMP manager to view the SNMP messages that are sent by IBM Storwize V7000.

To view the SNMP configuration, use the **System** window move the mouse pointer over the **Settings** and click **Notification** → **SNMP**.

From this window (Figure 5-52), you can view and configure an SNMP server to receive various informational, error, or warning notifications by setting the following information:

- IP Address

  The address for the SNMP server.

- Server Port

  The remote port number for the SNMP server. The remote port number must be a value of 1 - 65535.

- Community

  The SNMP community is the name of the group to which devices and management stations that run SNMP belong.

- Event Notifications:

  Consider the following points about event notifications:

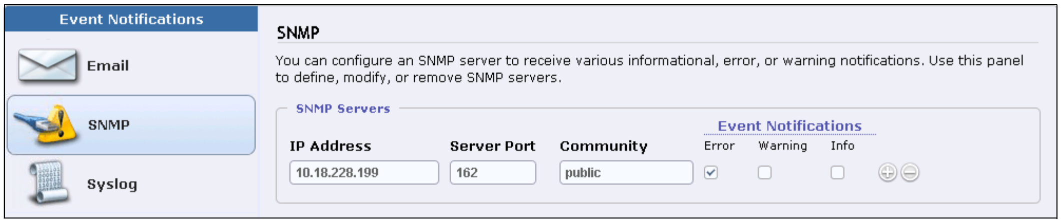
  – Select Error if you want the user to receive messages about problems, such as hardware failures, that must be resolved immediately.

    > **Important:** Browse to Recommended Actions to run the fix procedures on these notifications.

  – Select Warning if you want the user to receive messages about problems and unexpected conditions. Investigate the cause immediately to determine any corrective action.

    > **Important:** Browse to Recommended Actions to run the fix procedures on these notifications.

  – Select Info (Figure 5-52) if you want the user to receive messages about expected events. No action is required for these events.



SNMP

You can configure an SNMP server to receive various informational, error, or warning notifications. Use this panel to define, modify, or remove SNMP servers.

SNMP Servers

|  |  |  | Event Notifications |  |  |
|---|---|---|---|---|---|
| IP Address | Server Port | Community | Error | Warning | Info |
| 10.18.228.199 | 162 | public | ☑ | ☐ | ☐ |

*Figure 5-52 SNMP configuration*

To remove an SNMP server, click the Minus sign (**-**). To add another SNMP server, click the Plus sign (**+**).

## 5.8.3 Syslog notifications

The syslog protocol is a standard protocol for forwarding log messages from a sender to a receiver on an IP network. The IP network can be IPv4 or IPv6. The system can send syslog messages that notify personnel about an event. You can use the **Syslog** pane to view the Syslog messages that are sent by the IBM Storwize V7000. To view the Syslog configuration, use the System window and move the mouse pointer over **Settings** and click **Notification** → **Syslog**.

From this window (Figure 5-53), you can view and configure a syslog server to receive log messages from various systems and store them in a central repository by entering the following information:

- IP Address

  The IP address for the syslog server.

- Facility

  The facility determines the format for the syslog messages. The facility can be used to determine the source of the message.

- Message Format

  The message format depends on the facility. The system can transmit syslog messages in the following formats:

  – The concise message format provides standard detail about the event.
  – The expanded format provides more details about the event.

- Event Notifications:

  Consider the following points about event notifications:

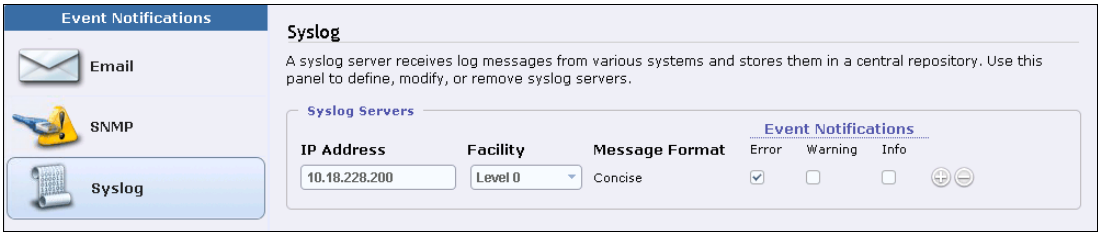
  – Select Error if you want the user to receive messages about problems, such as hardware failures, that must be resolved immediately.

  > **Important:** Browse to Recommended Actions to run the fix procedures on these notifications.

  – Select Warning if you want the user to receive messages about problems and unexpected conditions. Investigate the cause immediately to determine whether any corrective action is necessary.

  > **Important:** Browse to Recommended Actions to run the fix procedures on these notifications.

– Select Info (Figure 5-53) if you want the user to receive messages about expected events. No action is required for these events.



Syslog

A syslog server receives log messages from various systems and stores them in a central repository. Use this panel to define, modify, or remove syslog servers.

Syslog Servers

|  |  |  | Event Notifications |  |  |
|---|---|---|---|---|---|
| IP Address | Facility | Message Format | Error | Warning | Info |
| 10.18.228.200 | Level 0 | Concise | ☑ | □ | □ |

*Figure 5-53 Syslog configuration*

To remove a syslog server, click the Minus sign (**-**). To add another syslog server, click the Plus sign (**+**).

The syslog messages can be sent in a concise message format or an expanded message format.

Example 5-1 shows a compact format syslog message.

*Example 5-1 Compact syslog message example*

```
IBM2145 #NotificationType=Error #ErrorID=077001 #ErrorCode=1070 #Description=Node
CPU fan failed #ClusterName=SVCCluster1 #Timestamp=Wed Jul 02 08:00:00 2014 BST
#ObjectType=Node #ObjectName=Node1 #CopyID=0 #ErrorSequenceNumber=100
```

Example 5-2 shows an expanded format syslog message.

*Example 5-2 Full format syslog message example*

```
IBM2145 #NotificationType=Error #ErrorID=077001 #ErrorCode=1070 #Description=Node
CPU fan failed #ClusterName=SVCCluster1 #Timestamp=Wed Jul 02 08:00:00 2014 BST
#ObjectType=Node #ObjectName=Node1 #CopyID=0 #ErrorSequenceNumber=100 #ObjectID=2
#NodeID=2 #MachineType=21454F2#SerialNumber=1234567 #SoftwareVersion=5.1.0.0
(build 8.14.0805280000)#FRU=fan 24P1118, system board 24P1234
#AdditionalData(0->63)=0000000021000000000000000000000000000000000000000000000
0000000000000000000000000000000000000000000000000000000000000000000#Additional
Data(64-127)=000000000000000000000000000000000000000000000000000000000000000000
0000000000000000000000000000000000000000000000000000000000
```

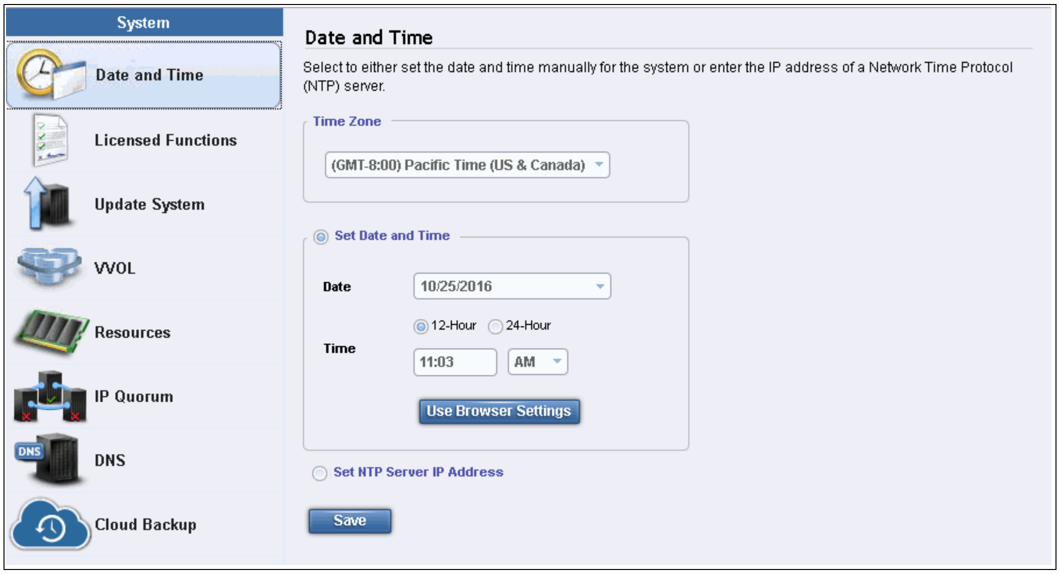
# 5.9 System menu

Use the System panel to view and change the time and date settings, work with licensing options, download configuration settings, work with VMware VVOLs, IP Quorum or download software upgrade packages.

## 5.9.1 Date and time

Complete the following steps to view or configure the date and time settings:

1. From the main System pane, move the pointer over **Settings** and click **System**.
2. In the left column, select **Date and Time**, as shown in Figure 5-54.



Date and Time

Select to either set the date and time manually for the system or enter the IP address of a Network Time Protocol (NTP) server.

Time Zone: (GMT-8:00) Pacific Time (US & Canada)

● Set Date and Time

Date: 10/25/2016

Time: ● 12-Hour ○ 24-Hour

11:03 AM

○ Set NTP Server IP Address

Figure 5-54 Date and Time window

3. From this pane, you can modify the following information:

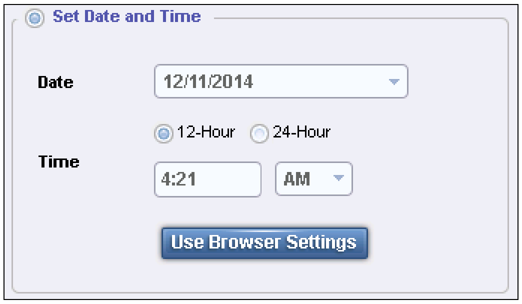
   – Time zone

     Select a time zone for your system by using the drop-down list.

   – Date and time

     The following options are available:

     - If you are not using a Network Time Protocol (NTP) server, select **Set Date and Time**, and then manually enter the date and time for your system, as shown in Figure 5-55. You can also click **Use Browser Settings** to automatically adjust the date and time of your IBM Storwize V7000 system with your local workstation date and time.



● Set Date and Time

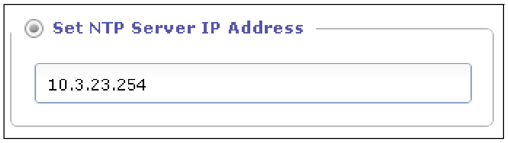
Date: 12/11/2014

Time: ● 12-Hour ○ 24-Hour

4:21 AM

Figure 5-55 Set Date and Time window

     - If you are using a Network Time Protocol (NTP) server, select **Set NTP Server IP Address** and then enter the IP address of the NTP server, as shown in Figure 5-56.



● Set NTP Server IP Address

10.3.23.254

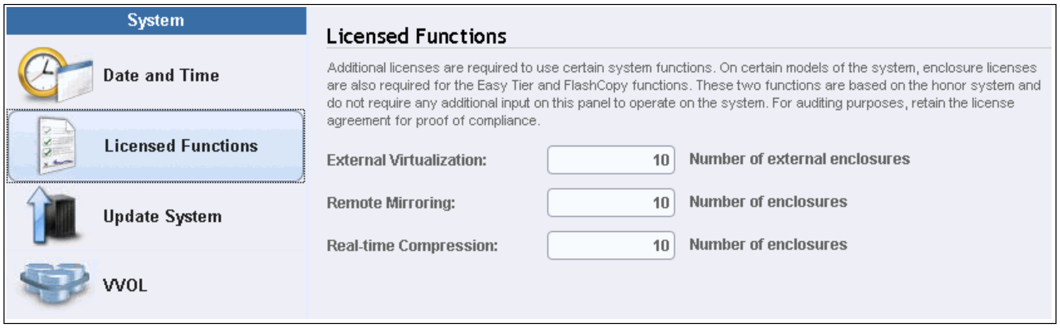
Figure 5-56 Set NTP Server IP Address window

4. Click **Save**.

## 5.9.2 Licensing

Complete the following steps to view or configure the licensing settings:

1. From the main Settings pane, move the pointer over Settings and click **System**.
2. In the left column, select **License Functions**, as shown in Figure 5-57.



Licensed Functions

Additional licenses are required to use certain system functions. On certain models of the system, enclosure licenses are also required for the Easy Tier and FlashCopy functions. These two functions are based on the honor system and do not require any additional input on this panel to operate on the system. For auditing purposes, retain the license agreement for proof of compliance.

External Virtualization: 10 Number of external enclosures

Remote Mirroring: 10 Number of enclosures

Real-time Compression: 10 Number of enclosures

*Figure 5-57 Licensing window*

3. In the Licensed Functions pane, you can view or set the licensing options for the IBM Storwize V7000 for the following elements (limits are in TiB):

   – External Virtualization

     Specifies the of external enclosures attached to your IBM Storwize V7000.

   – Remote Mirroring Limit

     Enter the capacity that is available for Metro Mirror and Global Mirror relationships.

     > **Important:** The Used capacity for Global Mirror and Metro Mirror is the sum of the capacities of all of the volumes that are in a Metro Mirror or Global Mirror relationship; both master volumes and auxiliary volumes are included.

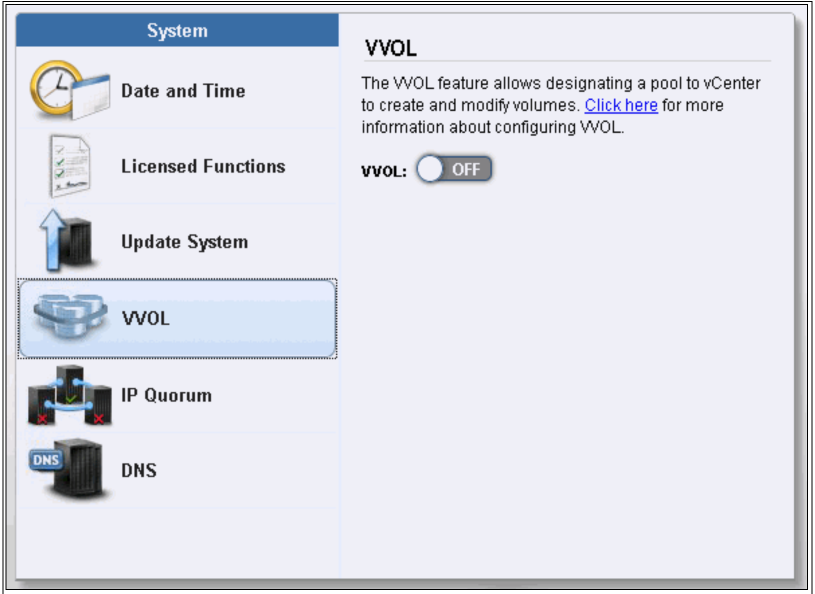
   – Real-time Compression Limit

     Enter the total number of TiB of virtual capacity that are licensed for compression.

## 5.9.3 VMware Virtual Volumes

IBM Spectrum Virtualize V7.6 and later is able to manage VMware vSphere Virtual Volumes (VVols) directly in cooperation with VMware. It enables VMware virtual machines to get the assigned disk capacity directly from IBM Storwize V7000 rather than from the ESXi datastore. That enables storage administrators to control the appropriate usage of storage capacity, and to enable enhanced features of storage virtualization directly to the virtual machine (such as replication, thin-provisioning, compression, encryption, and so on).

VVols management is enabled in IBM Storwize V7000 in the System section, as shown in Figure 5-58. The NTP server must be configured before enabling VVols management. It is strongly advised to use the same NTP server for ESXi and for IBM Storwize V7000.



The VVOL feature allows designating a pool to vCenter to create and modify volumes. Click here for more information about configuring VVOL.

VVOL: OFF

*Figure 5-58 Enabling VVOLs management*

A quick-start guide to VVols, *Quick-start Guide to Configuring VMware Virtual Volumes for Systems Powered by IBM Spectrum Virtualize*, REDP-5321 is available at:

`https://www.redbooks.ibm.com/Redbooks.nsf/RedpieceAbstracts/redp5321.html`

An IBM Redbook is also available, *Configuring VMware Virtual Volumes for Systems Powered by IBM Spectrum Virtualize*, SG24-8328:

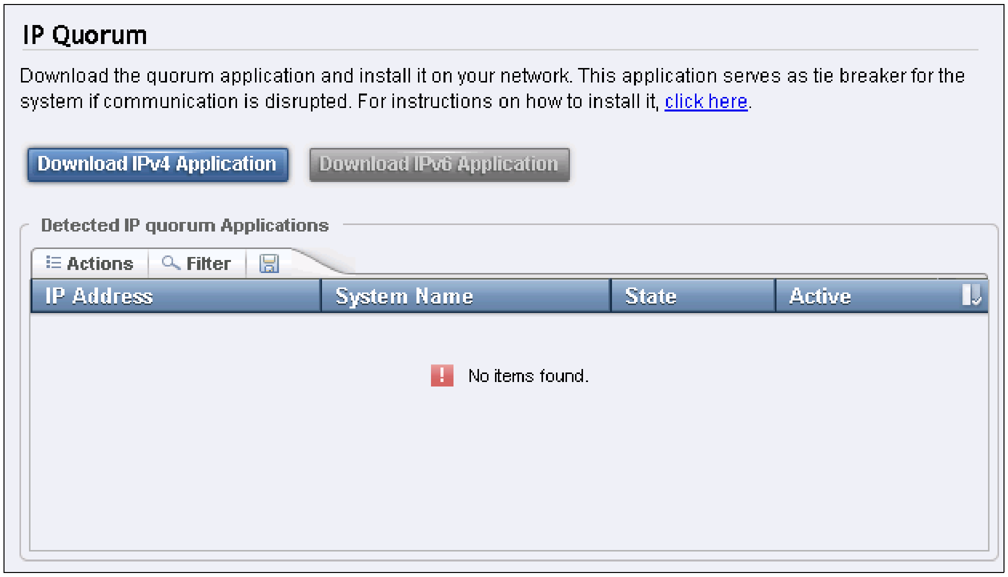
`http://www.redbooks.ibm.com/abstracts/sg248328.html?Open`

## 5.9.4 IP Quorum

Starting with IBM Spectrum Virtualize V7.6, a new feature was introduced for enhanced stretched systems, the IP Quorum application. Using an IP-based quorum application as the quorum device for the third site, no Fibre Channel connectivity is required. Java applications run on hosts at the third site.

To start with IP Quorum, complete the following steps:

1. If your IIBM Storwize V7000 is configured with IP addresses version 4, click on **Download IPv4 Application,** or select **Download IPv6 Application** for systems running with IP version 6. In our case, IPv4 is the option as shown in Figure 5-59.



IP Quorum

Download the quorum application and install it on your network. This application serves as tie breaker for the system if communication is disrupted. For instructions on how to install it, click here.

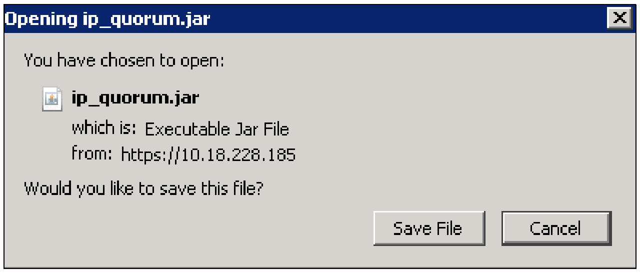
Download IPv4 Application
Download IPv6 Application
Detected IP quorum Applications
Actions
Filter

| IP Address | System Name | State | Active |
|---|---|---|---|

No items found.

Figure 5-59 IP Quorum

2. Click **Download IPv4 Application** and the IBM Spectrum Virtualize will generate an IP Quorum Java application as shown in Figure 5-60. The application can be saved and installed in a host that is to run the IP quorum application.



Opening ip_quorum.jar

You have chosen to open:

ip_quorum.jar

which is: Executable Jar File

from: https://10.18.228.185

Would you like to save this file?

Figure 5-60 IP Quorum Java Application

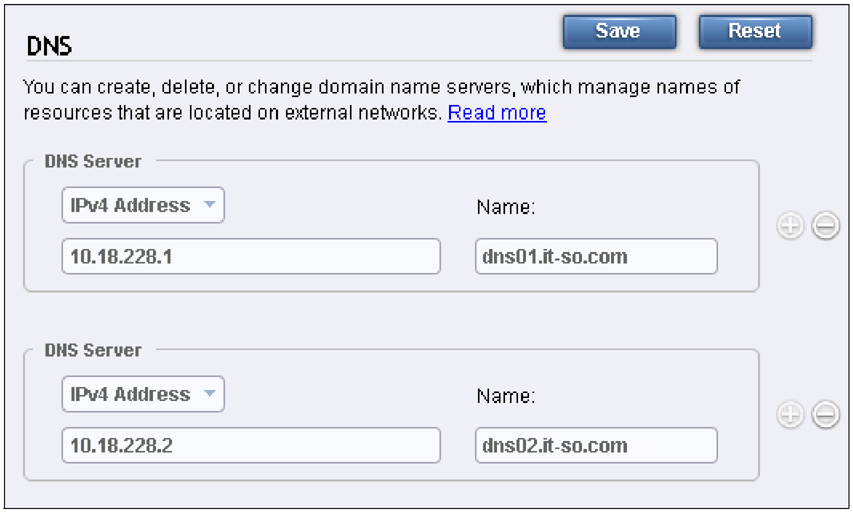
3. On the host, you must use the Java command line to initialize the IP quorum application. Change to the folder where the application is located and run **java -jar ip_quorum.jar**.

## 5.9.5 Domain Name Server

Introduced within V7.8, IBM Spectrum Virtualize allows Domain Name Server (DNS) entries to be manually setup in the IBM Storwize V7000. The information about the DNS servers in the IBM Storwize V7000 helps the system to access the DNS servers to resolve names of the computer resources that are in the external network.

To view and configure DNS server information in IBM Spectrum Virtualize, complete the following steps:

1. In the left panel, click on **DNS** icon and insert the IP address and the Name of each DNS server. The IBM Spectrum Virtualize supports up 2 DNS Servers, IPv4 or IPv6. See Figure 5-61.



DNS

You can create, delete, or change domain name servers, which manage names of resources that are located on external networks. Read more

DNS Server

IPv4 Address

Name:

10.18.228.1

dns01.it-so.com

DNS Server

IPv4 Address

Name:

10.18.228.2

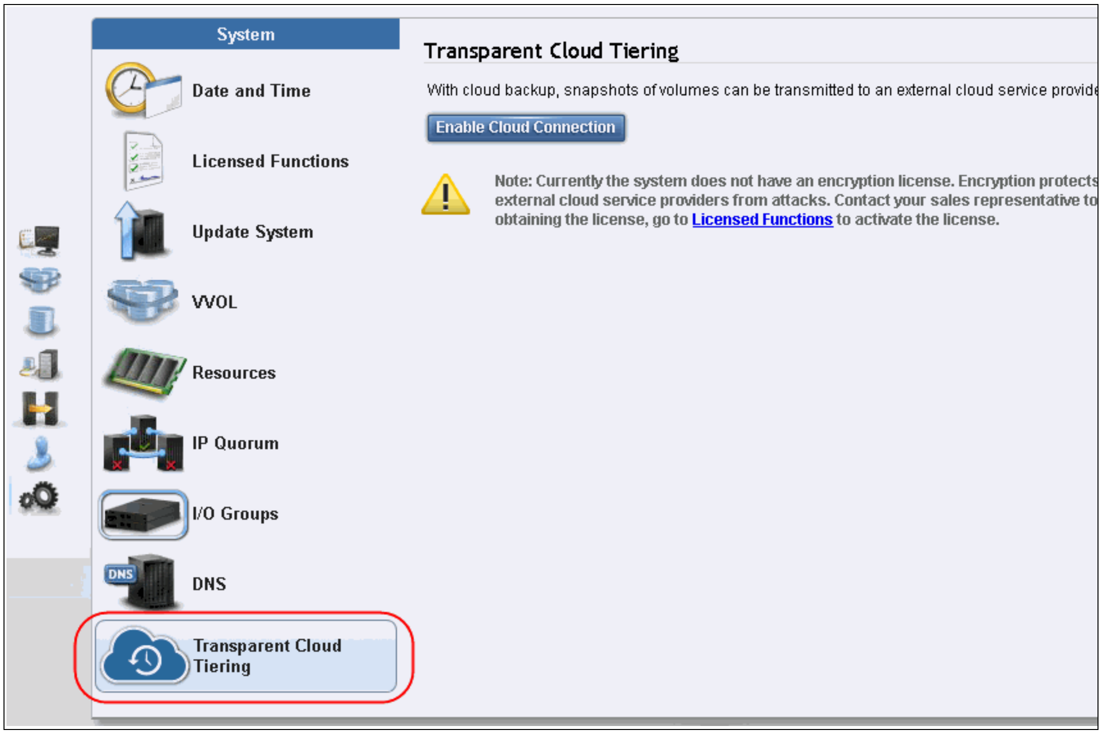
dns02.it-so.com

*Figure 5-61 DNS information*

2. Click **Save** after you complete typing the DNS server information.

## 5.9.6 Transparent Cloud Tiering

To view your IBM Spectrum Virtualize cloud provider settings, from the IBM Storwize V7000 Settings pane, move the pointer over **Settings** and click **System,** then select **Transparent Cloud Tiering** as shown in Figure 5-62.



Transparent Cloud Tiering

With cloud backup, snapshots of volumes can be transmitted to an external cloud service provide

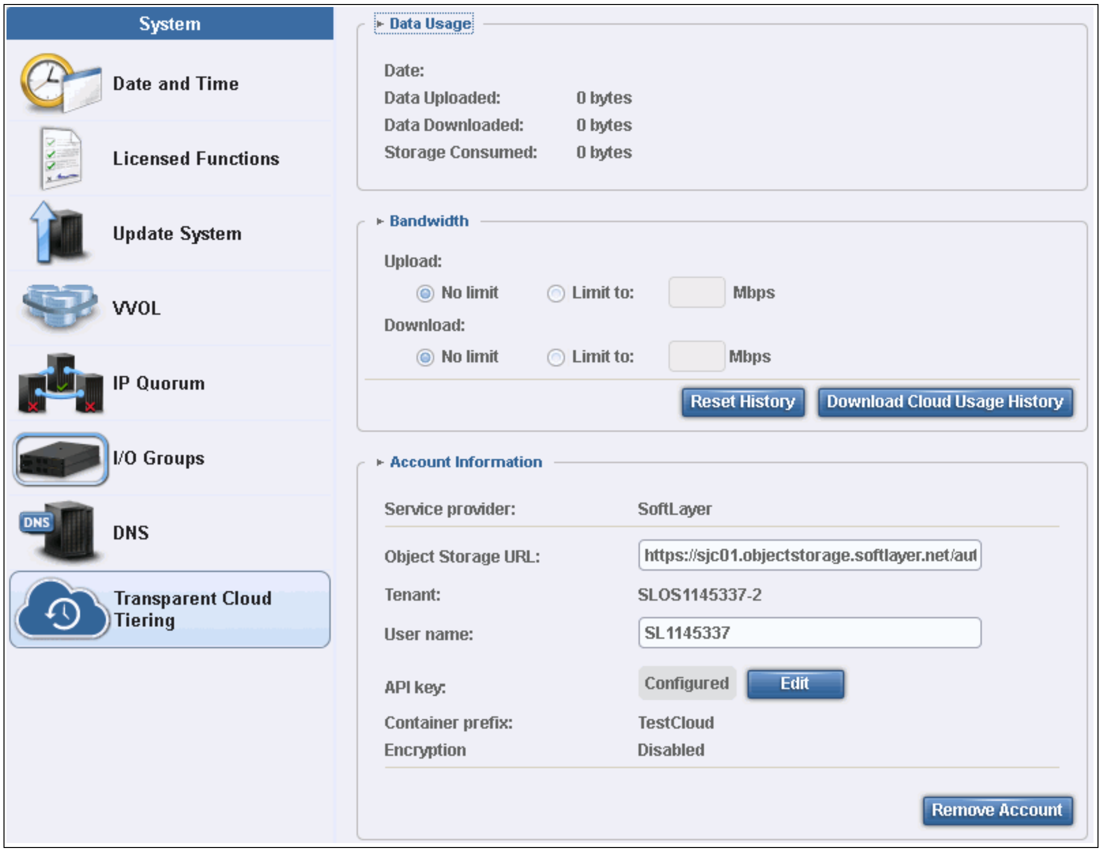
Note: Currently the system does not have an encryption license. Encryption protects external cloud service providers from attacks. Contact your sales representative to obtaining the license, go to Licensed Functions to activate the license.

*Figure 5-62 Transparent cloud tiering settings*

Using this view, you can enable and disable features of your Transparent Cloud Tiering settings and update the system information with respect to your cloud service provider. This pane allows you to set a number of options as follows:

- Cloud service provider
- Object Storage URL
- The Tenant or the container information that is associated to your cloud object storage
- User name of the cloud object account
- API Key
- The container prefix or location of your object
- Encryption
- Bandwidth

Figure 5-63 shows an example of Transparent Cloud Tiering configured in IBM Spectrum Virtualize.



Data Usage

Date:
Data Uploaded: 0 bytes
Data Downloaded: 0 bytes
Storage Consumed: 0 bytes

Bandwidth

Upload:
● No limit ○ Limit to: Mbps
Download:
● No limit ○ Limit to: Mbps

Account Information

Service provider: SoftLayer
Object Storage URL: https://sjc01.objectstorage.softlayer.net/aut
Tenant: SLOS1145337-2
User name: SL1145337
API key: Configured
Container prefix: TestCloud
Encryption Disabled

*Figure 5-63 Transparent Cloud Tiering settings*

## 5.10 Setting GUI preferences

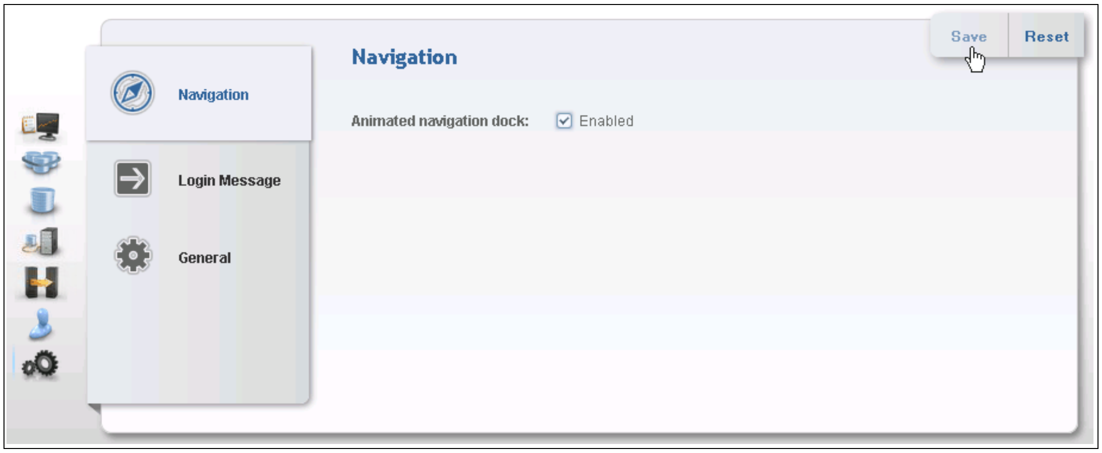
The menu GUI Preferences consists of three options:

- Navigation
- Login/Message
- General

### 5.10.1 Navigation

Starting with IBM Spectrum Virtualize V7.6., the dynamic function of the left menu is disabled by default and the icons are of a static size. To enable dynamic appearance, navigate to **Settings** → **GUI Preferences** → **Navigation** as shown in Figure 5-64. Select the check box to enable animation (dynamic function) and click the **Save** button.

For changes to take effect log in again or refresh the GUI cache from the General pane in GUI Preferences.



Navigation

Animated navigation dock: ☑ Enabled

*Figure 5-64 Navigation menu*

### 5.10.2 Login Message

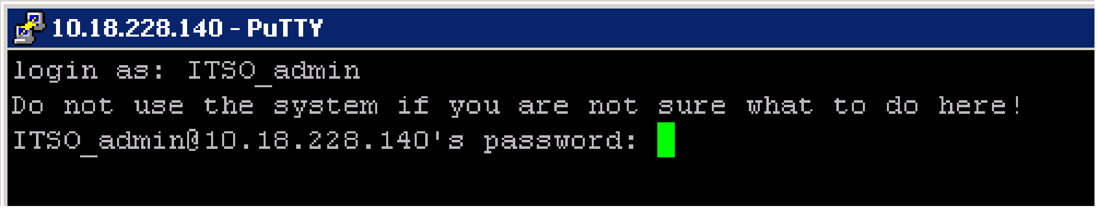
IBM Spectrum Virtualize V7.6 and higher enables administrators to configure the welcome banner, which is a text message that displays either in the GUI login window or at the command-line interface (CLI) login prompt. The content of the welcome message is helpful when you need to notify users about some important information about system, for example, security warnings or location description.

The welcome banner (or login message) can be enabled from the GUI or the CLI. To define a message, use the commands outlined in Example 5-3. Define (copy) the text file that contains the welcome message to your configuration node and enable it in the CLI. In our case, the content of the file is located in `/tmp/banner`.

*Example 5-3 Configure welcome message*

```
IBM_2076:ITSO_V7000_01:ITSO_admin>chbanner -file=/tmp/banner -enable
IBM_2076:ITSO_V7000_01:ITSO_admin>

## where the file contains:
node1:/tmp # cat /tmp/banner
Do not use the system if you are not sure what to do here!

## To disable showing of the message use:
IBM_2076:ITSO_V7000_01:ITSO_admin>chbanner -disable
```

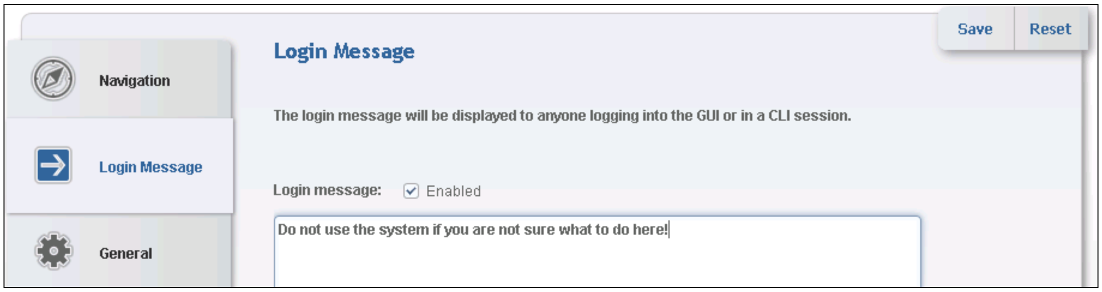
The banner message is also illustrated in the CLI login prompt window as shown in Figure 5-65.



```
login as: ITSO_admin
Do not use the system if you are not sure what to do here!
ITSO_admin@10.18.228.140's password:
```

*Figure 5-65 Banner message in the CLI*

It can be defined and enabled using the GUI by entering the text message and click **Save** as shown Figure 5-66.



Login Message

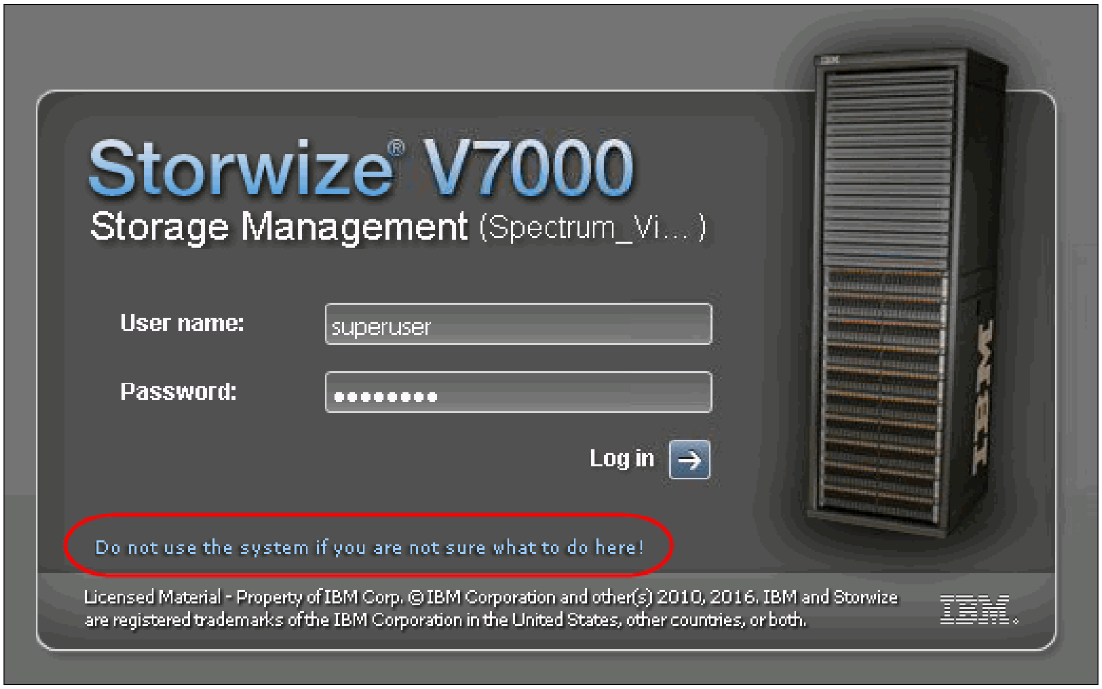
The login message will be displayed to anyone logging into the GUI or in a CLI session.

Login message: ☑ Enabled

Do not use the system if you are not sure what to do here!

*Figure 5-66 Enabling login message*

The result of the action before is as illustrated in Figure 5-67.



User name: superuser

Password: ••••••••

Do not use the system if you are not sure what to do here!

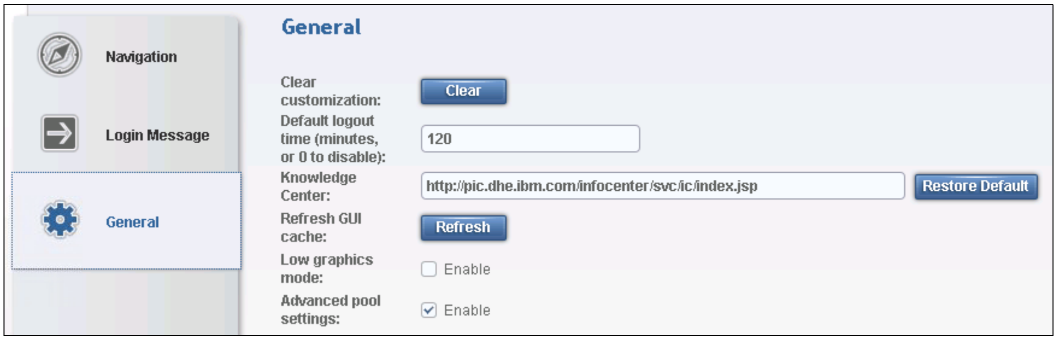
Licensed Material - Property of IBM Corp. © IBM Corporation and other(s) 2010, 2016. IBM and Storwize are registered trademarks of the IBM Corporation in the United States, other countries, or both.

*Figure 5-67 GUI banner*

## 5.10.3 General settings

Complete the following steps to configure general GUI preferences:

1. From the Settings window, move the pointer over Settings and click **GUI Preferences** (Figure 5-68 on page 164).



General

Clear customization:

Default logout time (minutes, or 0 to disable): 120

Knowledge Center: http://pic.dhe.ibm.com/infocenter/svc/ic/index.jsp

Refresh GUI cache:

Low graphics mode: ☐ Enable

Advanced pool settings: ☑ Enable

*Figure 5-68 General GUI Preferences window*

2. You can configure the following elements:

- Refresh GUI cache

  This option causes the GUI to refresh all of its views and clears the GUI cache. The GUI looks up every object again.

- Clear Customization
- This option deletes all GUI preferences that are stored in the browser and restores the default preferences.
- IBM Knowledge Center
- You can change the URL of the IBM Knowledge Center for IBM Spectrum Virtualize.
- The accessibility option enables Low graphic mode when the system is connected through a slower network.
- Advanced pool settings allow you to select the extent size during storage pool creation.
- Default logout time in minutes after inactivity in the established session.

6

# Storage pools

This chapter describes how IBM Storwize V7000 manages physical storage resources. All storage resources that are under the system control are managed using *storage pools*. Storage pools aggregate internal and external capacity and provide the containers in which volumes can be created. Storage pools make it easy to dynamically allocate resources, maximize productivity, and reduce costs.

This chapter includes the following topics:

- Working with storage pools
- Working with managed disks
- Working with internal drives
- Working with external storage controllers

## 6.1 Working with storage pools

Storage pools act as containers for MDisks and provision the capacity to volumes. MDisks can be provisioned through internal or external storage. MDisks created from internal storage are created as RAID arrays.

Figure 6-1 provides an overview of how storage pools, MDisks, and volumes are related. This panel is available by browsing to **Monitoring** → **System** and clicking **Overview** on the upper-right corner of the panel.

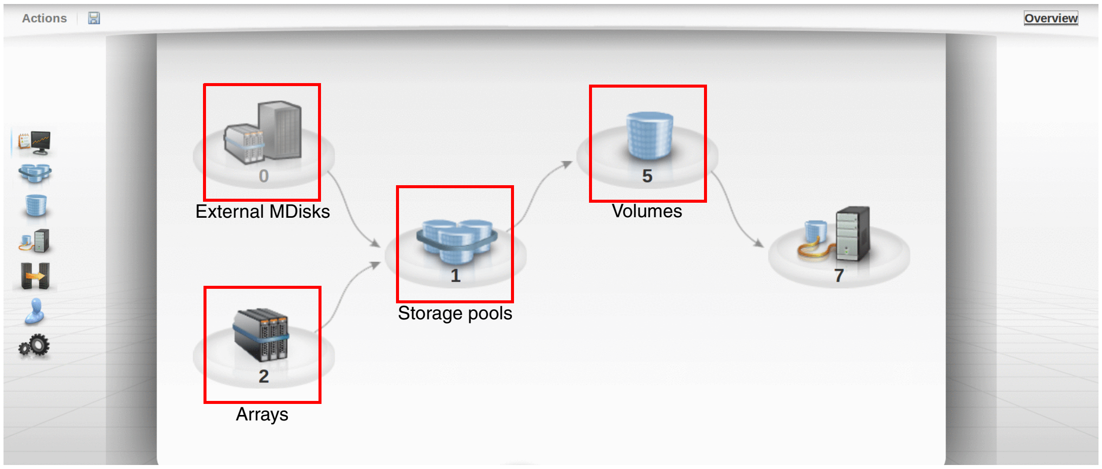
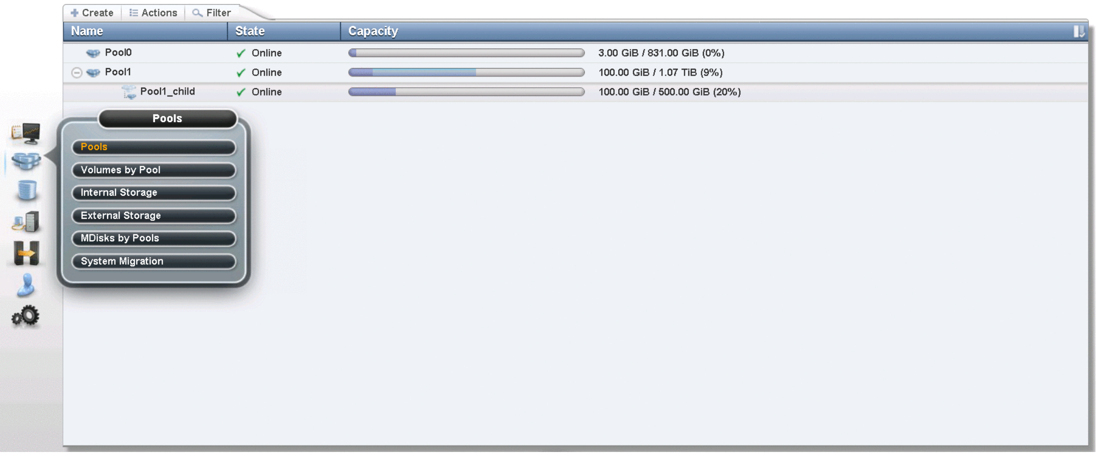


*Figure 6-1 Relationship between MDisks, storage pools, and volumes*

Storwize V7000 organizes storage into pools to ease storage management and make it more efficient. All MDisks in a pool are split into extents of the same size and volumes are created from the extents available. The extent size is a property of the storage pool and when an MDisk is added to a pool the size of the extents the MDisk is divided into depends on the attribute of the pool to which the MDisk was added.

Storage pools can be further divided into sub-containers called child pools. Child pools inherit the properties of the parent pool and can also be used to provision volumes.

Storage pools are managed either via the Pools panel or via the MDisks by Pool panel. Both panels allow you to execute the same actions on parent pools; however, actions on child pools can only be performed through the Pools panel. To access the Pools panel browse to **Pools** → **Pools**, as shown in Figure 6-2 on page 167.



| Name | State | Capacity |
|---|---|---|
| Pool0 | ✓ Online | 3.00 GiB / 831.00 GiB (0%) |
| Pool1 | ✓ Online | 100.00 GiB / 1.07 TiB (9%) |
| Pool1_child | ✓ Online | 100.00 GiB / 500.00 GiB (20%) |

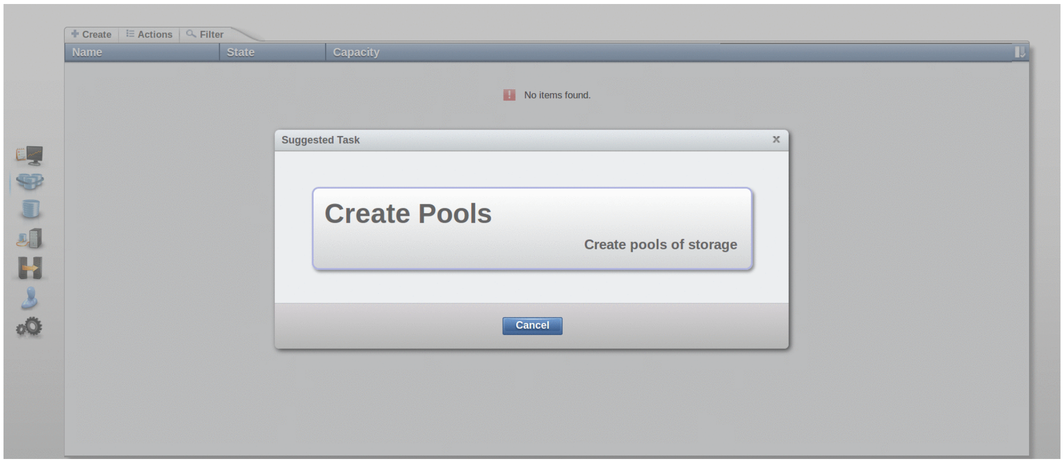
*Figure 6-2 Accessing the Pools panel*

The panel lists all storage pools available in the system. If a storage pool has child pools, you can toggle the sign to the left of the storage pool icon to either show or hide the child pools.

## 6.1.1 Creating storage pools

There are three different alternatives to create a new storage pool:

- If no storage pools are configured you are prompted to create them when you log into the management GUI, as shown in Figure 6-3. Click **Create Pools**.



No items found.

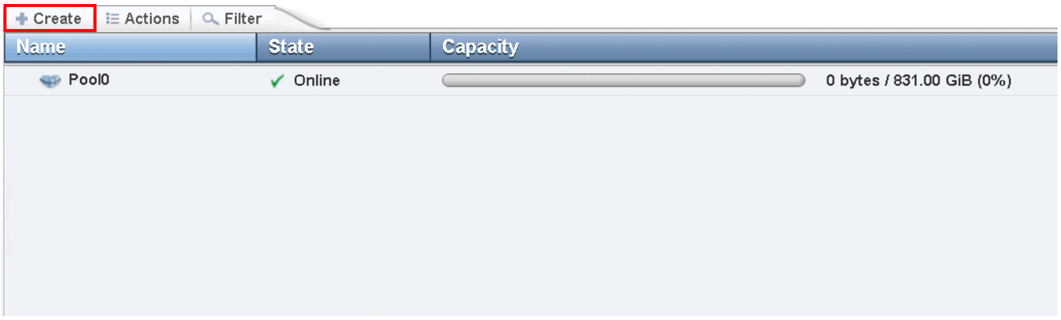
Suggested Task

Create Pools

Create pools of storage

*Figure 6-3 Suggested task to create pools*

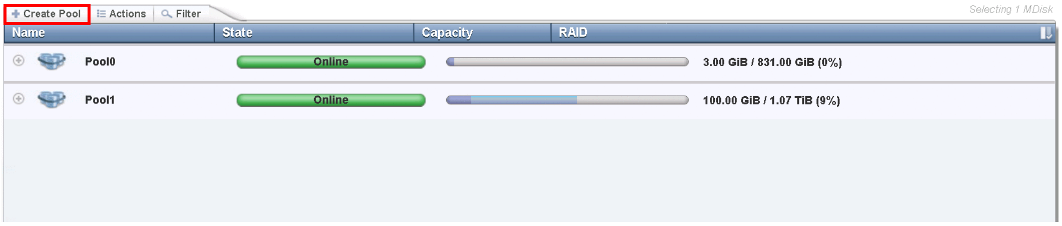
- Navigate to **Pools** → **Pools** and click **Create**, as shown in Figure 6-4 on page 168.



| Name | State | Capacity |
|---|---|---|
| Pool0 | ✓ Online | 0 bytes / 831.00 GiB (0%) |

*Figure 6-4 Option to create a storage pool in the Pools panel*

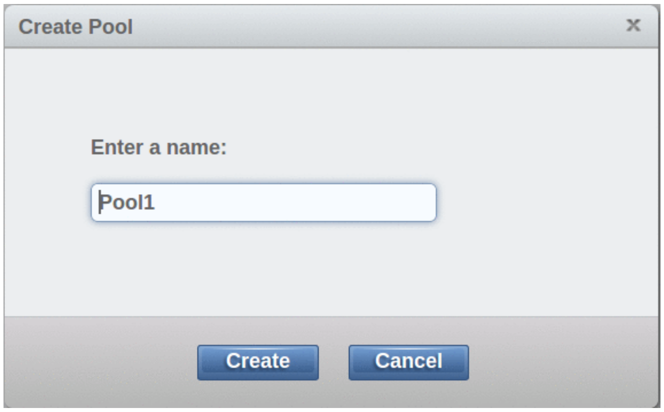
- Navigate to Pools → MDisks by Pools and click **Create Pool**, as shown in Figure 6-5.



| Name | State | Capacity | RAID |
|---|---|---|---|
| Pool0 | Online | 3.00 GiB / 831.00 GiB (0%) | |
| Pool1 | Online | 100.00 GiB / 1.07 TiB (9%) | |

*Figure 6-5 Option to create a storage pool in the MDisks by Pools panel*

All alternatives open the dialog box shown in Figure 6-6.

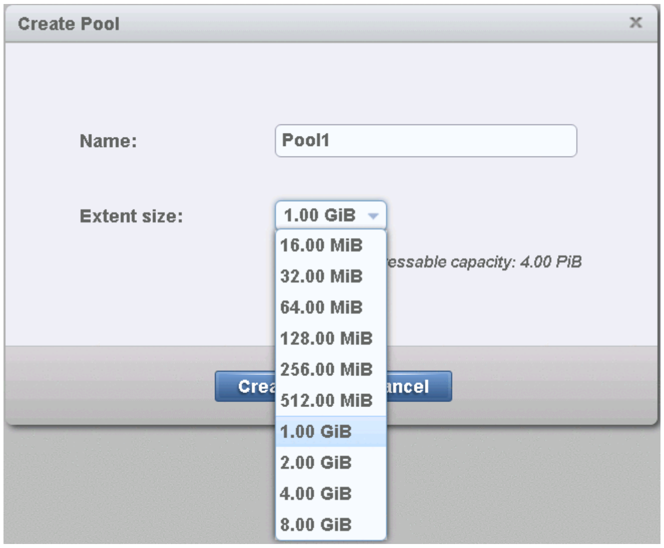

Create Pool

Enter a name: Pool1

*Figure 6-6 Creating a new storage pool*

If advanced pool settings are enabled, you can additionally select a extent size at creation time, as shown in Figure 6-7 on page 169.

> **Note:** Every storage pool created via the GUI has a default extent size of 1 GB. The size of the extent is selected at creation time and cannot be changed later. If you want to specify a different extent size, browse to **Settings** → **GUI Preferences** and check **Advanced pool settings**.



Name: Pool1

Extent size: 1.00 GiB

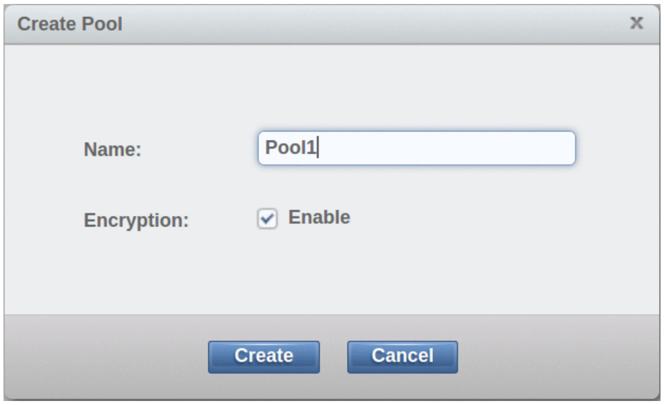
- 16.00 MiB
- 32.00 MiB
- 64.00 MiB
- 128.00 MiB
- 256.00 MiB
- 512.00 MiB
- 1.00 GiB
- 2.00 GiB
- 4.00 GiB
- 8.00 GiB

essable capacity: 4.00 PiB

Figure 6-7 Choosing the extent size of a new storage pool

If encryption is enabled, you can additionally select whether the storage pool is encrypted, as shown in Figure 6-8 on page 169.

**Note:** The encryption setting of a storage pool is selected at creation time and cannot be changed later. By default, if encryption is enabled, encryption is selected. For more information about encryption and encrypted storage pools see Chapter 12, “Encryption” on page 563.

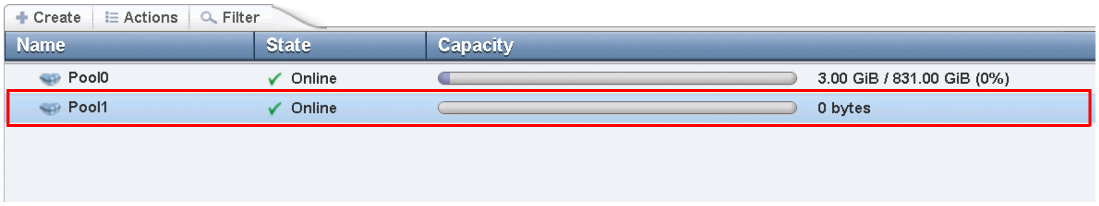


Name: Pool1

Encryption: ☑ Enable

Figure 6-8 Choosing the encryption setting of a new storage pool

Enter the name for the pool and click **Create**. The new pool is created and is included in the list of storage pools with zero bytes, as shown in Figure 6-9.



| Name | State | Capacity |
|---|---|---|
| Pool0 | ✓ Online | 3.00 GiB / 831.00 GiB (0%) |
| Pool1 | ✓ Online | 0 bytes |

Figure 6-9 New empty pool

## 6.1.2 Actions on storage pools

There are a number of actions that can be performed on storage pools. To select an action, select the storage pool and click **Actions**, as shown in Figure 6-10. Alternatively, right-click the storage pool.

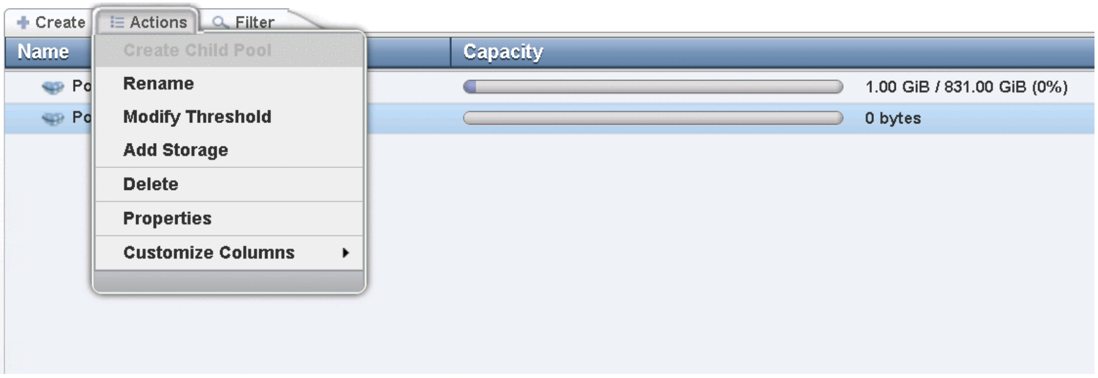


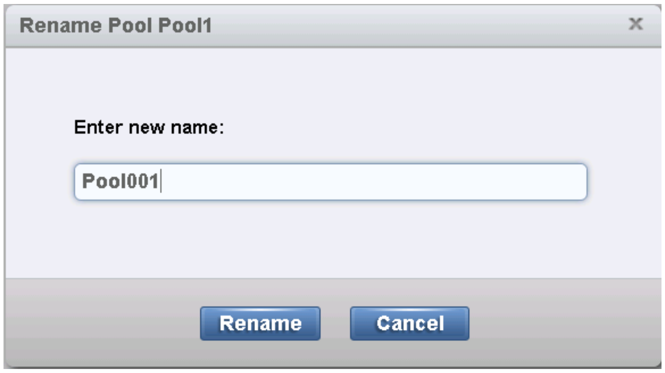
*Figure 6-10 Storage pool actions*

### Create child pool

Selecting **Create Child Pool** starts the wizard to create a child storage pool. For information about child storage pools and a detailed description of this wizard see “Child storage pools” on page 171.

### Rename

Selecting **Rename** allows you to modify the name of a storage pool, as shown in Figure 6-11. Enter the new name and click **Rename**.

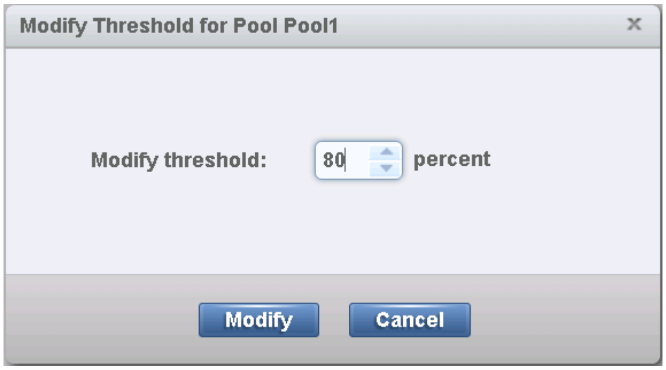

Rename Pool Pool1

Enter new name: Pool001

*Figure 6-11 Renaming a storage pool*

### Modify threshold

The storage pool threshold refers to the percentage of storage capacity that must be in use for a warning event to be generated. The threshold is especially useful when using thin-provisioned volumes that are configured to expand automatically. The threshold can be modified by selecting **Modify Threshold** and entering the new value, as shown in Figure 6-12. The default threshold is 80%. Warnings can be disabled by setting the threshold to 0%.


Modify Threshold for Pool Pool1

Modify threshold: 80 percent

*Figure 6-12 Modifying a storage pool threshold*

### Add storage

Selecting **Add Storage** starts the wizard to assign storage to the pool. For a detailed description of this wizard see “Assigning managed disks to storage pools” on page 176.

### Delete

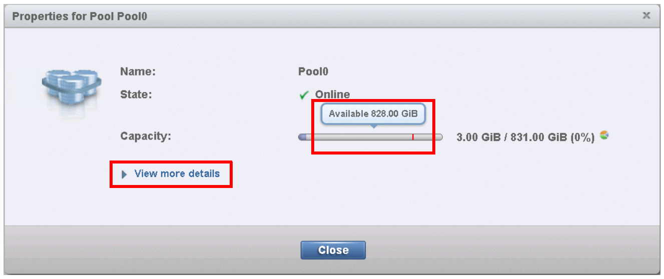
A storage pool can be deleted via the GUI only if there are no volumes associated with it. Selecting **Delete** deletes the pool immediately without any additional confirmation.

> **Note:** If there are volumes in the pool **Delete** cannot be selected. If that is the case, either delete the volumes or move them to another storage pool before proceeding. To move a volume you can either migrate it or use volume mirroring. For information about volume migration and volume mirroring, see Chapter 7, “Volumes” on page 213.

After you delete a pool, all the managed or image mode MDisks in the pool return to a status of unmanaged, all the array mode MDisks in the pool are deleted, and all member drives return to a status of candidate.

### Properties

Selecting **Properties** displays information about the storage pool. Additional information is available by clicking **View more details** and by hovering over the elements on the window, as shown in Figure 6-13.


Properties for Pool Pool0

Name: Pool0

State: ✓ Online

Capacity: 3.00 GiB / 831.00 GiB (0%)

Available 828.00 GiB

*Figure 6-13 Storage pool properties*

## 6.1.3 Child storage pools

A *child storage pool* is a storage pool created within a storage pool. The storage pool in which the child storage pool is created is called *parent storage pool*.

Unlike a parent pool, a child pool does not contain MDisks; its capacity is provided exclusively by the parent pool in the form of extents. The capacity of a child pool is set at creation time, but can be modified later nondisruptively. The capacity must be a multiple of the parent pool extent size and must be smaller than the free capacity of the parent pool.

Child pools are useful when the capacity allocated to a specific set of volumes must be controlled. For example, child pools are used with VMware vSphere Virtual Volumes so part of the storage pool can be assigned to and managed by the IBM Spectrum Control Base.

**Note:** Creation and management of child pools used for Virtual Volumes must be done either through the command-line interface or through the IBM Spectrum Control. You can use the management GUI to view these child pools and their properties.

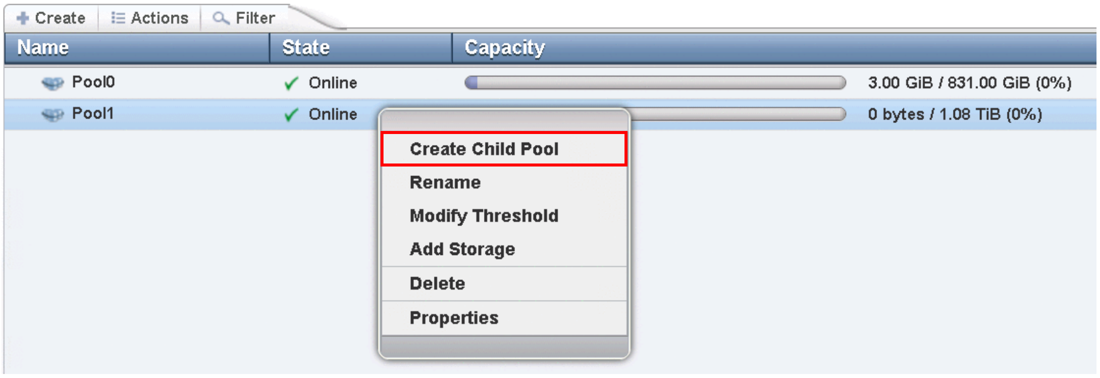
Child pools inherit most properties from their parent pools and these cannot be changed. The inherited properties include:

- Extent size
- Easy Tier setting
- Encryption setting, but only if the parent pool is encrypted

**Note:** For information about encryption and encrypted child storage pools, see Chapter 12, “Encryption” on page 563.

## Creating a child storage pool

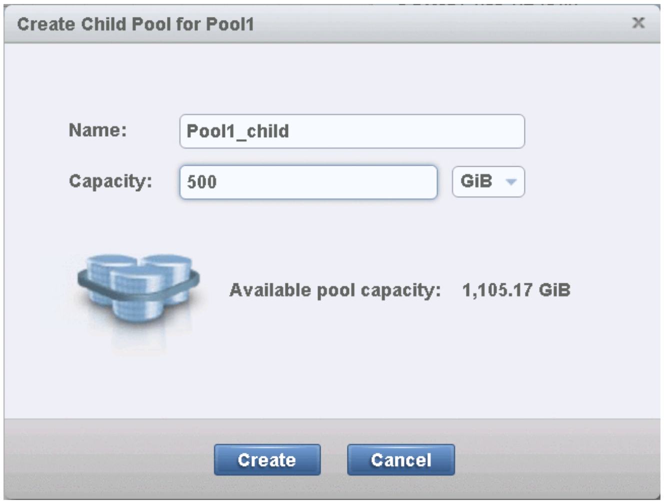
To create a child pool browse to **Pools** → **Pools**, right-click the parent pool, and select **Create Child Pool**, as shown in Figure 6-14 on page 172.



| Name | State | Capacity |
|---|---|---|
| Pool0 | ✓ Online | 3.00 GiB / 831.00 GiB (0%) |
| Pool1 | ✓ Online | 0 bytes / 1.08 TiB (0%) |

*Figure 6-14 Action to create a child pool*

Enter the name and capacity of the child pool and click **Create**, as shown in Figure 6-15.



Create Child Pool for Pool1

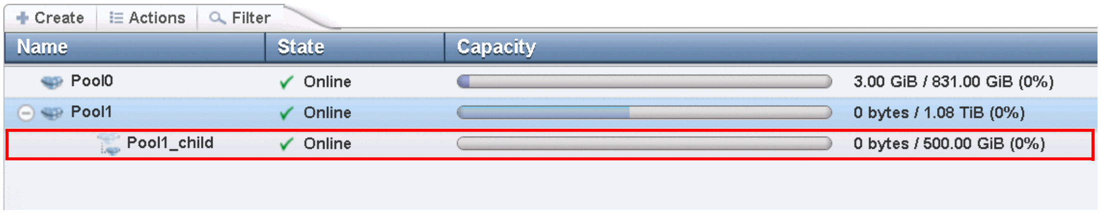
Name: Pool1_child

Capacity: 500 GiB

Available pool capacity: 1,105.17 GiB

*Figure 6-15 Creating a child pool*

After the child pool is created it is listed in the Pools panel under its parent pool, as shown in Figure 6-16. Toggle the sign to the left of the storage pool icon to either show or hide the child pools.



| Name | State | Capacity |
|---|---|---|
| Pool0 | ✓ Online | 3.00 GiB / 831.00 GiB (0%) |
| Pool1 | ✓ Online | 0 bytes / 1.08 TiB (0%) |
| Pool1_child | ✓ Online | 0 bytes / 500.00 GiB (0%) |

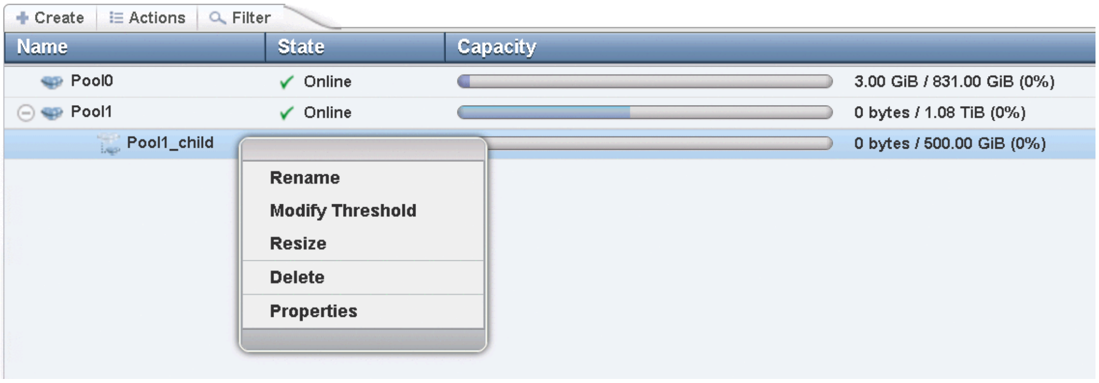
*Figure 6-16 List of pools and associated child pools*

Creation of a child pool within a child pool is not possible.

## Actions on child storage pools

All actions supported for parent storage pools are supported for child storage pools, with the exception of **Add Storage**. Child pools additionally support the **Resize** action.

To select an action right-click the child storage pool, as shown in Figure 6-17. Alternatively, select the storage pool and click **Actions**.



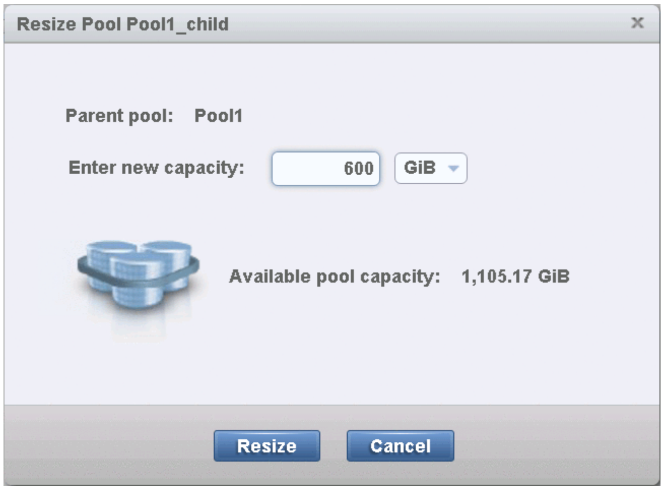
| Name | State | Capacity |
|---|---|---|
| Pool0 | ✓ Online | 3.00 GiB / 831.00 GiB (0%) |
| Pool1 | ✓ Online | 0 bytes / 1.08 TiB (0%) |
| Pool1_child | | 0 bytes / 500.00 GiB (0%) |

*Figure 6-17 Child pool actions*

### *Resize*

Selecting **Resize** allows you to increase or decrease the capacity of the child storage pool, as shown in Figure 6-18. Enter the new pool capacity and click **Resize**.

> **Note:** You cannot shrink a child pool below its real capacity. This means that the new size of a child pool needs to be larger than the capacity used by its volumes.



Resize Pool Pool1_child

Parent pool: Pool1

Enter new capacity: 600 GiB

Available pool capacity: 1,105.17 GiB

*Figure 6-18 Resizing a child pool*

When the child pool is shrunk the system resets the warning threshold and issues a warning if the threshold is reached.

### *Delete*

Deleting a child pool is a task quite similar to deleting a parent pool. As with a parent pool, the **Delete** action is disabled if the child pool contains volumes. To move a volume to another pool you can use migration or volume mirroring in the same way you use them for parent pools. For information about volume migration and volume mirroring, see Chapter 7, “Volumes” on page 213.

> **Note:** A volume in a child pool can only be migrated to another child pool within the same parent pool or to its own parent pool. In any other case use volume mirroring instead.

> **Note:** During migration from a child pool to its parent pool, or vice versa, there is no real data movement. There is only a reassignment of extents between the pools.

After deleting a child pool the extents that it occupied return to the parent pool as free capacity.

## 6.1.4 Encrypted storage pools

Storwize V7000 supports two types of encryption: hardware encryption and software encryption.

Hardware encryption is implemented at an array level, whereas software encryption is implemented at a storage pool level. For information about encryption and encrypted storage pools, see Chapter 12, “Encryption” on page 563.

## 6.2 Working with managed disks

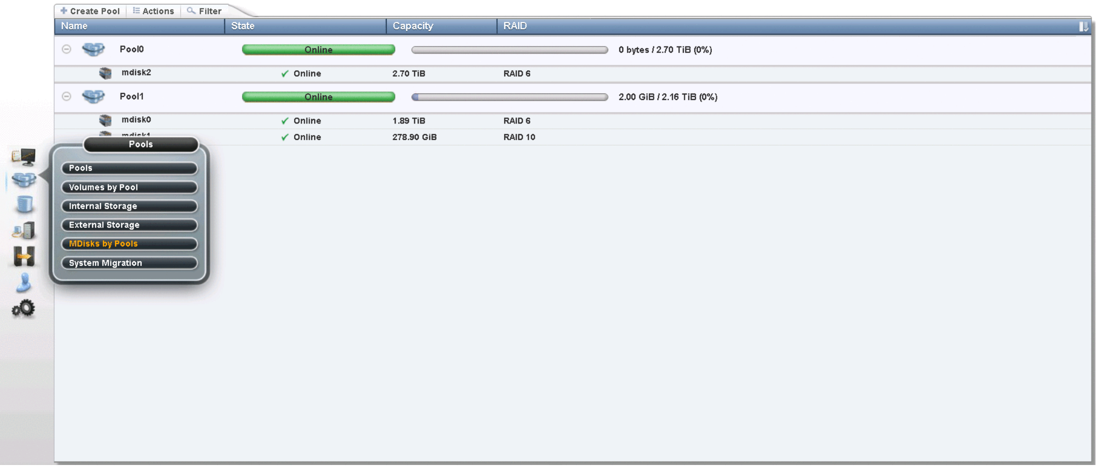
A storage pool is created as an empty container, with no storage assigned to it. Storage is then added in the form of MDisks. An MDisk can be either an array from internal storage or an LU from an external storage system. The same storage pool can include both internal and external MDisks.

Arrays are created from internal storage using RAID technology to provide redundancy and increased performance. The system supports two types of RAID: traditional RAID and distributed RAID. Arrays are assigned to storage pools at creation time and cannot be moved between storage pools. You cannot have an array that does not belong to any storage pool.

External MDisks can have one of three modes:

- **Unmanaged**. External MDisks are discovered by the system as unmanaged MDisks. An unmanaged MDisk is not a member of any storage pool, it is not associated with any volumes, and has no metadata stored on it. The system does not write to an MDisk that is in unmanaged mode, except when it attempts to change the mode of the MDisk to one of the other modes.
- **Managed**. When unmanaged MDisks are added to storage pools, they become managed. Managed mode MDisks are always members of a storage pool, and they contribute extents to the storage pool. This mode is the most common and normal mode for an MDisk.
- **Image**. Image mode provides a direct block-for-block translation from the MDisk to a volume. This mode is provided to satisfy the following major usage scenarios:
  - Virtualization of external LUs that contain data not written through the Storwize V7000.
  - Exporting MDisks from the Storwize V7000 after migration of volumes to image mode MDisks.

MDisks are managed via the MDisks by Pools panel. To access the MDisks by Pools panel browse to **Pools** → **MDisks by Pools**, as shown in Figure 6-19.



| Name | State | Capacity | RAID |
|---|---|---|---|
| Pool0 | Online | 0 bytes / 2.70 TiB (0%) | |
| mdisk2 | ✓ Online | 2.70 TiB | RAID 6 |
| Pool1 | Online | 2.00 GiB / 2.16 TiB (0%) | |
| mdisk0 | ✓ Online | 1.89 TiB | RAID 6 |
| mdisk1 | ✓ Online | 278.90 GiB | RAID 10 |

*Figure 6-19 Accessing the MDisks by Pools panel*

The panel lists all the MDisks available in the system under the storage pool to which they belong. Both arrays and external MDisks are listed.

External MDisks can additionally be managed through the External Storage panel. To access the External Storage panel browse to **Pools** → **External Storage**.

## 6.2.1 Assigning managed disks to storage pools

MDisks can be assigned to a storage pool at any time to increase the number of extents available in the pool. The system automatically balances volume extents between the MDisks to provide the best performance to the volumes.

Arrays are created and assigned to a storage pool at the same time.

To assign MDisks to a storage pool navigate to **Pools** → **MDisks by Pools** and choose one of the following options:

- Option 1: Select **Add Storage** on the right side of the storage pool, as shown in Figure 6-20. The Add Storage button is shown only when the pool has no capacity assigned or when the pool capacity usage is over the warning threshold.

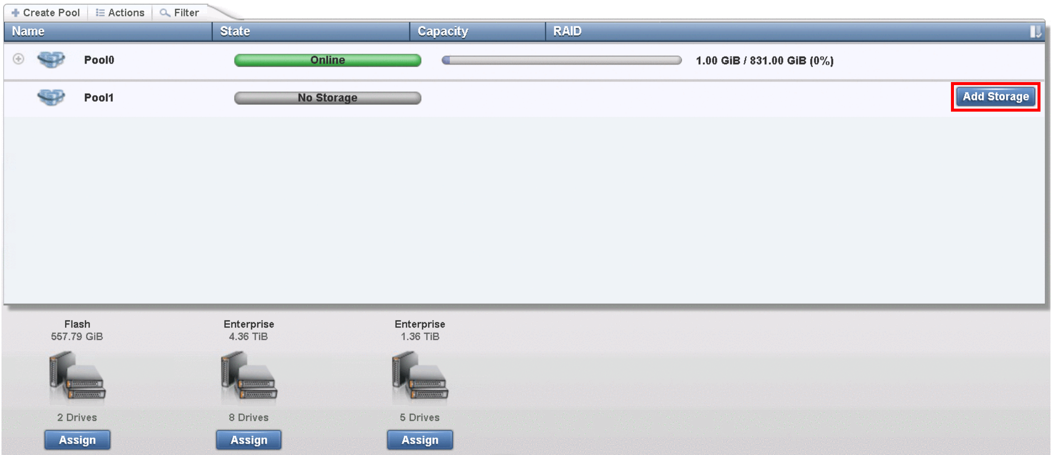


*Figure 6-20 Adding storage: option 1*

- Option 2: Right-click the pool and select **Add Storage**, as shown in Figure 6-21.

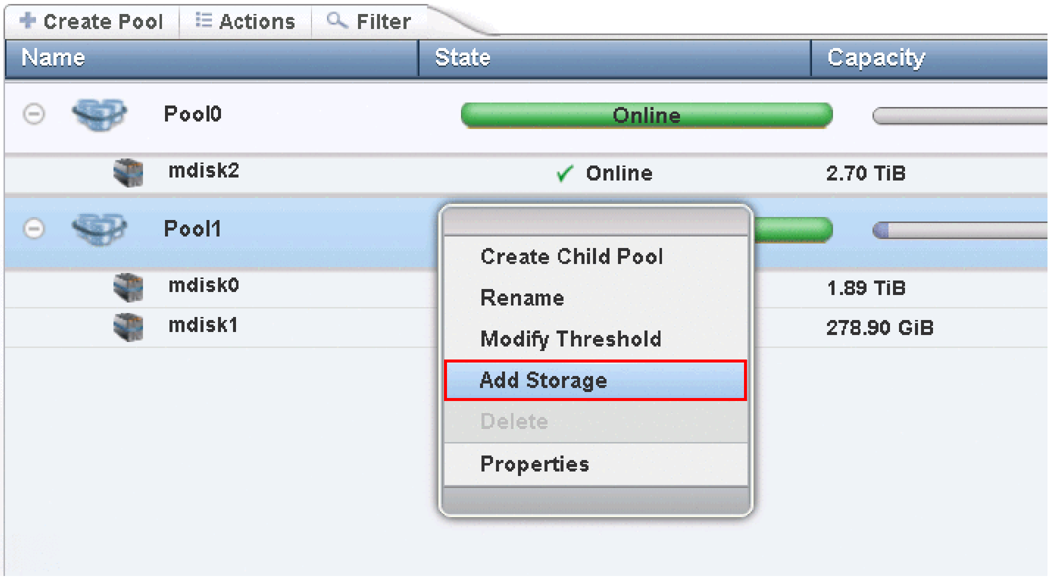


*Figure 6-21 Adding storage: option 2*

- Option 3: Select **Assign** under a specific drive class or external storage controller, as shown in Figure 6-22.

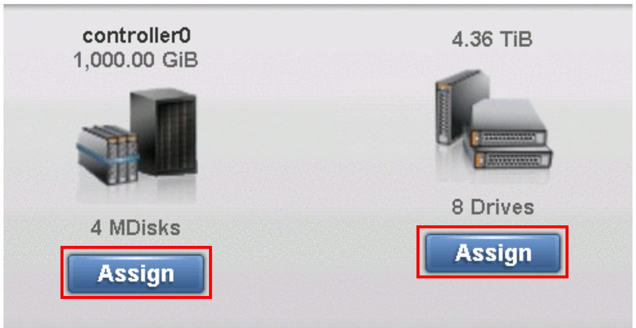


*Figure 6-22 Adding storage: option 3*

Both options 1 and 2 start the configuration wizard shown in Figure 6-23. If no external storage is attached the External option is not shown.

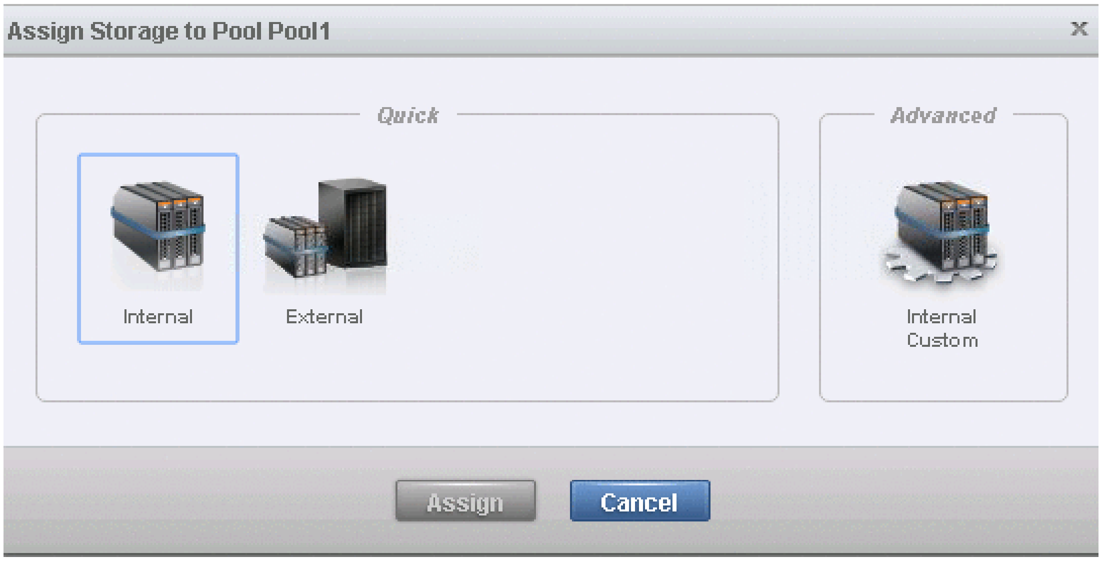


*Figure 6-23 Assigning storage to storage pool*

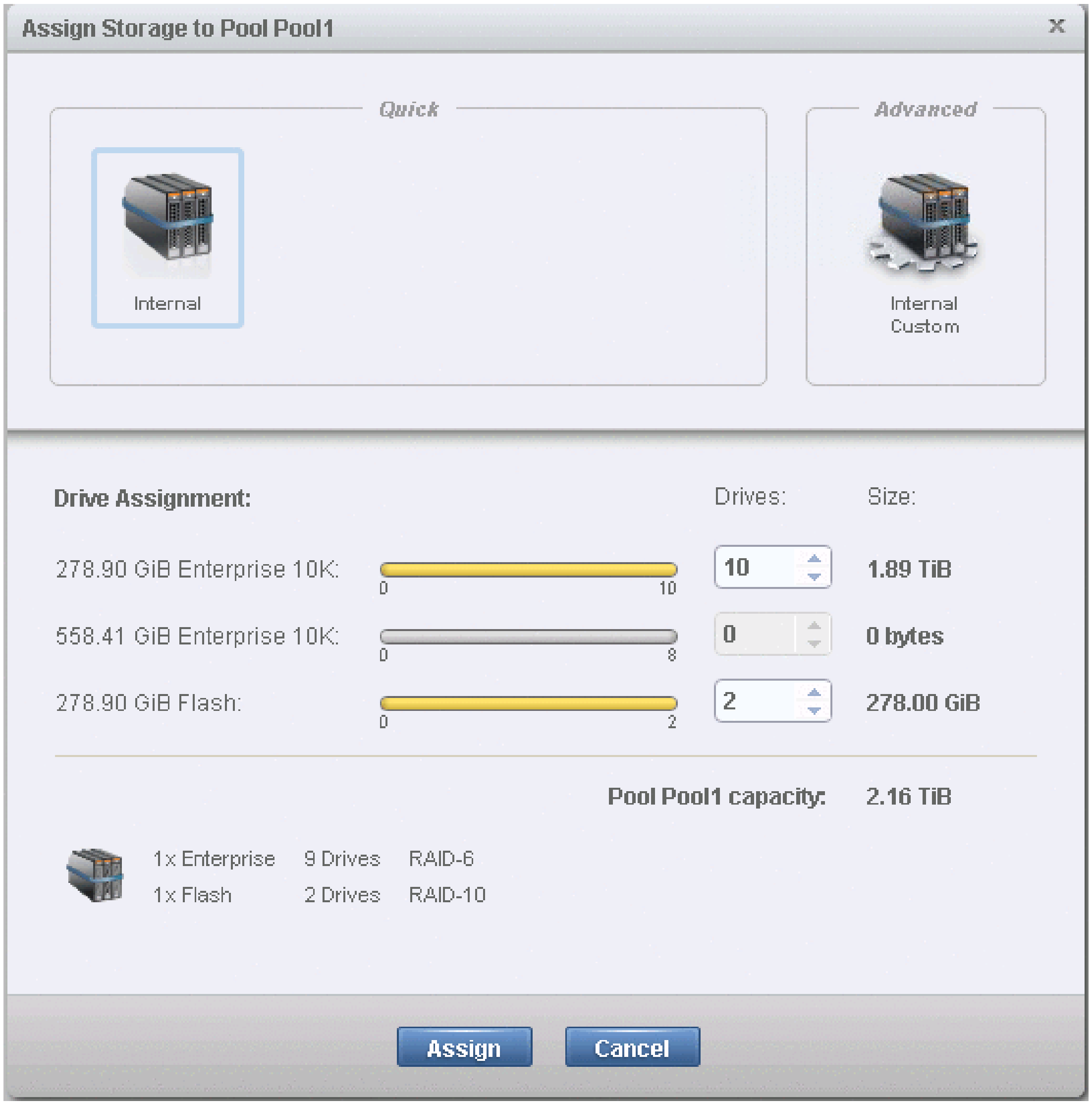
Option 3 starts the quick internal wizard for the selected drive class only or the quick external wizard for the selected external controller only.

## Quick internal configuration

Selecting **Internal** suggests a configuration for internal drives based on RAID configuration presets, taking into account drive class and number of drives available. It automatically defaults parameters such as stripe width, number of spares (for traditional RAID), number of rebuild areas (for distributed RAID), and number of drives of each class. The number of drives is the only value that can be adjusted.

By default, the system recommends distributed arrays for most new configurations. There are exceptions, however; distributed RAID is not suggested if the drive count is lower than an internally determined value, for example. For information about traditional and distributed RAID see Chapter 6.2.2, “Traditional and distributed RAID” on page 183.

Figure 6-24 on page 179 shows an example of a quick configuration.



Drive Assignment:

| | Drives: | Size: |
|---|---|---|
| 278.90 GiB Enterprise 10K: | 10 | 1.89 TiB |
| 558.41 GiB Enterprise 10K: | 0 | 0 bytes |
| 278.90 GiB Flash: | 2 | 278.00 GiB |

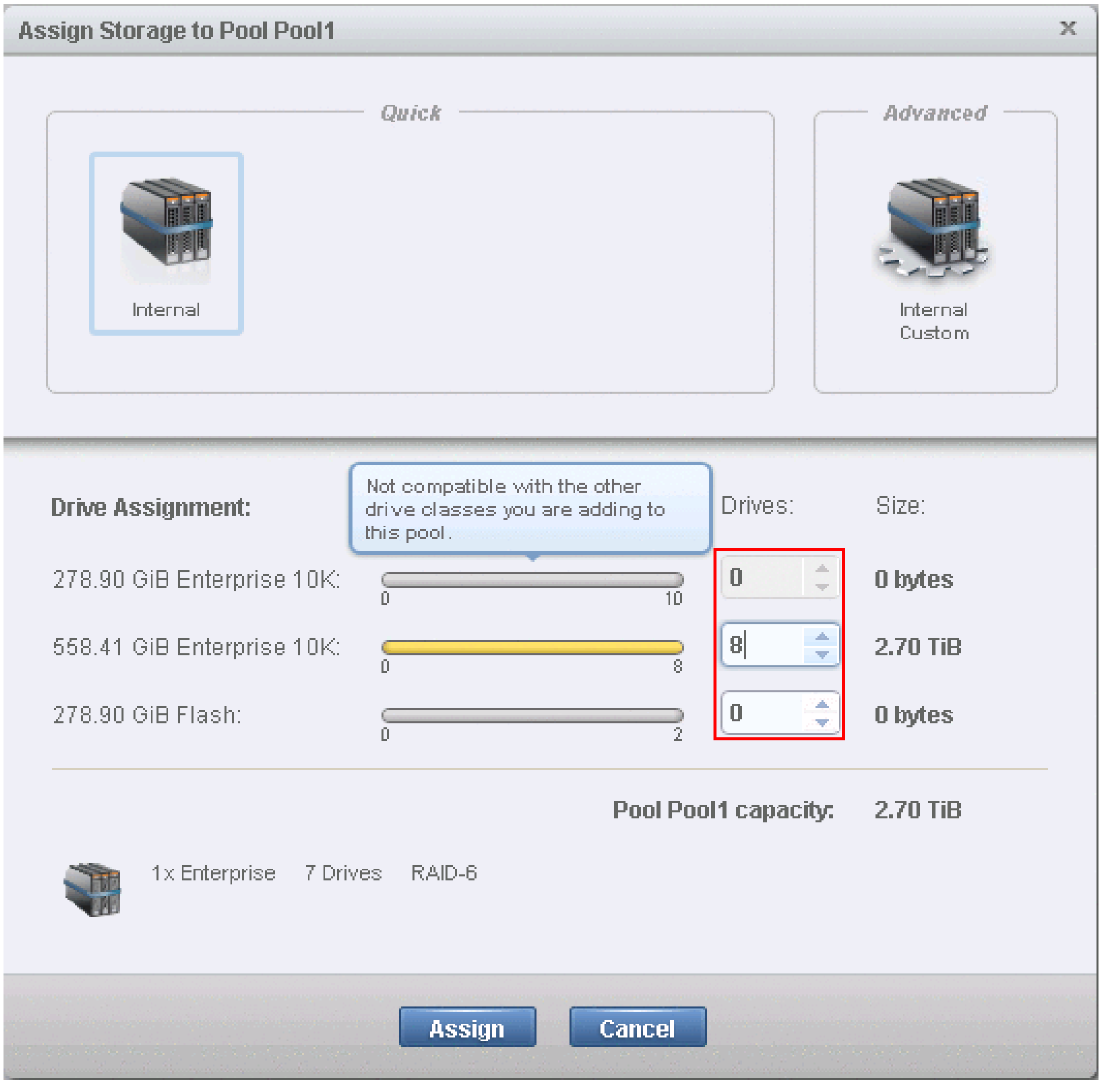
Pool Pool1 capacity: 2.16 TiB

| | | |
|---|---|---|
| 1x Enterprise | 9 Drives | RAID-6 |
| 1x Flash | 2 Drives | RAID-10 |

*Figure 6-24 Quick internal configuration: pool with multiple drive classes*

This configuration combines two drive classes, belonging to two different tiers of storage (Flash and Enterprise). This is the default option and takes advantage of the Easy Tier functionality. However, this can be adjusted by setting the number of drives of different classes to zero, as shown in Figure 6-25 on page 180. For information about Easy Tier see Chapter 10, “Advanced features for storage efficiency” on page 363.

Note that if any drive class is not compatible with the drives being assigned that drive class cannot be selected.



Drive Assignment:

Not compatible with the other drive classes you are adding to this pool.

| Drive Assignment: | Drives: | Size: |
|---|---|---|
| 278.90 GiB Enterprise 10K: | 0 | 0 bytes |
| 558.41 GiB Enterprise 10K: | 8 | 2.70 TiB |
| 278.90 GiB Flash: | 0 | 0 bytes |

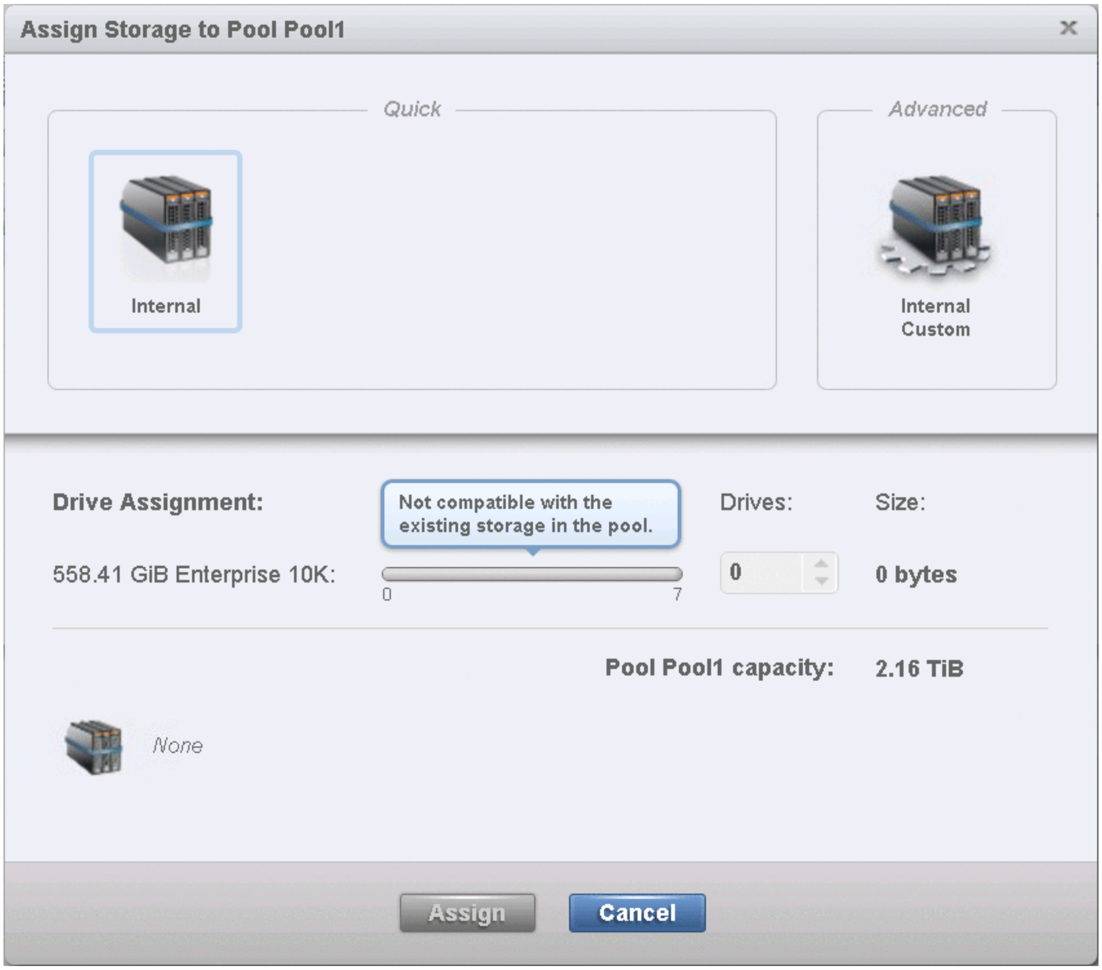
Pool Pool1 capacity: 2.70 TiB

1x Enterprise 7 Drives RAID-6

*Figure 6-25 Quick internal configuration: pool with single drive classes*

If you are adding storage to a pool with storage already assigned, the existing storage is also taken into consideration, with some properties being inherited from existing arrays for a given drive class. Drive classes incompatible with the classes already in the pool are disabled as well.

Figure 6-26 on page 181 shows an example of a quick configuration after the arrays suggested in Figure 6-24 are created. As expected, the drive class that was not compatible then is not compatible now.



Assign Storage to Pool Pool1

Drive Assignment:

Not compatible with the existing storage in the pool.

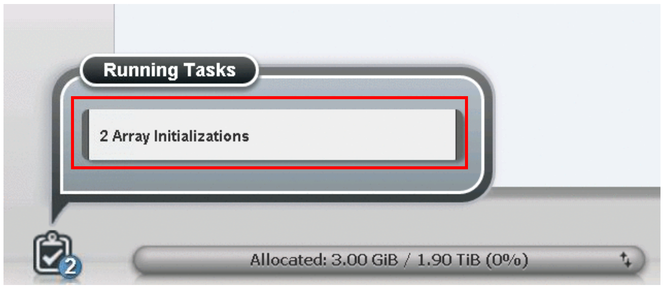
| Drive Assignment: | | Drives: | Size: |
|---|---|---|---|
| 558.41 GiB Enterprise 10K: | 0 — 7 | 0 | 0 bytes |

Pool Pool1 capacity: 2.16 TiB

None

Figure 6-26 Quick internal configuration: adding more capacity

When you are satisfied with the configuration presented click **Assign**. The array MDisks are then created and start initializing on the background. The progress of the initialization process can be monitored by selecting the correct task under Running Tasks on the lower-left corner, as shown in Figure 6-27. The array is available for I/O during this process.



2 Array Initializations

Allocated: 3.00 GiB / 1.90 TiB (0%)

Figure 6-27 Array initialization running task

## Advanced internal configuration

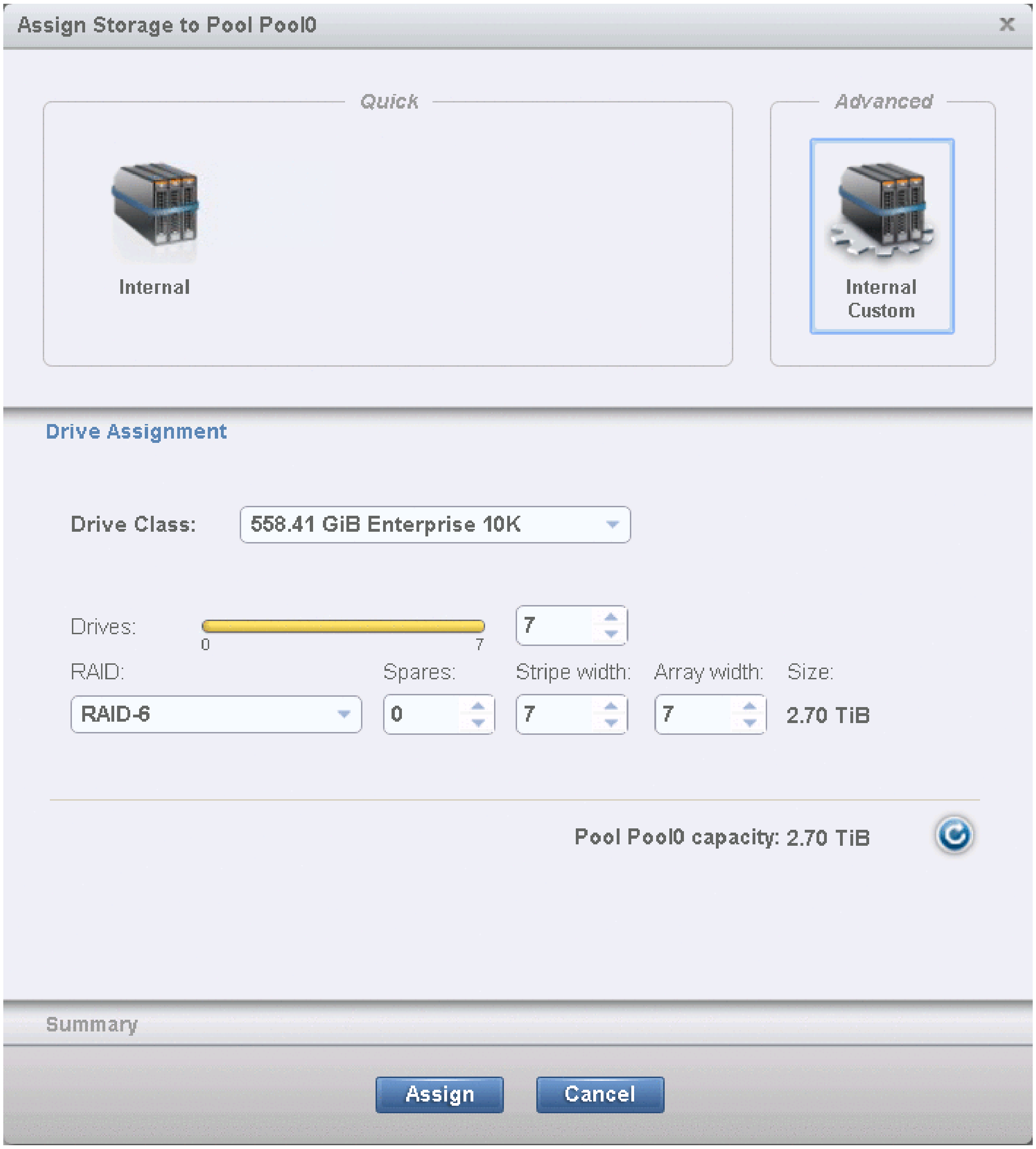
Selecting **Internal Custom** allows the user to customize the configuration for internal drives.

> **Tip:** It is advised to use the advanced configuration only when the quick configuration suggested does not fit your business requirements.

The following values can be customized:

- RAID level
- Number of spares
- Array width
- Stripe width
- Number of drives of each class

Figure 6-28 on page 182 shows an example with 7 drives ready to be configured as RAID 6. Click **Summary** to see the list of MDisks arrays to create.



Drive Class: 558.41 GiB Enterprise 10K

Drives: 7

RAID: RAID-6

Spares: 0

Stripe width: 7

Array width: 7

Size: 2.70 TiB

Pool Pool0 capacity: 2.70 TiB

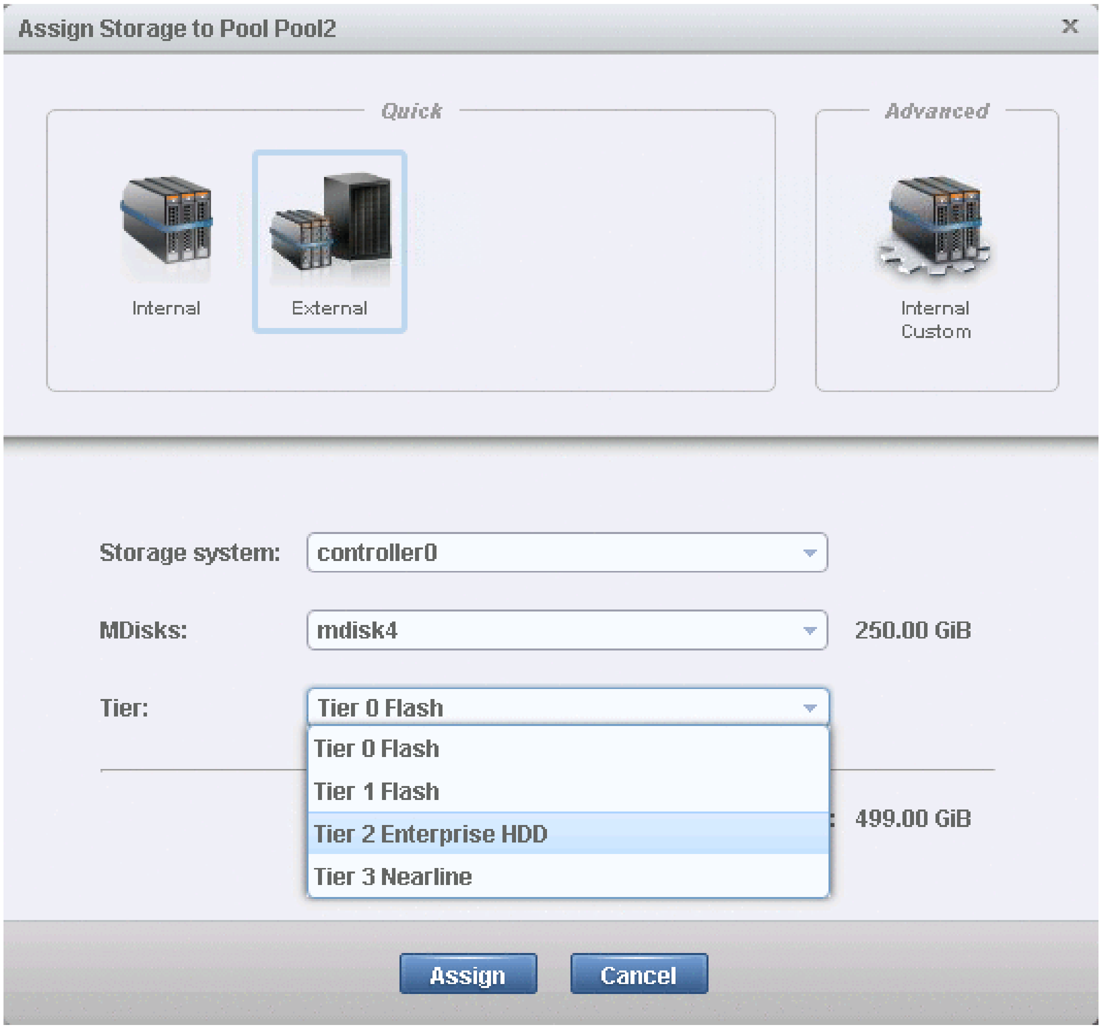
Figure 6-28 Custom internal configuration

To return to the default settings, select the refresh button next to the pool capacity. To create and assign the arrays, click **Assign**.

### Quick external configuration

Selecting **External** allows the user to assign external MDisks to storage pools. Contrarily to array MDisks, the tier associated with an external MDisk cannot be determined automatically and needs to be set manually. Select the external storage controller, the MDisks you want to assign, and the tier to which they belong, as shown in Figure 6-29 on page 183. The tier of an external MDisk can be modified after creation.

> **Attention:** If you need to preserve existing data on an unmanaged MDisk do *not* assign it to a storage pool because this action *deletes the data* on the MDisk. Use **Import** instead. See “Import” on page 190 for information about this action.



Storage system: controller0

MDisks: mdisk4 250.00 GiB

Tier: Tier 0 Flash

*Figure 6-29 Quick external configuration*

## 6.2.2 Traditional and distributed RAID

Release 7.6 of IBM Spectrum Virtualize introduced a new way of managing physical drives, as an alternative to the traditional Redundant Array of Independent Disks (RAID). It is called distributed RAID.

Storwize V7000 supports the following traditional and distributed RAID levels:

- RAID0

- RAID1
- RAID5
- RAID6
- RAID10
- Distributed RAID5 (DRAID5)
- Distributed RAID6 (DRAID6)

### Traditional RAID

In a traditional RAID approach, whether it is RAID10, RAID5, or RAID6, data is spread among drives in an array. However, the spare space is constituted by spare drives, which are global and sit outside of the array. When one of the drives within the array fails, all data is read from the mirrored copy (for RAID10), or is calculated from remaining data stripes and parity (for RAID5 or RAID6), and written to one single spare drive.

Figure 6-30 shows a traditional RAID6 array with two global spare drives and data and parity striped among five drives.

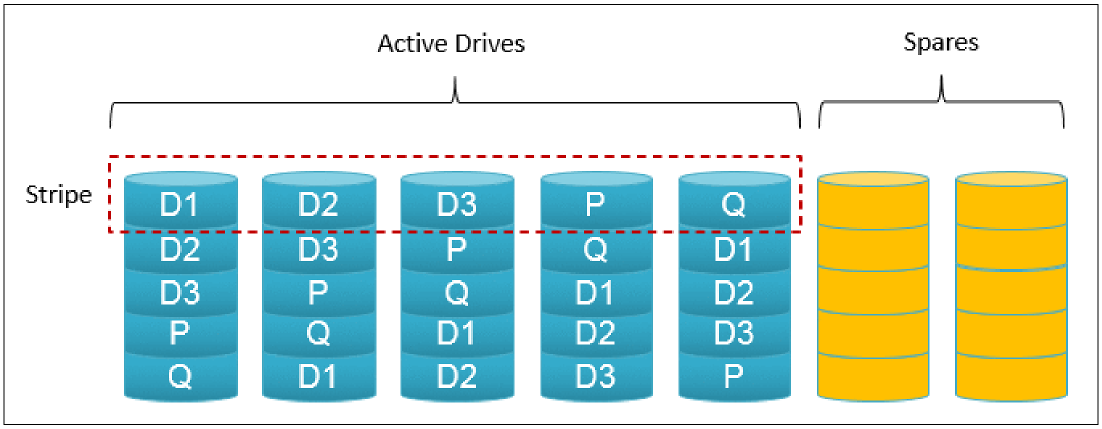


*Figure 6-30 Traditional RAID6 with two global spare drives*

If a drive fails, data is calculated from the remaining strips in a stripe and written to the spare strip in the same stripe on a spare drive, as shown in Figure 6-31.

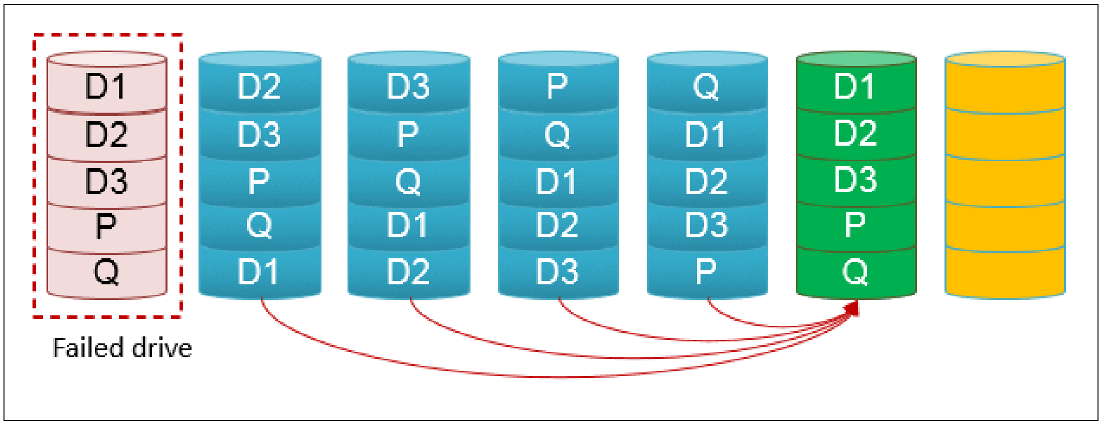


*Figure 6-31 Traditional RAID6 after a drive failure*

This model has two main disadvantages:

- In case of a drive failure data is read from many drives but written to only one. This can affect performance of foreground operations and means the rebuild process can take a long time, depending on the size of the drives.
- The spare drives are idle and do not perform any operations until one of the array drives fails.

## Distributed RAID

In distributed RAID, the spare drive (or drives) are included in the array as spare space. All drives in the array have spare space reserved that is used if one of the drives among the same array fails. This means there are no idling spare drives and all drives in the array take part in data processing. In case of a drive failure, the data is written to several drives and this reduces recovery time and consequently the probability of a second drive failure occurring during rebuild.

Distributed RAID also has the ability to distribute data and parity strips among more drives than traditional RAID. This means more drives can be used to create one array, therefore improving performance of a single managed disk.

Figure 6-32 shows a distributed RAID6 array with the stripe width of five distributed among 10 physical drives. The reserved spare space is marked in yellow and is equivalent to two spare drives.

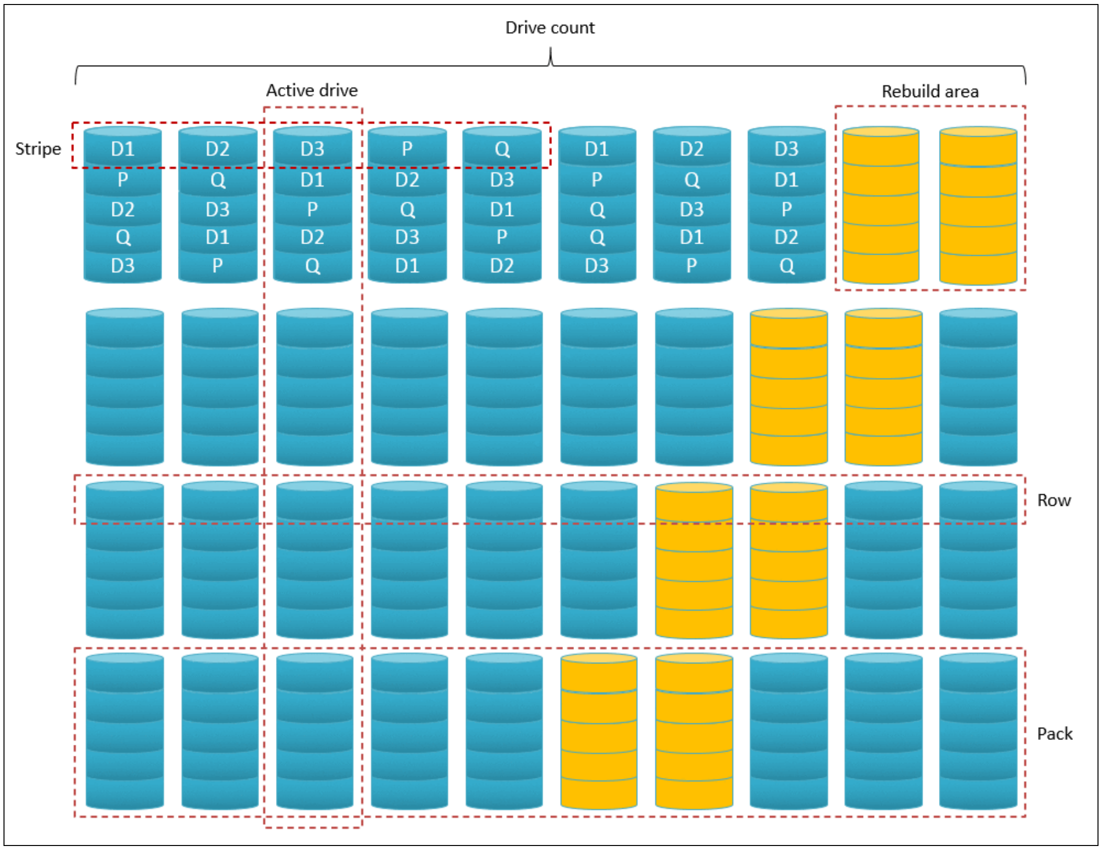


*Figure 6-32 Distributed RAID6*

Both distributed RAID5 and distributed RAID6 divide the physical drives into rows and packs. The row has the size of the array width and has only one strip from each drive in an array. A pack is a group of several continuous rows and its size depends on the number of strips in a stripe.

In case of a drive failure, all data is calculated using the remaining data stripes and parities and written to a spare space within each row, as shown in Figure 6-33.

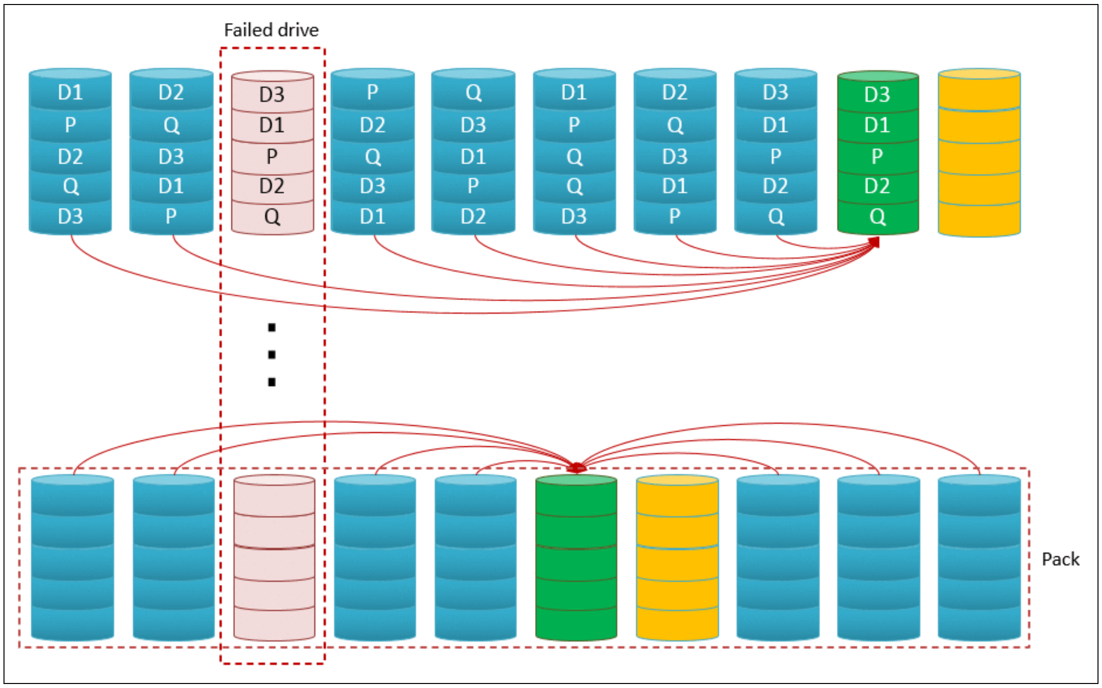


*Figure 6-33 Distributed RAID6 after a drive failure*

This model addresses the two main disadvantages of traditional RAID:

- In case of a drive failure data is read from many drives and written to many drives. This minimizes the effect on performance during the rebuild process and significantly reduces rebuild time (depending on the distributed array configuration and drive sizes the rebuild process can be up to 10 times faster).
- Spare space is distributed throughout the array, so there are more drives processing I/O and no idle spare drives.

There are also the following additional advantages:

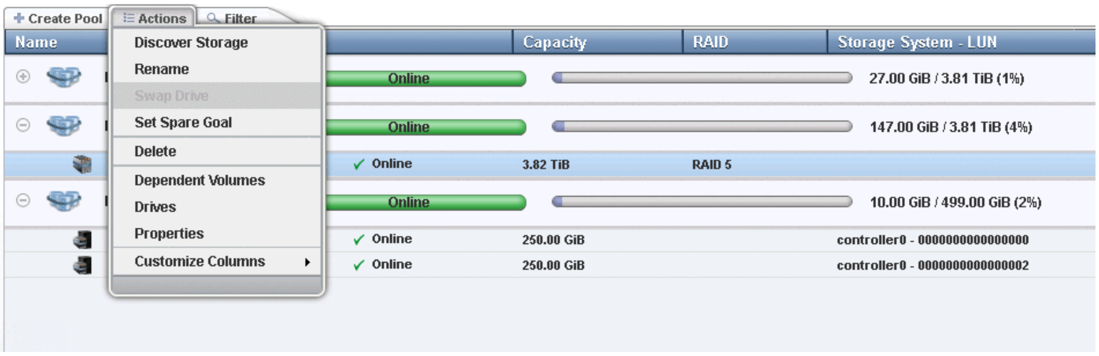
- In case of a drive failure, only the actual data is rebuilt. Space that is not allocated to volumes is not re-created to the spare regions of the array.
- Arrays can be much larger than before, spanning over many more drives and therefore improving the performance of the array.

**Note:** Distributed RAID does not change the number of failed drives an array can endure. Just like in traditional RAID, a distributed RAID5 array can only lose one physical drive and survive. If another drive fails in the same array before the array finishes rebuilding both the managed disk and storage pool go offline.

## 6.2.3 Actions on arrays

MDisks created from internal storage are RAID arrays and support specific actions that are not supported on external MDisks. Some actions supported on traditional RAID arrays are not supported on distributed RAID arrays and vice versa.

To choose an action select the array and click **Actions**, as shown in Figure 6-34. Alternatively, right-click the array.



| Name | | Capacity | RAID | Storage System - LUN |
|---|---|---|---|---|
| | Online | | | 27.00 GiB / 3.81 TiB (1%) |
| | Online | | | 147.00 GiB / 3.81 TiB (4%) |
| | ✓ Online | 3.82 TiB | RAID 5 | |
| | Online | | | 10.00 GiB / 499.00 GiB (2%) |
| | ✓ Online | 250.00 GiB | | controller0 - 0000000000000000 |
| | ✓ Online | 250.00 GiB | | controller0 - 0000000000000002 |

*Figure 6-34 Actions on arrays*

## Set spare goal

This action is available only for traditional RAID arrays. Selecting **Set Spare Goal** allows you to set the number of spare drives required to protect the array from drive failures. If the number of spare drives available does not meet the configured goal an error is logged in the event log. This error can be fixed by adding more drives of a compatible drive class as spares.

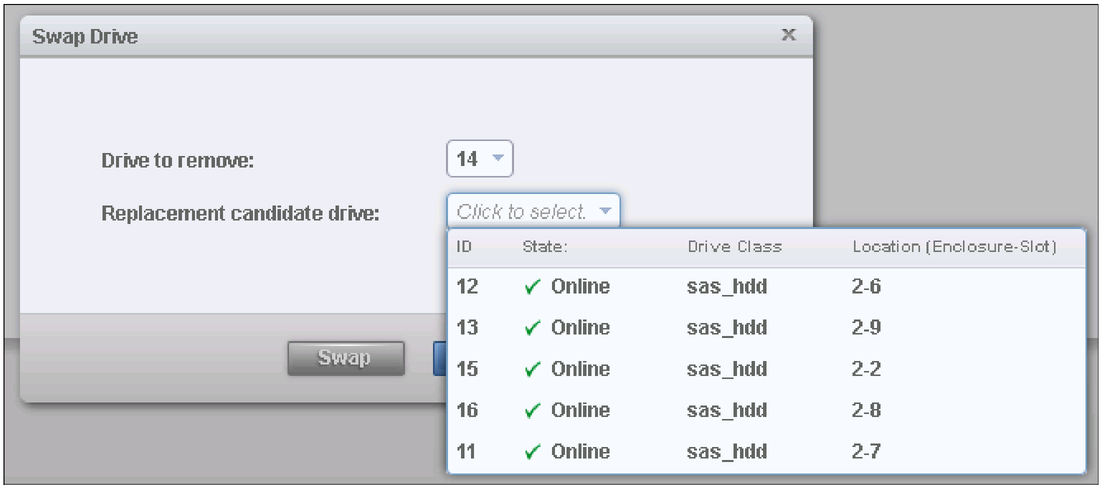
## Set rebuild areas goal

This action is available only for distributed RAID arrays. Selecting **Set Rebuild Areas Goal** allows you to set the number of rebuild areas required to protect the array from drive failures. If the number of rebuild areas available does not meet the configured goal, an error is logged in the event log. This error can be fixed by replacing the failed drives in the array with new drives of a compatible drive class.

## Swap drive

Selecting **Swap Drive** allows the user to replace a drive in the array with another drive. The other drive needs to have a use of `Candidate` or `Spare`. This action can be used to replace a drive that is expected to fail soon, for example, as indicated by an error message in the event log.

Figure 6-35 on page 187 shows the dialog box that opens. Select the member drive to replaced and the replacement drive and click **Swap**.



Drive to remove: 14

Replacement candidate drive: Click to select.

| ID | State: | Drive Class | Location (Enclosure-Slot) |
|---|---|---|---|
| 12 | ✓ Online | sas_hdd | 2-6 |
| 13 | ✓ Online | sas_hdd | 2-9 |
| 15 | ✓ Online | sas_hdd | 2-2 |
| 16 | ✓ Online | sas_hdd | 2-8 |
| 11 | ✓ Online | sas_hdd | 2-7 |

*Figure 6-35 Swapping two drives*

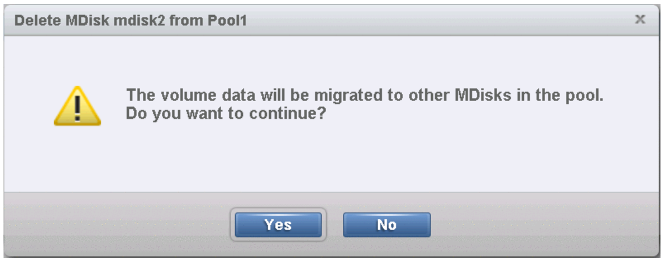
The exchange of the drives starts running in the background. The volumes on the affected MDisk remain accessible during the process.

### Delete

Selecting **Delete** removes the array from the storage pool and deletes it.

> **Remember:** An array does not exist outside of a storage pool. Therefore an array cannot be removed from the pool without being deleted.

If there are no volumes using extents from this array the command runs immediately without additional confirmation. If there are volumes using extents from this array, you are prompted to confirm the action, as shown in Figure 6-36.



Delete MDisk mdisk2 from Pool1

The volume data will be migrated to other MDisks in the pool.
Do you want to continue?

*Figure 6-36 Removing an array*

Confirming the action starts the migration of the volumes to extents from other MDisks that remain in the pool; after the action completes the array is removed from the storage pool and deleted.

> **Note:** Ensure that you have enough available capacity remaining in the storage pool to allocate the data being migrated from the removed array or else the command fails.

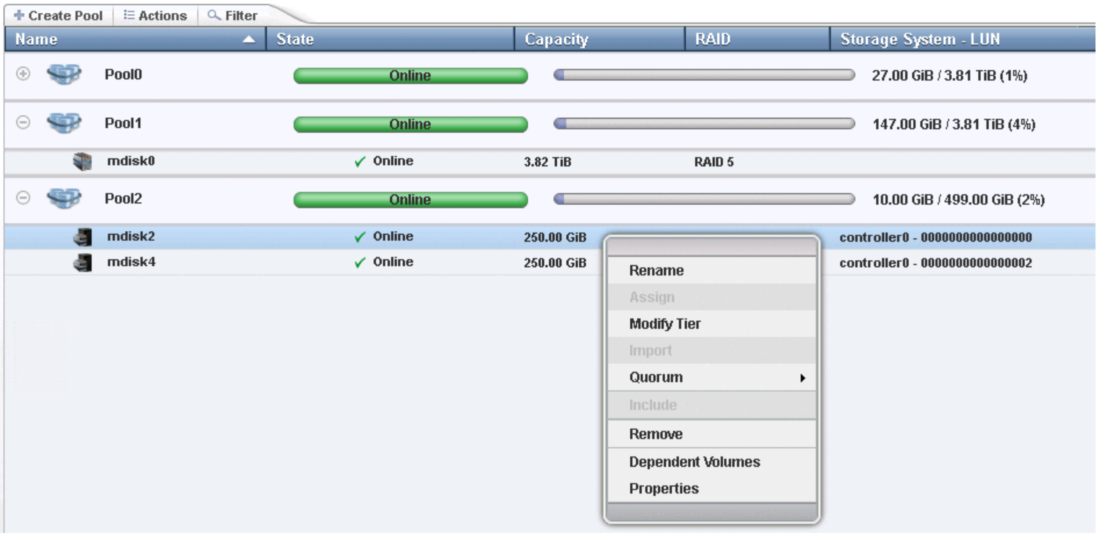
### Drives

Selecting **Drives** shows information about the drives that are included in the array.

## 6.2.4 Actions on external MDisks

External MDisks support specific actions that are not supported on arrays. Some actions are supported only on unmanaged external MDisks and some are supported only on managed external MDisks.

To choose an action right-click the external MDisk, as shown in Figure 6-37 on page 189. Alternatively, select the external MDisk and click **Actions**.



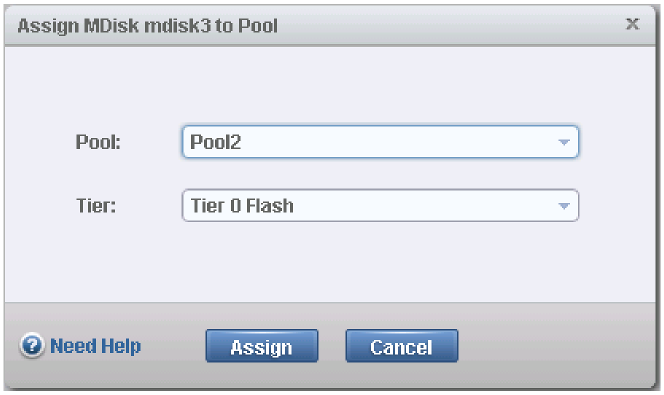
| Name | State | Capacity | RAID | Storage System - LUN |
|---|---|---|---|---|
| Pool0 | Online | 27.00 GiB / 3.81 TiB (1%) | | |
| Pool1 | Online | 147.00 GiB / 3.81 TiB (4%) | | |
| mdisk0 | ✓ Online | 3.82 TiB | RAID 5 | |
| Pool2 | Online | 10.00 GiB / 499.00 GiB (2%) | | |
| mdisk2 | ✓ Online | 250.00 GiB | | controller0 - 0000000000000000 |
| mdisk4 | ✓ Online | 250.00 GiB | | controller0 - 0000000000000002 |

*Figure 6-37 Actions on external MDisks*

## Assign

This action is available only for unmanaged MDisks. Selecting **Assign** opens the dialog box shown in Figure 6-38. This action is equivalent to the wizard described in Quick external configuration, but acts only on the selected MDisk or MDisks.

> **Attention:** If you need to preserve existing data on an unmanaged MDisk do *not* assign it to a storage pool because this action *deletes the data* on the MDisk. Use **Import** instead.



Assign MDisk mdisk3 to Pool

Pool: Pool2

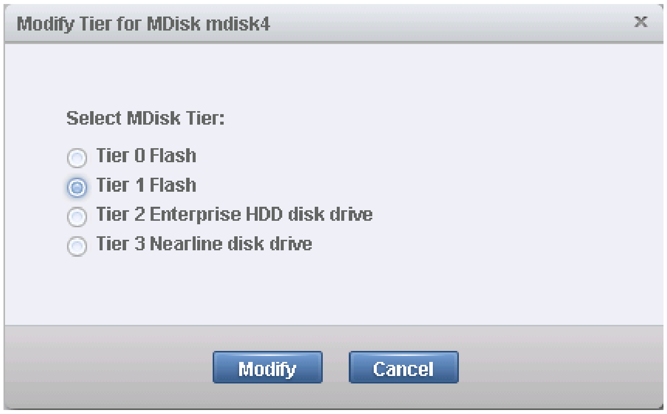
Tier: Tier 0 Flash

*Figure 6-38 Assigning an external MDisk to a storage pool*

## Modify tier

Selecting **Modify Tier** allows the user to modify the tier to which the external MDisk is assigned, as shown in Figure 6-39 on page 190. This setting is adjustable because the system cannot detect the tiers associated with external storage automatically. **Tier 2 Enterprise HDD disk drive** is the option selected by default.

Figure 6-39 on page 190 shows the dialog box that opens when **Modify Tier** is selected.



Modify Tier for MDisk mdisk4

Select MDisk Tier:

○ Tier 0 Flash
● Tier 1 Flash
○ Tier 2 Enterprise HDD disk drive
○ Tier 3 Nearline disk drive

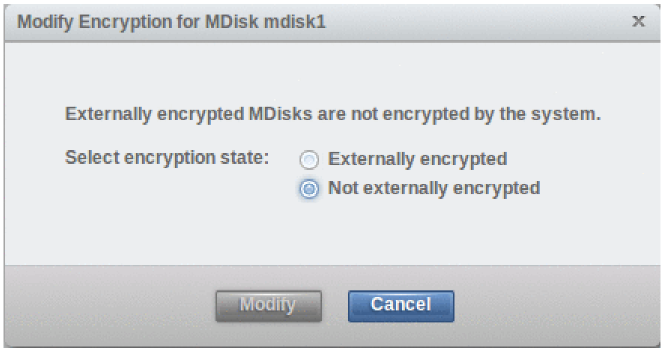
*Figure 6-39 Modifying the tier of an external MDisk*

For information about storage tiers and their importance see Chapter 10, “Advanced features for storage efficiency” on page 363.

## Modify encryption

This option is available only when encryption is enabled. Selecting **Modify Encryption** allows the user to modify the encryption setting for the MDisk, as shown in Figure 6-40.

For example, if the external MDisk is already encrypted by the external storage system, change the encryption state of the MDisk to **Externally encrypted**. This stops the system from encrypting the MDisk again if the MDisk is part of an encrypted storage pool.



Modify Encryption for MDisk mdisk1

Externally encrypted MDisks are not encrypted by the system.

Select encryption state: ○ Externally encrypted
● Not externally encrypted

*Figure 6-40 Modifying the encryption setting of an external MDisk*

For information about encryption, encrypted storage pools and self-encrypting MDisks see Chapter 12, “Encryption” on page 563.

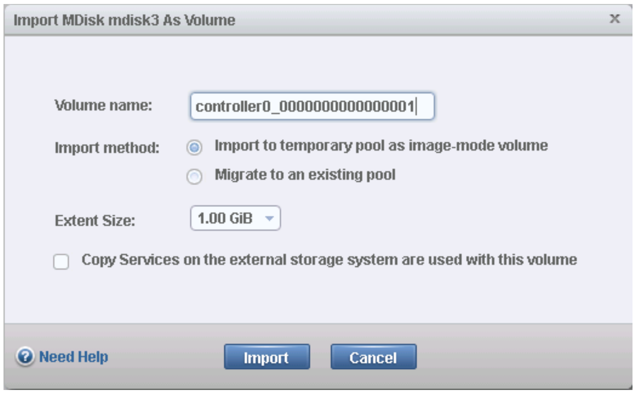
## Import

This action is available only for unmanaged MDisks. Importing an unmanaged MDisk allows the user to preserve the data on the MDisk, either by migrating the data to a new volume or by keeping the data on the external system.

Selecting **Import** allows the user to choose one of the migration methods described below:

- **Import to temporary pool as image-mode volume** does not migrate data from the source MDisk. It creates an *image-mode volume* that has a direct block-for-block translation of the MDisk. The existing data is preserved on the external storage system, but it is also accessible from the Storwize V7000 system.

If this method is selected the image-mode volume is created in a temporary migration pool and presented through the Storwize V7000. Choose the extent size of the temporary pool and click **Import**, as shown in Figure 6-41.



Import MDisk mdisk3 As Volume

Volume name: controller0_0000000000000001

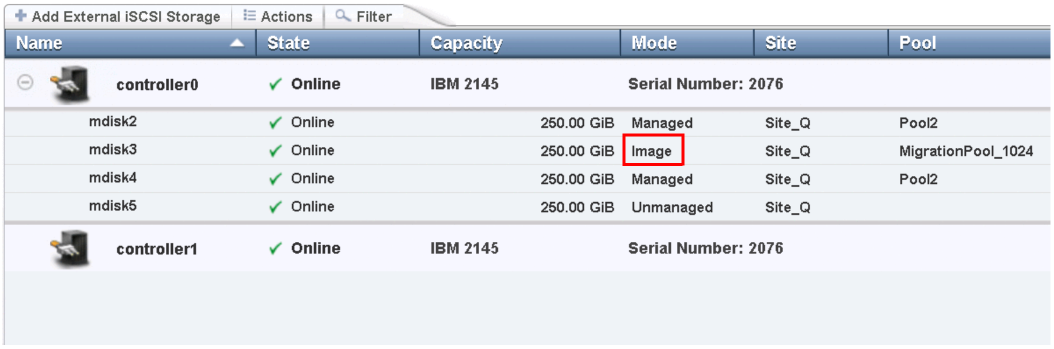
Import method: ● Import to temporary pool as image-mode volume
○ Migrate to an existing pool

Extent Size: 1.00 GiB

□ Copy Services on the external storage system are used with this volume

*Figure 6-41 Importing an external MDisk as an image-mode volume*

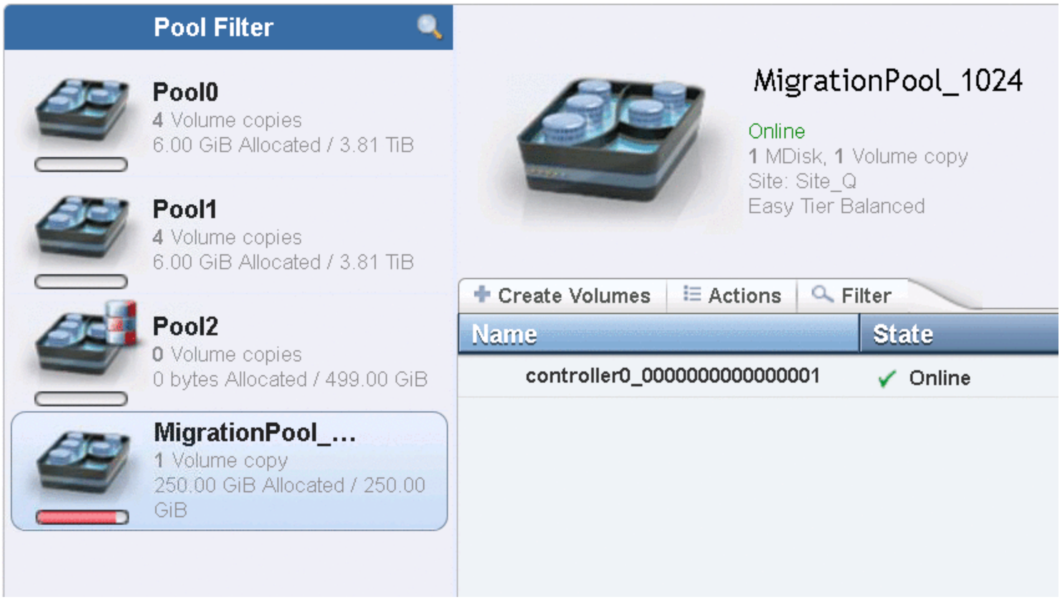
The MDisk is imported and listed as an `image` mode MDisk in the temporary migration pool, as shown in Figure 6-42.



| Name | State | Capacity | Mode | Site | Pool |
|---|---|---|---|---|---|
| controller0 | ✓ Online | IBM 2145 | Serial Number: 2076 | | |
| mdisk2 | ✓ Online | 250.00 GiB | Managed | Site_Q | Pool2 |
| mdisk3 | ✓ Online | 250.00 GiB | Image | Site_Q | MigrationPool_1024 |
| mdisk4 | ✓ Online | 250.00 GiB | Managed | Site_Q | Pool2 |
| mdisk5 | ✓ Online | 250.00 GiB | Unmanaged | Site_Q | |
| controller1 | ✓ Online | IBM 2145 | Serial Number: 2076 | | |

*Figure 6-42 Image mode MDisk*

A corresponding image-mode volume is now available in the same migration pool, as shown in Figure 6-43.



Pool0
4 Volume copies
6.00 GiB Allocated / 3.81 TiB

Pool1
4 Volume copies
6.00 GiB Allocated / 3.81 TiB

Pool2
0 Volume copies
0 bytes Allocated / 499.00 GiB

MigrationPool_...
1 Volume copy
250.00 GiB Allocated / 250.00 GiB

MigrationPool_1024
Online
1 MDisk, 1 Volume copy
Site: Site_Q
Easy Tier Balanced

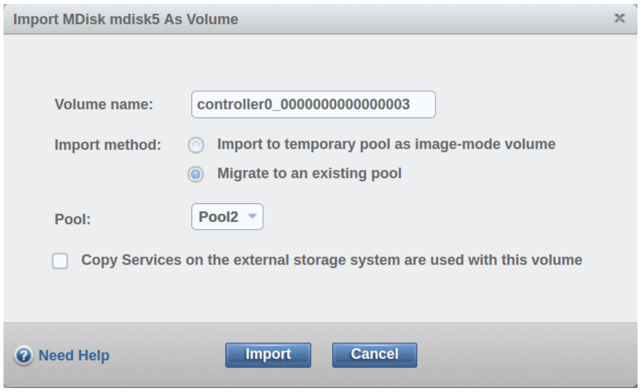
| Name | State |
|---|---|
| controller0_0000000000000001 | ✓ Online |

*Figure 6-43 Image-mode volume*

The image-mode volume can then be mapped to the original host mode. The data is still physically present on the physical disk of the original external storage controller system and no automatic migration process is currently running. If needed, the image-mode volume can be migrated manually to another storage pool using volume migration or volume mirroring later.

- **Migrate to an existing pool** starts by creating an image-mode volume as the first method. However, it then migrates the data from the image-mode volume onto another volume in the selected storage pool. After the migration process completes the image-mode volume and temporary migration pool are deleted.

  If this method is selected, choose the storage pool to hold the new volume and click **Import**, as shown in Figure 6-44 on page 192. Only pools with sufficient free extent capacity are listed.



Import MDisk mdisk5 As Volume

Volume name: controller0_0000000000000003

Import method: ○ Import to temporary pool as image-mode volume
● Migrate to an existing pool

Pool: Pool2

□ Copy Services on the external storage system are used with this volume

*Figure 6-44 Importing and migrating an external MDisk*

The data migration begins automatically after the MDisk is imported successfully as a image-mode volume. You can check the migration progress by clicking the task under Running Tasks, as shown in Figure 6-45. Alternatively, navigate to **Pools** → **System Migration**.

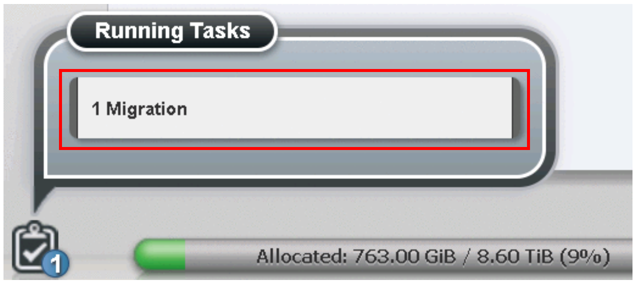
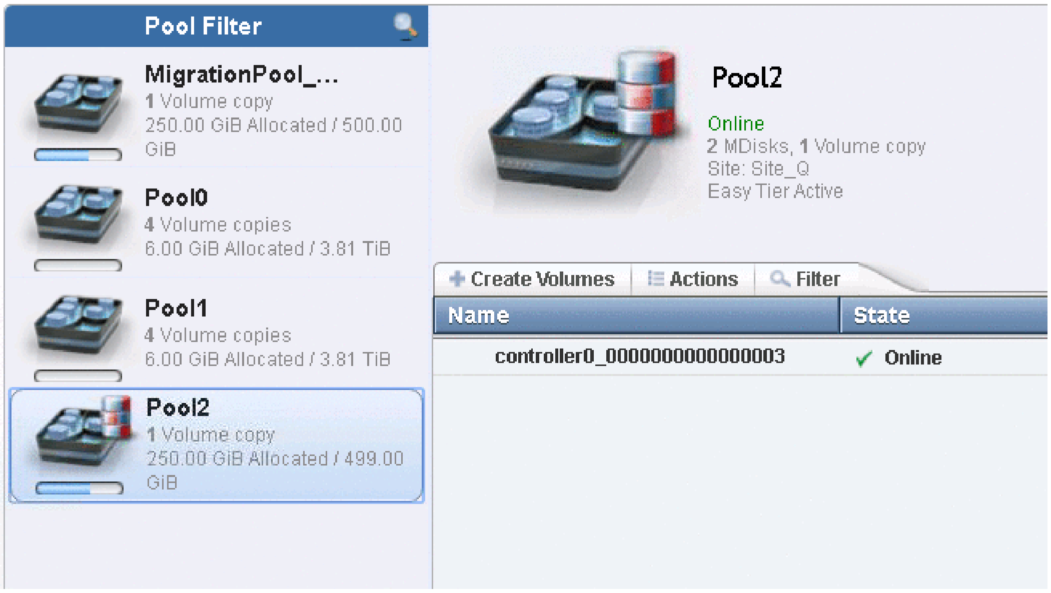


*Figure 6-45 Migration running task*

After the migration completes, the volume is available in the chosen destination pool, as shown in Figure 6-46 on page 193. This volume is no longer an image-mode volume; it is a normal striped volume.



MigrationPool_...
1 Volume copy
250.00 GiB Allocated / 500.00 GiB

Pool0
4 Volume copies
6.00 GiB Allocated / 3.81 TiB

Pool1
4 Volume copies
6.00 GiB Allocated / 3.81 TiB

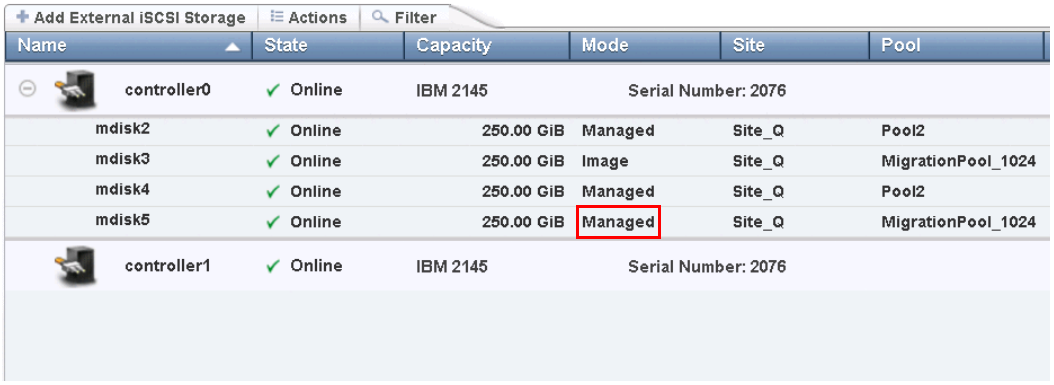
Pool2
1 Volume copy
250.00 GiB Allocated / 499.00 GiB

Pool2
Online
2 MDisks, 1 Volume copy
Site: Site_Q
Easy Tier Active

| Name | State |
|---|---|
| controller0_0000000000000003 | ✓ Online |

*Figure 6-46 Striped volume after migration*

At this point all data has been migrated off the source MDisk and the MDisk is no longer in image mode, as shown in Figure 6-47. The MDisk can be removed from the temporary pool and used as a regular MDisk to host volumes.



| Name | State | Capacity | Mode | Site | Pool |
|---|---|---|---|---|---|
| controller0 | ✓ Online | IBM 2145 | Serial Number: 2076 | | |
| mdisk2 | ✓ Online | 250.00 GiB | Managed | Site_Q | Pool2 |
| mdisk3 | ✓ Online | 250.00 GiB | Image | Site_Q | MigrationPool_1024 |
| mdisk4 | ✓ Online | 250.00 GiB | Managed | Site_Q | Pool2 |
| mdisk5 | ✓ Online | 250.00 GiB | Managed | Site_Q | MigrationPool_1024 |
| controller1 | ✓ Online | IBM 2145 | Serial Number: 2076 | | |

*Figure 6-47 External MDisk after migration*

Alternatively, import and migration of external MDisks to another pool can be done by selecting **Pools** → **System Migration**. Migration and the system migration wizard are described in more detail in Chapter 9, “Storage migration” on page 347.

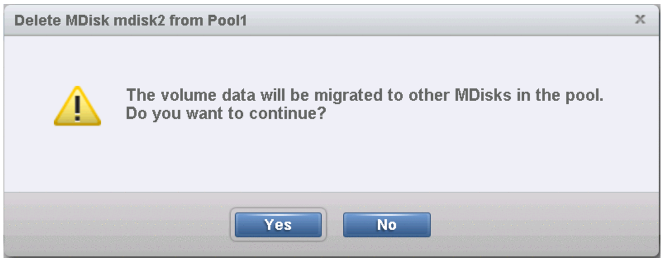
### Include

The system can exclude an MDisk with multiple I/O failures or persistent connection errors from its storage pool to ensure these errors do not interfere with data access. If an MDisk has been automatically excluded, run the fix procedures to resolve any connection and I/O failure errors. Drives used by the excluded MDisk with multiple errors might require replacing or reseating.

After the problems have been fixed, select **Include** to add the excluded MDisk back into the storage pool.

### Remove

In some cases you might want to remove external MDisks from storage pools to reorganize your storage allocation. Selecting **Remove** removes the MDisk from the storage pool. After the MDisk is removed it goes back to `unmanaged`. If there are no volumes in the storage pool to which this MDisk is allocated the command runs immediately without additional confirmation. If there are volumes in the pool, you are prompted to confirm the action, as shown in Figure 6-36.



Delete MDisk mdisk2 from Pool1

The volume data will be migrated to other MDisks in the pool.
Do you want to continue?

*Figure 6-48 Removing an external MDisk*

Confirming the action starts the migration of the volumes to extents from other MDisks that remain in the pool; when the action completes the MDisk is removed from the storage pool and returns to `unmanaged`.

> **Note:** Ensure that you have enough available capacity remaining in the storage pool to allocate the data being migrated from the removed MDisk or else the command fails.

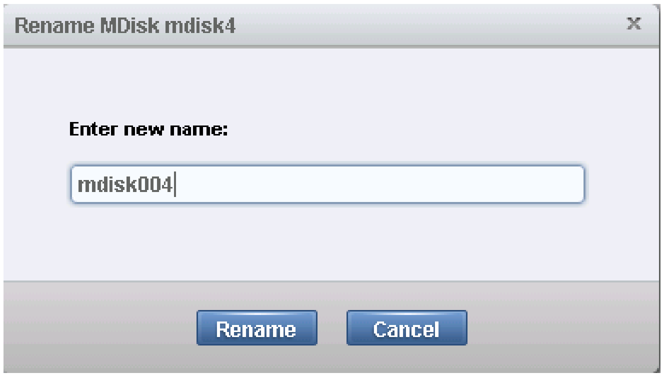
> **Important:** The MDisk being removed must remain accessible to the system while all data is copied to other MDisks in the same storage pool. If the MDisk is unmapped before the migration finishes all volumes in the storage pool go offline and remain in this state until the removed MDisk is connected again.

## 6.2.5 Additional actions on MDisks

There are a few additional actions supported both on arrays and external MDisks.

### Rename

Selecting **Rename** allows the user to modify the name of the MDisk, as shown in Figure 6-49. Enter the new name a click **Rename**.



Rename MDisk mdisk4

Enter new name: mdisk004

*Figure 6-49 Renaming an MDisk*

### Dependent volumes

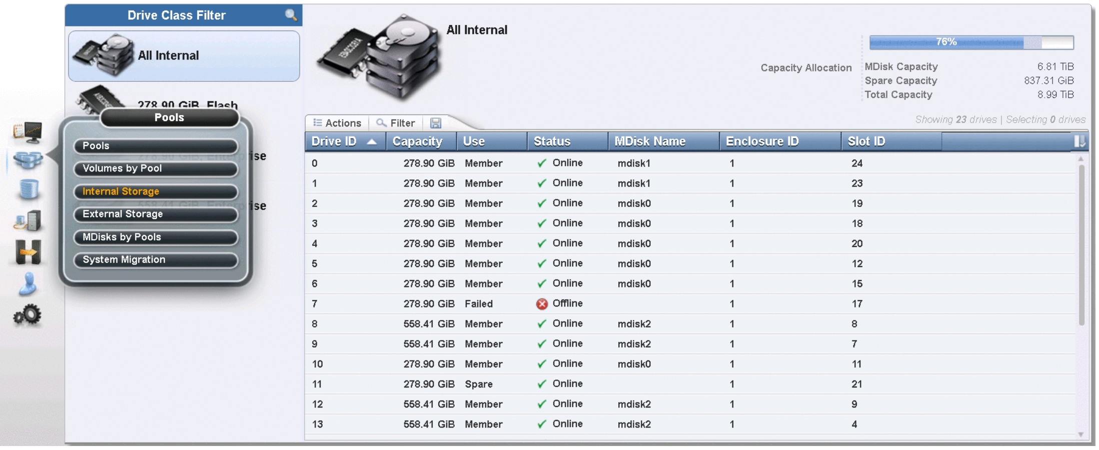
This action is not supported on unmanaged MDisks. Selecting **Dependent Volumes** lists the volumes that are dependent on the selected MDisk. A volume is dependent on an MDisk when the MDisk going offline causes the volume to go offline as well. Use this option before performing maintenance operations to determine which volumes are affected.

### Properties

Selecting **Properties** displays information about the MDisk. Additional information is available by clicking **View more details**.

## 6.3 Working with internal drives

The Storwize V7000 system provides an Internal Storage panel for managing all internal drives. To access the Internal Storage panel, browse to **Pools** → **Internal Storage**, as shown in Figure 6-50 on page 195.



Capacity Allocation

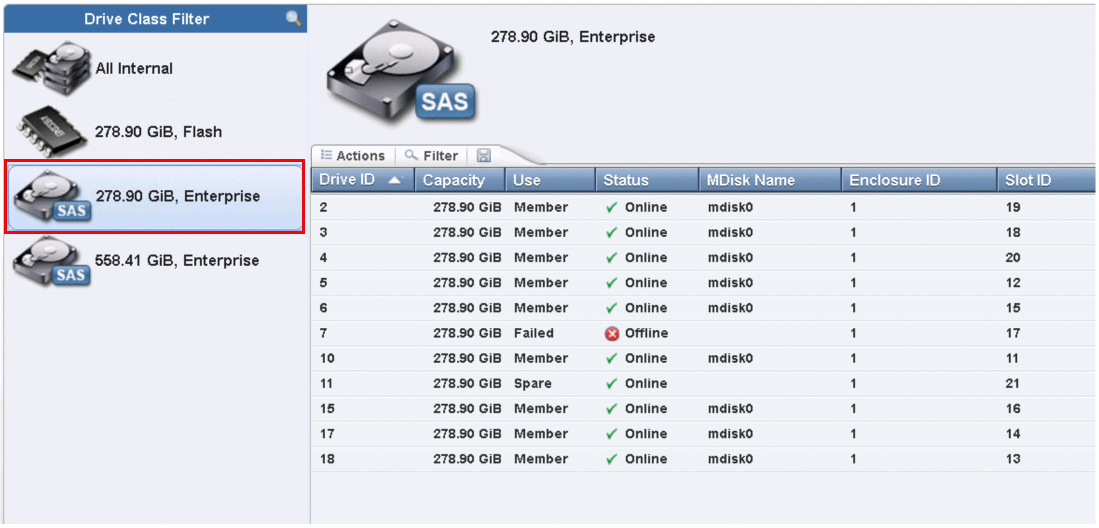
| | |
|---|---|
| MDisk Capacity | 6.81 TiB |
| Spare Capacity | 837.31 GiB |
| Total Capacity | 8.99 TiB |

Showing 23 drives | Selecting 0 drives

| Drive ID | Capacity | Use | Status | MDisk Name | Enclosure ID | Slot ID |
|---|---|---|---|---|---|---|
| 0 | 278.90 GiB | Member | ✓ Online | mdisk1 | 1 | 24 |
| 1 | 278.90 GiB | Member | ✓ Online | mdisk1 | 1 | 23 |
| 2 | 278.90 GiB | Member | ✓ Online | mdisk0 | 1 | 19 |
| 3 | 278.90 GiB | Member | ✓ Online | mdisk0 | 1 | 18 |
| 4 | 278.90 GiB | Member | ✓ Online | mdisk0 | 1 | 20 |
| 5 | 278.90 GiB | Member | ✓ Online | mdisk0 | 1 | 12 |
| 6 | 278.90 GiB | Member | ✓ Online | mdisk0 | 1 | 15 |
| 7 | 278.90 GiB | Failed | ☒ Offline | | 1 | 17 |
| 8 | 558.41 GiB | Member | ✓ Online | mdisk2 | 1 | 8 |
| 9 | 558.41 GiB | Member | ✓ Online | mdisk2 | 1 | 7 |
| 10 | 278.90 GiB | Member | ✓ Online | mdisk0 | 1 | 11 |
| 11 | 278.90 GiB | Spare | ✓ Online | | 1 | 21 |
| 12 | 558.41 GiB | Member | ✓ Online | mdisk2 | 1 | 9 |
| 13 | 558.41 GiB | Member | ✓ Online | mdisk2 | 1 | 4 |

*Figure 6-50 Accessing the internal storage panel*

The panel gives an overview of the internal drives in the Storwize V7000 system. Selecting **All Internal** in the drive class filter displays all drives managed in the system, including all I/O groups and expansion enclosures. Alternatively, you can filter the drives by their class, as shown in Figure 6-51 on page 196. The right side of the internal storage panel lists the internal disk drives of the class selected.

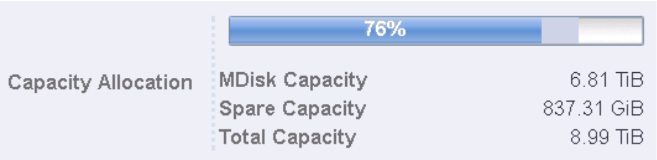


| Drive ID | Capacity | Use | Status | MDisk Name | Enclosure ID | Slot ID |
|---|---|---|---|---|---|---|
| 2 | 278.90 GiB | Member | ✓ Online | mdisk0 | 1 | 19 |
| 3 | 278.90 GiB | Member | ✓ Online | mdisk0 | 1 | 18 |
| 4 | 278.90 GiB | Member | ✓ Online | mdisk0 | 1 | 20 |
| 5 | 278.90 GiB | Member | ✓ Online | mdisk0 | 1 | 12 |
| 6 | 278.90 GiB | Member | ✓ Online | mdisk0 | 1 | 15 |
| 7 | 278.90 GiB | Failed | ✖ Offline | | 1 | 17 |
| 10 | 278.90 GiB | Member | ✓ Online | mdisk0 | 1 | 11 |
| 11 | 278.90 GiB | Spare | ✓ Online | | 1 | 21 |
| 15 | 278.90 GiB | Member | ✓ Online | mdisk0 | 1 | 16 |
| 17 | 278.90 GiB | Member | ✓ Online | mdisk0 | 1 | 14 |
| 18 | 278.90 GiB | Member | ✓ Online | mdisk0 | 1 | 13 |

*Figure 6-51 Filtering drives by class*

You can also find information regarding the capacity allocation of each drive class in the upper right corner, as shown in Figure 6-52. **Total Capacity** shows the overall capacity of the selected drive class; **MDisk Capacity** shows the storage capacity of the selected drive class that is assigned to MDisks; and **Spare Capacity** shows the storage capacity of the selected drive class that is used for spare drives. If **All Internal** is selected under the drive class filter, the values shown refer to the entire internal storage.

The percentage bar indicates how much of the total capacity is allocated to MDisks and spare drives, with MDisk capacity being represented by dark blue and spare capacity by light blue.



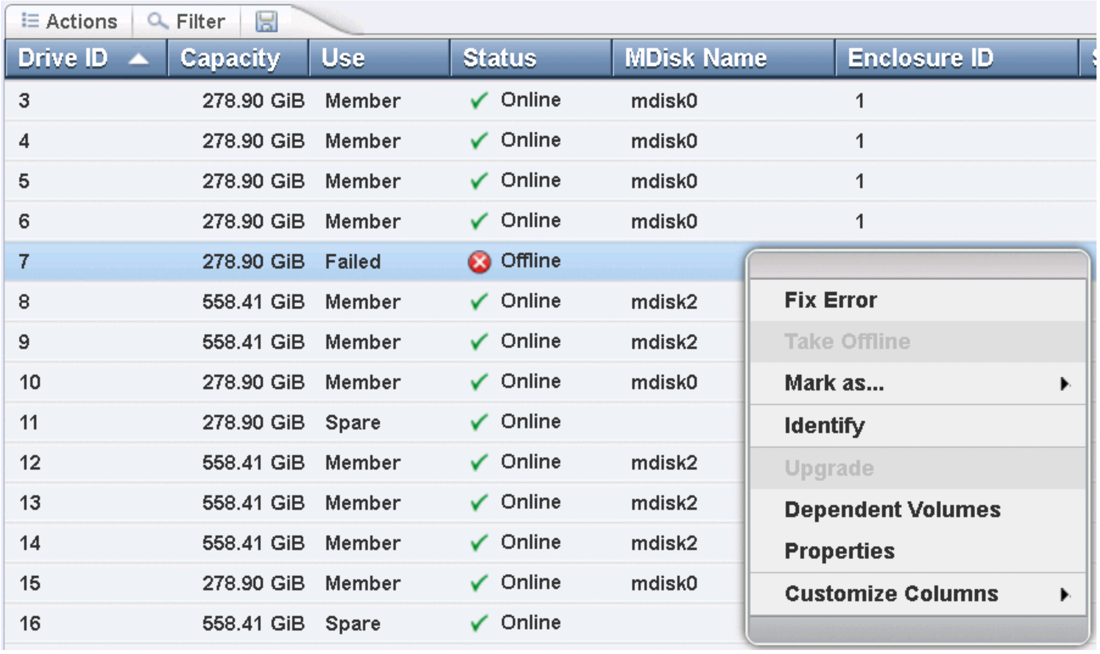
Capacity Allocation

| | |
|---|---|
| MDisk Capacity | 6.81 TiB |
| Spare Capacity | 837.31 GiB |
| Total Capacity | 8.99 TiB |

*Figure 6-52 Internal storage allocation indicator*

## 6.3.1 Actions on internal drives

There are a number of actions that can be performed on internal drives. To perform any action select the drives and right-click the selection, as shown in Figure 6-53. Alternatively, select the drives and click **Actions**.



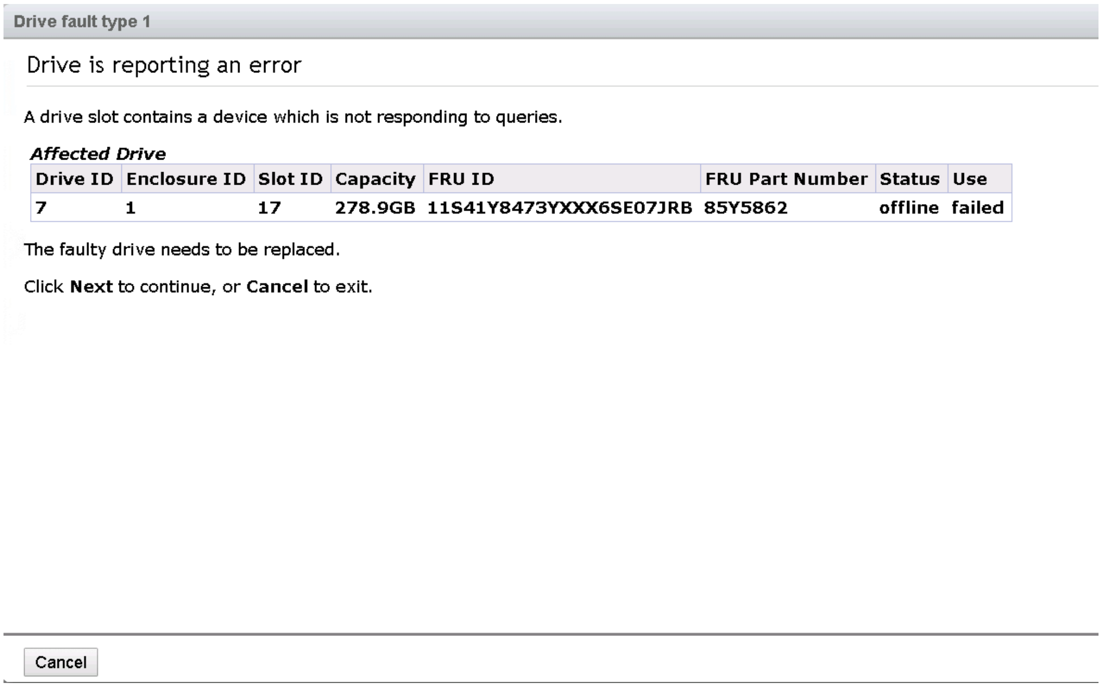
| Drive ID | Capacity | Use | Status | MDisk Name | Enclosure ID |
|---|---|---|---|---|---|
| 3 | 278.90 GiB | Member | ✓ Online | mdisk0 | 1 |
| 4 | 278.90 GiB | Member | ✓ Online | mdisk0 | 1 |
| 5 | 278.90 GiB | Member | ✓ Online | mdisk0 | 1 |
| 6 | 278.90 GiB | Member | ✓ Online | mdisk0 | 1 |
| 7 | 278.90 GiB | Failed | ✗ Offline | | |
| 8 | 558.41 GiB | Member | ✓ Online | mdisk2 | |
| 9 | 558.41 GiB | Member | ✓ Online | mdisk2 | |
| 10 | 278.90 GiB | Member | ✓ Online | mdisk0 | |
| 11 | 278.90 GiB | Spare | ✓ Online | | |
| 12 | 558.41 GiB | Member | ✓ Online | mdisk2 | |
| 13 | 558.41 GiB | Member | ✓ Online | mdisk2 | |
| 14 | 558.41 GiB | Member | ✓ Online | mdisk2 | |
| 15 | 278.90 GiB | Member | ✓ Online | mdisk0 | |
| 16 | 558.41 GiB | Spare | ✓ Online | | |

*Figure 6-53 Actions on internal drives*

The actions available depend on the status of the drive or drives selected. Some actions can only be run for individual drives.

## Fix error

This action is only available if the drive selected is in an error condition. Selecting **Fix Error** starts the Directed Maintenance Procedure (DMP) for the defective drive. For the drive listed as `offline` in Figure 6-53, for example, this action starts the DMP shown in Figure 6-54 on page 197. This DMP guides the user through the steps needed to replace the faulty drive.



Drive fault type 1

Drive is reporting an error

A drive slot contains a device which is not responding to queries.

***Affected Drive***

| Drive ID | Enclosure ID | Slot ID | Capacity | FRU ID | FRU Part Number | Status | Use |
|---|---|---|---|---|---|---|---|
| 7 | 1 | 17 | 278.9GB | 11S41Y8473YXXX6SE07JRB | 85Y5862 | offline | failed |

The faulty drive needs to be replaced.

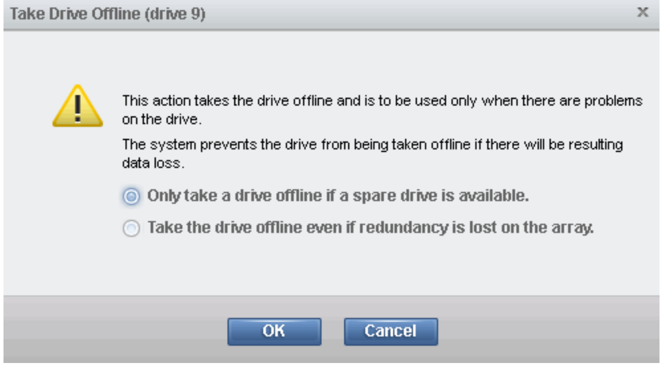
Click **Next** to continue, or **Cancel** to exit.

*Figure 6-54 DMP for drive fault type 1*

For more information about DMPs, see Chapter 13, “RAS, monitoring, and troubleshooting” on page 605.

### Take offline

Selecting **Take Offline** allows the user to take a drive offline. Select this action *only if* there is a problem with the drive and a spare drive is available. When selected you are prompted to confirm the action, as shown in Figure 6-55.



This action takes the drive offline and is to be used only when there are problems on the drive.

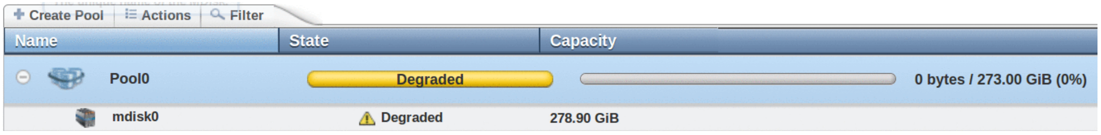
The system prevents the drive from being taken offline if there will be resulting data loss.

● Only take a drive offline if a spare drive is available.

○ Take the drive offline even if redundancy is lost on the array.

*Figure 6-55 Confirmation dialog box before taking an internal drive offline*

If a spare drive is available and the drive is taken offline, the MDisk of which the failed drive is a member remains `Online` and the spare is automatically reassigned. If no spare drive is available and the drive is taken offline, the MDisk of which the failed drive is a member gets `Degraded`; consequently, the storage pool to which the MDisk belongs gets `Degraded` as well, as shown in Figure 6-56 on page 198.

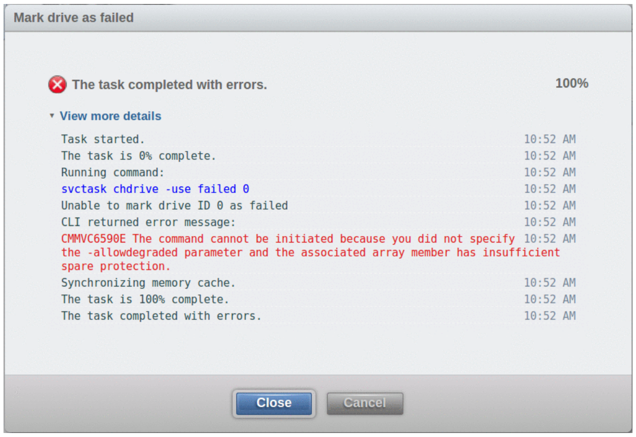


| Name | State | Capacity |
|---|---|---|
| Pool0 | Degraded | 0 bytes / 273.00 GiB (0%) |
| mdisk0 | ⚠ Degraded | 278.90 GiB |

*Figure 6-56 Degraded MDisk and storage pool*

The system prevents you from taking the drive offline if one of the following conditions is true:

- The first option was selected and no suitable spares are available. Figure 6-57 shows the error thrown in this case.

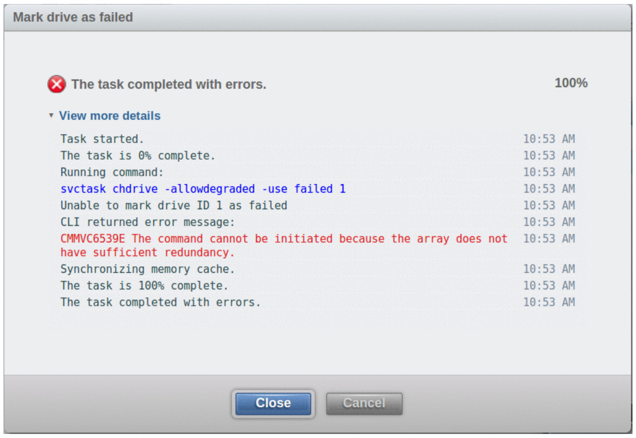


The task completed with errors. 100%

```
Task started.                                                          10:52 AM
The task is 0% complete.                                               10:52 AM
Running command:                                                       10:52 AM
svctask chdrive -use failed 0                                          10:52 AM
Unable to mark drive ID 0 as failed                                    10:52 AM
CLI returned error message:                                            10:52 AM
CMMVC6590E The command cannot be initiated because you did not specify 10:52 AM
the -allowdegraded parameter and the associated array member has insufficient
spare protection.
Synchronizing memory cache.                                            10:52 AM
The task is 100% complete.                                             10:52 AM
The task completed with errors.                                        10:52 AM
```

*Figure 6-57 Error thrown for insufficient spares*

- Losing another drive in the MDisk results in data loss. Figure 6-58 on page 199 shows the error thrown in this case.

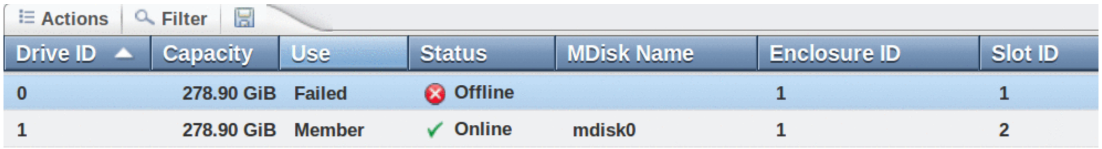


The task completed with errors. 100%

```
Task started.                                                          10:53 AM
The task is 0% complete.                                               10:53 AM
Running command:                                                       10:53 AM
svctask chdrive -allowdegraded -use failed 1                           10:53 AM
Unable to mark drive ID 1 as failed                                    10:53 AM
CLI returned error message:                                            10:53 AM
CMMVC6539E The command cannot be initiated because the array does not  10:53 AM
have sufficient redundancy.
Synchronizing memory cache.                                            10:53 AM
The task is 100% complete.                                             10:53 AM
The task completed with errors.                                        10:53 AM
```

*Figure 6-58 Error thrown for insufficient redundancy*

A drive that is taken offline is considered `Failed`, as shown in Figure 6-59.



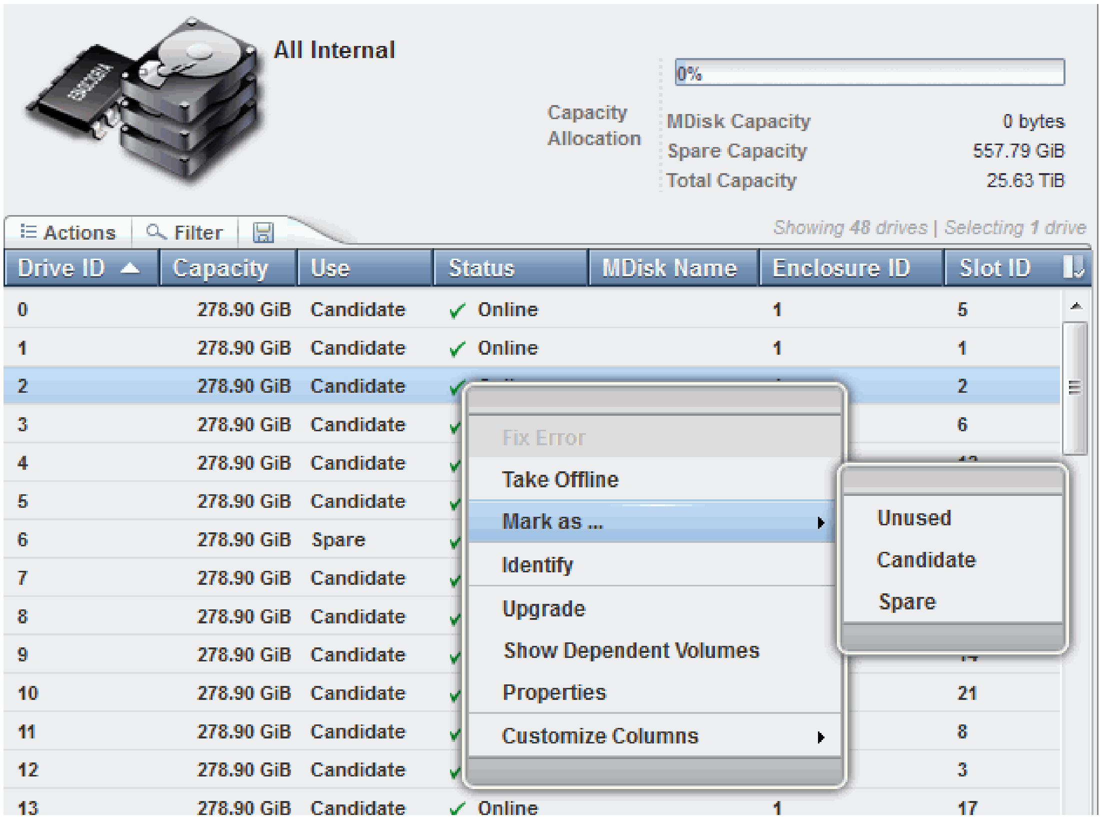
| Drive ID ▲ | Capacity | Use | Status | MDisk Name | Enclosure ID | Slot ID |
|---|---|---|---|---|---|---|
| 0 | 278.90 GiB | Failed | ✖ Offline | | 1 | 1 |
| 1 | 278.90 GiB | Member | ✓ Online | mdisk0 | 1 | 2 |

*Figure 6-59 Internal drive taken offline*

## Mark as

Selecting **Mark as** allows the user to change the usage assigned to the drive. The use options available are shown in Figure 6-60 and are described below:

- **Unused**: the drive is not in use and cannot be used as a spare.
- **Candidate**: the drive is available to be used in an MDisk.
- **Spare**: the drive can be used as a hot spare, if required.



All Internal

Capacity Allocation 0%

MDisk Capacity 0 bytes
Spare Capacity 557.79 GiB
Total Capacity 25.63 TiB

Showing 48 drives | Selecting 1 drive

| Drive ID | Capacity | Use | Status | MDisk Name | Enclosure ID | Slot ID |
|---|---|---|---|---|---|---|
| 0 | 278.90 GiB | Candidate | ✓ Online | | 1 | 5 |
| 1 | 278.90 GiB | Candidate | ✓ Online | | 1 | 1 |
| 2 | 278.90 GiB | Candidate | ✓ | | | 2 |
| 3 | 278.90 GiB | Candidate | ✓ | | | 6 |
| 4 | 278.90 GiB | Candidate | ✓ | | | |
| 5 | 278.90 GiB | Candidate | ✓ | | | |
| 6 | 278.90 GiB | Spare | ✓ | | | |
| 7 | 278.90 GiB | Candidate | ✓ | | | |
| 8 | 278.90 GiB | Candidate | ✓ | | | |
| 9 | 278.90 GiB | Candidate | ✓ | | | |
| 10 | 278.90 GiB | Candidate | ✓ | | | 21 |
| 11 | 278.90 GiB | Candidate | ✓ | | | 8 |
| 12 | 278.90 GiB | Candidate | ✓ | | | 3 |
| 13 | 278.90 GiB | Candidate | ✓ Online | | 1 | 17 |

*Figure 6-60 Changing the usage of internal drives*

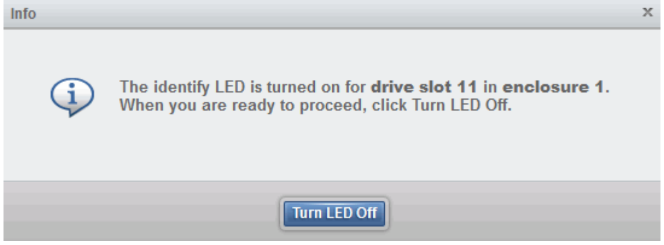
The use that can be assigned depends on the current drive use. These dependencies are shown in Table 6-1.

*Table 6-1 Allowed usage changes for internal drives*

| | | To | | | | |
|---|---|---|---|---|---|---|
| | | Unused | Candidate | Spare | Failed | Member |
| From | Unused | allowed | allowed | not allowed | no option | |
| | Candidate | allowed | allowed | allowed | no option | |
| | Spare | not allowed | allowed | allowed | no option | |
| | Failed | allowed | allowed | not allowed | no option | |
| | Member | no change allowed for member drives | | | | |

## Identify

Selecting **Identify** turns on the LED light so you can easily identify a drive that must be replaced or that you want to troubleshoot. Selecting this action opens a dialog box like the one shown in Figure 6-61 on page 201.


Info

The identify LED is turned on for **drive slot 11** in **enclosure 1**.
When you are ready to proceed, click Turn LED Off.

*Figure 6-61 Identifying an internal drive*

Click **Turn LED Off** when you are finished.

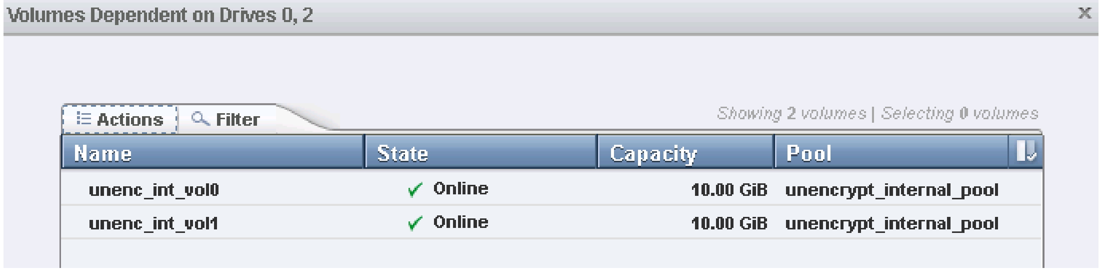
## Upgrade

Selecting **Upgrade** allows the user to update the drive firmware. You can choose to update individual drives or all the drives that have available updates. For information about updating drive firmware see Chapter 13, “RAS, monitoring, and troubleshooting” on page 605.

## Show dependent volumes

Selecting **Show Dependent Volumes** lists the volumes that are dependent on the selected drive. A volume is dependent on a drive when removal or failure of that drive causes the volume to become unavailable. Use this option before performing maintenance operations to determine which volumes are affected.

Figure 6-62 shows the list of volumes dependent on two drives that belong to the same MDisk. This means both volumes listed will go offline if *both* drives go offline.


Volumes Dependent on Drives 0, 2

Showing 2 volumes | Selecting 0 volumes

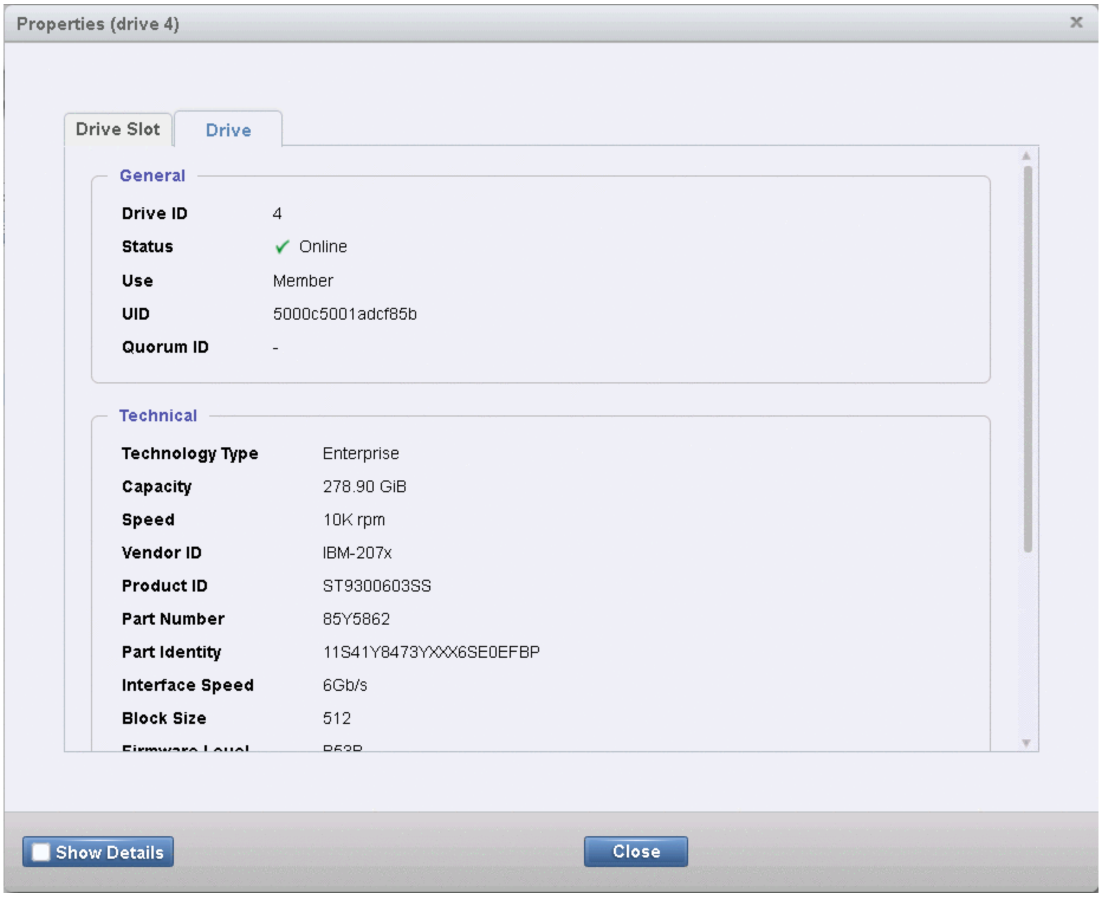
| Name | State | Capacity | Pool |
|---|---|---|---|
| unenc_int_vol0 | ✓ Online | 10.00 GiB | unencrypt_internal_pool |
| unenc_int_vol1 | ✓ Online | 10.00 GiB | unencrypt_internal_pool |

*Figure 6-62 Volumes dependent on internal drives*

> **Note:** A lack of dependent volumes does not imply that there are no volumes using the drive.

## Properties

Selecting **Properties** provides more information about the drive, as shown in Figure 6-63 on page 202.



General

| | |
|---|---|
| **Drive ID** | 4 |
| **Status** | ✓ Online |
| **Use** | Member |
| **UID** | 5000c5001adcf85b |
| **Quorum ID** | - |

Technical

| | |
|---|---|
| **Technology Type** | Enterprise |
| **Capacity** | 278.90 GiB |
| **Speed** | 10K rpm |
| **Vendor ID** | IBM-207x |
| **Product ID** | ST9300603SS |
| **Part Number** | 85Y5862 |
| **Part Identity** | 11S41Y8473YXXX6SE0EFBP |
| **Interface Speed** | 6Gb/s |
| **Block Size** | 512 |
| **Firmware Level** | B53B |

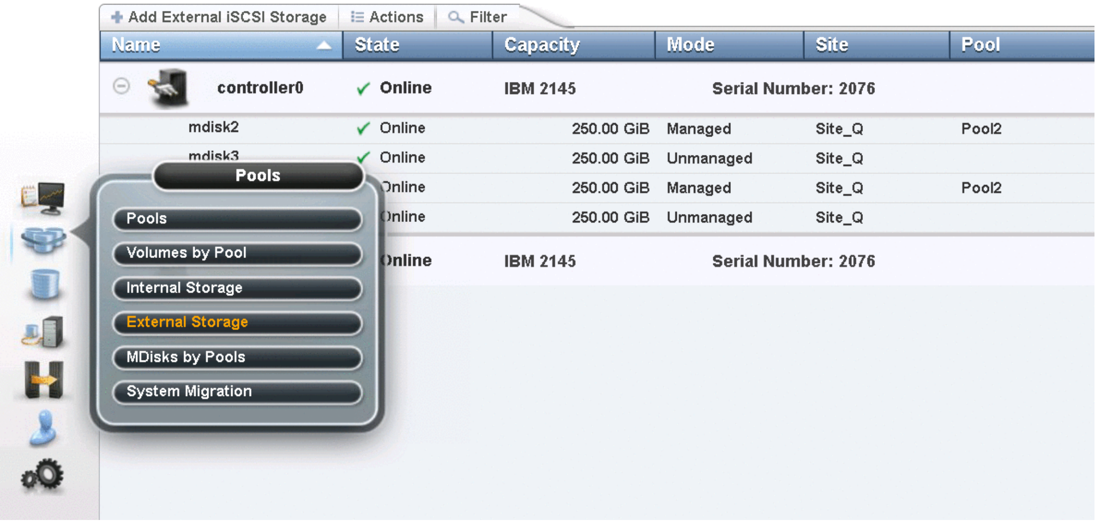
*Figure 6-63 Drive properties*

Checking **Show Details** on the left corner of the window shows more details, including vendor ID, product ID, and part number. You can also display drive slot details by selecting **Drive Slot**.

## 6.4 Working with external storage controllers

Release 7.7 of IBM Spectrum Virtualize introduced support for external storage controllers attached through iSCSI. Attachment through Fibre Channel is supported as before.

External storage controllers with both types of attachment can be managed through the External Storage panel. To access the External Storage panel, browse to **Pools** → **External Storage**, as shown in Figure 6-64 on page 203.



| Name | State | Capacity | Mode | Site | Pool |
|---|---|---|---|---|---|
| controller0 | ✓ Online | IBM 2145 | Serial Number: 2076 | | |
| mdisk2 | ✓ Online | 250.00 GiB | Managed | Site_Q | Pool2 |
| mdisk3 | ✓ Online | 250.00 GiB | Unmanaged | Site_Q | |
| | Online | 250.00 GiB | Managed | Site_Q | Pool2 |
| | Online | 250.00 GiB | Unmanaged | Site_Q | |
| | Online | IBM 2145 | Serial Number: 2076 | | |

*Figure 6-64 Accessing the external storage panel*

The panel lists the external controllers connected to the Storwize V7000 system and all the external MDisks detected by the system. The MDisks are organized by the external storage system that presents them. You can toggle the sign to the left of the controller icon to either show or hide the MDisks associated with the controller.

**Note:** A controller connected through Fibre Channel is detected automatically by the system, provided the cabling, the zoning, and the system layer are configured correctly. A controller connected through iSCSI must be added to the system manually.

If you have configured logical unit names on your external storage systems, it is not possible for the system to determine this name, because it is local to the external storage system. However, you can use the external storage system WWNNs and the LU number to identify each device.

## 6.4.1 Fibre Channel external storage controllers

A controller connected through Fibre Channel is detected automatically by the system, provided the cabling, the zoning, and the system layer are configured correctly.

If the external controller is not detected, ensure the Storwize V7000 is cabled and zoned into the same storage area network (SAN) as the external storage system. If you are using Fibre Channel, connect the Fibre Channel cables to the Fibre Channel ports of the canisters in your system, and then to the Fibre Channel network. If you are using Fibre Channel over Ethernet, connect Ethernet cables to the 10 Gbps Ethernet ports.

**Attention:** If the external controller is a Storwize system, the Storwize V7000 must be configured at the *replication* layer and the external controller must be configured at the *storage* layer. The default layer for a Storwize system is *storage*. Make sure the layers are correct before zoning the two systems together.

Ensure the layer of both systems is correct by entering the following command: **`svcinfo lssystem`**

If needed, change the layer of the Storwize V7000 to replication by entering the following command: **`chsystem -layer replication`**

If needed, change the layer of the Storwize controller to storage by entering the following command: **`chsystem -layer storage`**

For more information about layers and how to change them see Chapter 11, “Advanced Copy Services” on page 413.

### 6.4.2 iSCSI external storage controllers

Unlike Fibre Channel connections, you must manually configure iSCSI connections between the Storwize V7000 and the external storage controller. Until then the controller is not listed in the External Storage panel.

Before adding an iSCSI-attached controller, ensure the following prerequisites are fulfilled:

- The Storwize V7000 and the external storage system are connected through one or more Ethernet switches. Symmetric ports on all nodes of the Storwize V7000 are connected to the same switch and configured on the same subnet. Optionally you can use a Virtual Local Area Network (VLAN) to define network traffic for the system ports.

  Direct attachment between this system and the external controller is not supported. To avoid a single point of failure, use a dual switch configuration. For full redundancy, a minimum of two paths between each initiator node and target node must be configured with each path on a separate switch.

  Figure 6-65 shows an example of a fully redundant iSCSI connection between two Storwize systems. In this example, the Storwize V7000 is composed of two I/O groups. Each node has a maximum of four initiator ports with two ports configured, through two switches, to the target ports on the other Storwize system. The first ports (orange) on each initiator and target nodes are connected through Ethernet switch 1. The second ports (blue) on each initiator and target nodes are connected through Ethernet switch 2. Each target node on the storage system has one iSCSI qualified name (IQN) that represents all the LUs on that node.

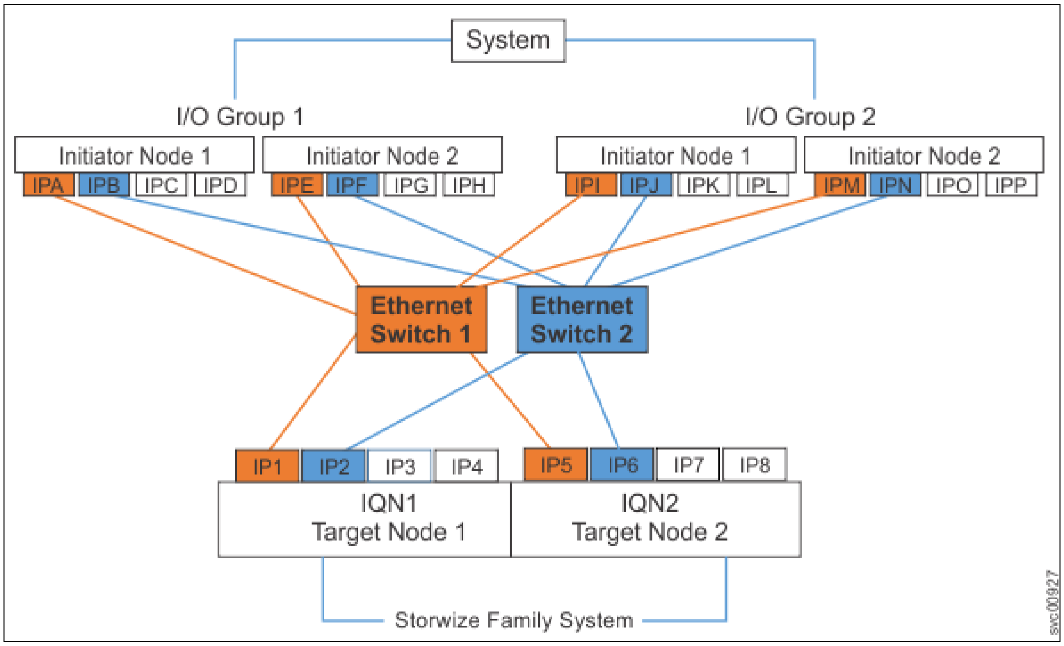


*Figure 6-65 Fully redundant iSCSI connection between two Storwize systems*

**IBM Spectrum Accelerate and Dell EqualLogic:**

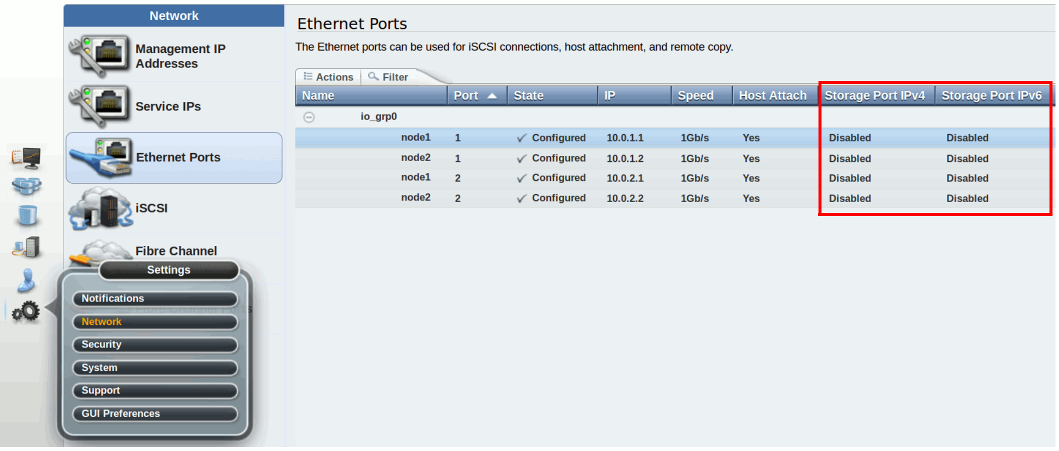
For an example of how to cable the IBM Spectrum Accelerate to the Storwize V7000 see the Storwize V7000 Knowledge Center at:

https://ibm.biz/BdsKSf

For an example of how to cable the Dell EqualLogic to the Storwize V7000 see the Storwize V7000 Knowledge Center at:

https://ibm.biz/BdsKSP

- The ports used for iSCSI attachment are enabled for external storage connections. By default, Ethernet ports are disabled for external storage connections. You can verify the setting of your Ethernet ports by navigating to **Settings** → **Network** and selecting **Ethernet Ports**, as shown in Figure 6-66.



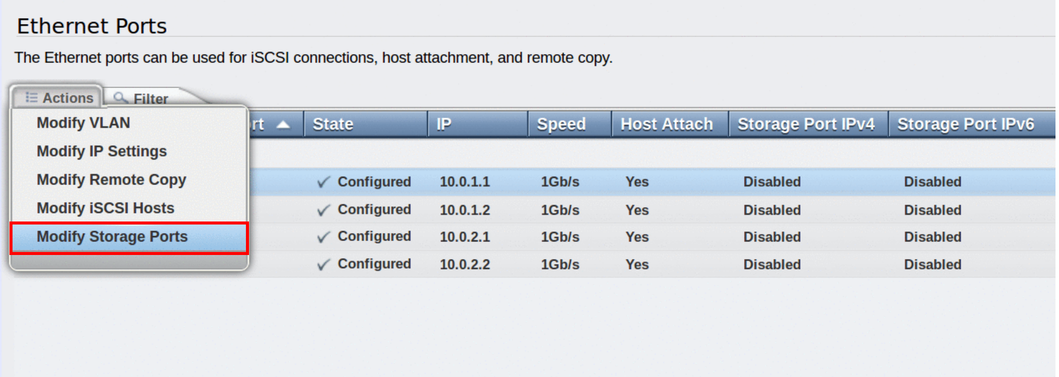
Ethernet Ports

The Ethernet ports can be used for iSCSI connections, host attachment, and remote copy.

| Name | Port | State | IP | Speed | Host Attach | Storage Port IPv4 | Storage Port IPv6 |
|---|---|---|---|---|---|---|---|
| io_grp0 | | | | | | | |
| node1 | 1 | ✓ Configured | 10.0.1.1 | 1Gb/s | Yes | Disabled | Disabled |
| node2 | 1 | ✓ Configured | 10.0.1.2 | 1Gb/s | Yes | Disabled | Disabled |
| node1 | 2 | ✓ Configured | 10.0.2.1 | 1Gb/s | Yes | Disabled | Disabled |
| node2 | 2 | ✓ Configured | 10.0.2.2 | 1Gb/s | Yes | Disabled | Disabled |

*Figure 6-66 Verifying Ethernet port settings*

To enable the port for external storage connections, select the port, click **Actions** and then **Modify Storage Ports**, as shown in Figure 6-67 on page 206. Alternatively, right-click the port and click **Modify Storage Ports**.



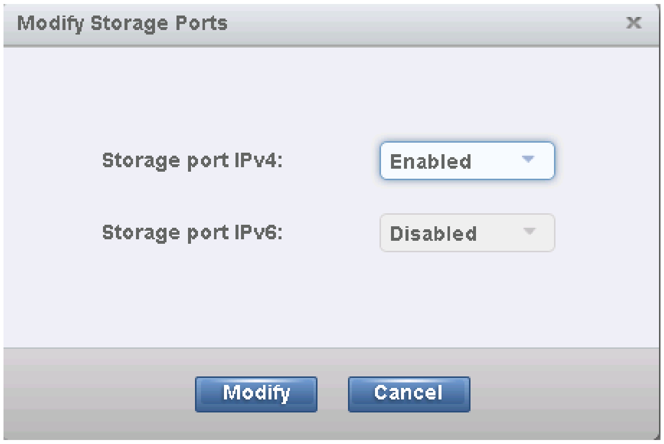
Ethernet Ports

The Ethernet ports can be used for iSCSI connections, host attachment, and remote copy.

| rt | State | IP | Speed | Host Attach | Storage Port IPv4 | Storage Port IPv6 |
|---|---|---|---|---|---|---|
| | ✓ Configured | 10.0.1.1 | 1Gb/s | Yes | Disabled | Disabled |
| | ✓ Configured | 10.0.1.2 | 1Gb/s | Yes | Disabled | Disabled |
| | ✓ Configured | 10.0.2.1 | 1Gb/s | Yes | Disabled | Disabled |
| | ✓ Configured | 10.0.2.2 | 1Gb/s | Yes | Disabled | Disabled |

*Figure 6-67 Modifying the external storage connection setting*

Set the port as **Enabled** for either IPv4 or IPv6, depending on the protocol version configured for the connection, as shown in Figure 6-68. Click **Modify**.

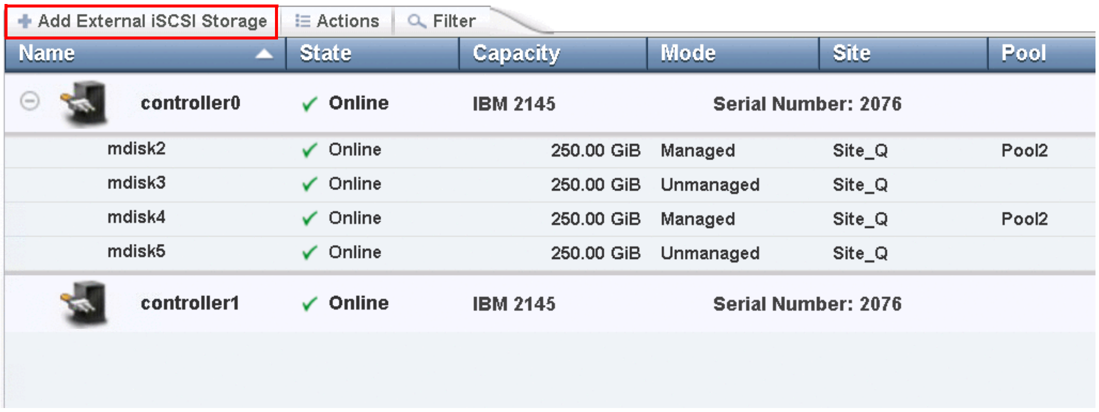


Storage port IPv4: Enabled

Storage port IPv6: Disabled

*Figure 6-68 Enabling an IPv4 port for external storage connections*

When all prerequisites are fulfilled, you are ready to add the iSCSI controller. To do so, navigate to **Pools** → **External Storage** and click **Add External iSCSI Storage**, as shown in Figure 6-69 on page 207.



| Name | State | Capacity | Mode | Site | Pool |
|---|---|---|---|---|---|
| controller0 | ✓ Online | IBM 2145 | Serial Number: 2076 | | |
| mdisk2 | ✓ Online | 250.00 GiB | Managed | Site_Q | Pool2 |
| mdisk3 | ✓ Online | 250.00 GiB | Unmanaged | Site_Q | |
| mdisk4 | ✓ Online | 250.00 GiB | Managed | Site_Q | Pool2 |
| mdisk5 | ✓ Online | 250.00 GiB | Unmanaged | Site_Q | |
| controller1 | ✓ Online | IBM 2145 | Serial Number: 2076 | | |

*Figure 6-69 Option to add an external iSCSI controller*

**Attention:** Unlike Fibre Channel connections, iSCSI connections require the Storwize V7000 to be configured at the *replication* layer for every type of external controller. However, as with Fibre Channel, if the external controller is a Storwize system, the controller must be configured at the *storage* layer. The default layer for a Storwize system is *storage*.

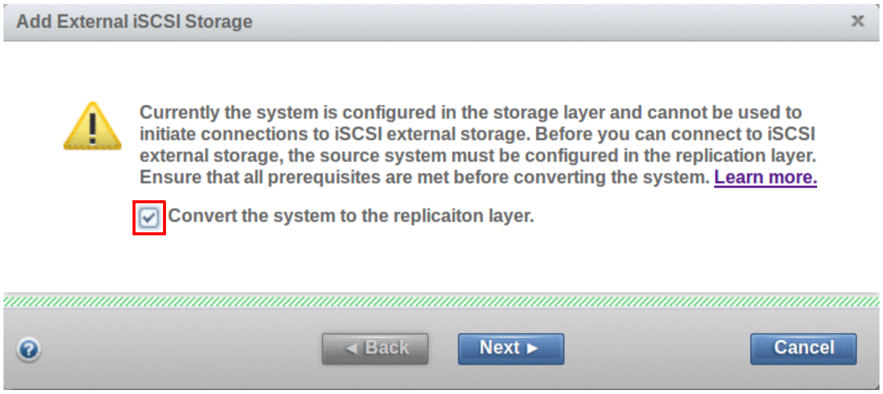
If the Storwize V7000 is not configured at the replication layer when **Add External iSCSI Storage** is selected, you are prompted to do so, as shown in Figure 6-70 on page 207.

If the Storwize controller is not configured at the storage layer this must be changed using the command-line interface.

Ensure the layer of the Storwize controller is correct by entering the following command: `svcinfo lssystem`

If needed, change the layer of the Storwize controller to storage by entering the following command: `chsystem -layer storage`

For more information about layers and how to change them see Chapter 11, “Advanced Copy Services” on page 413.



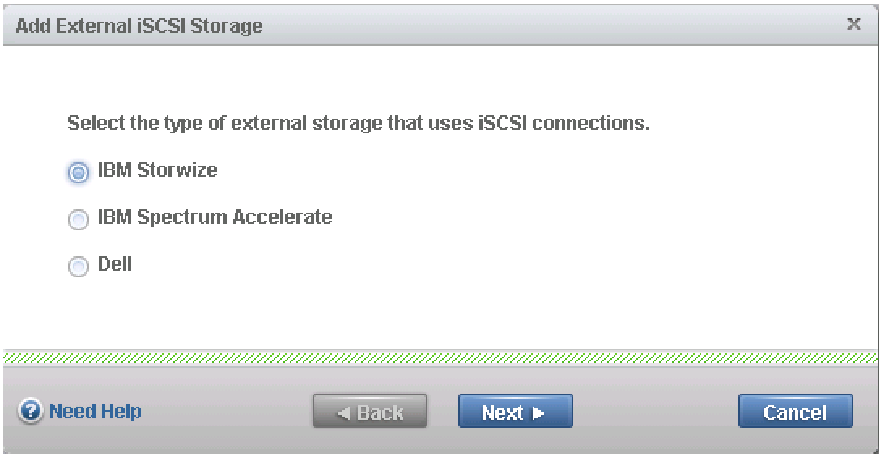
Currently the system is configured in the storage layer and cannot be used to initiate connections to iSCSI external storage. Before you can connect to iSCSI external storage, the source system must be configured in the replication layer. Ensure that all prerequisites are met before converting the system. Learn more.

☑ Convert the system to the replicaiton layer.

*Figure 6-70 Adding an external iSCSI controller: converting the system layer to replication*

Check **Convert the system to the replication layer** and click **Next**.

Select the type of external storage, as shown in Figure 6-71. For the purpose of this example the IBM Storwize type is chosen. Click **Next.**



Add External iSCSI Storage

Select the type of external storage that uses iSCSI connections.

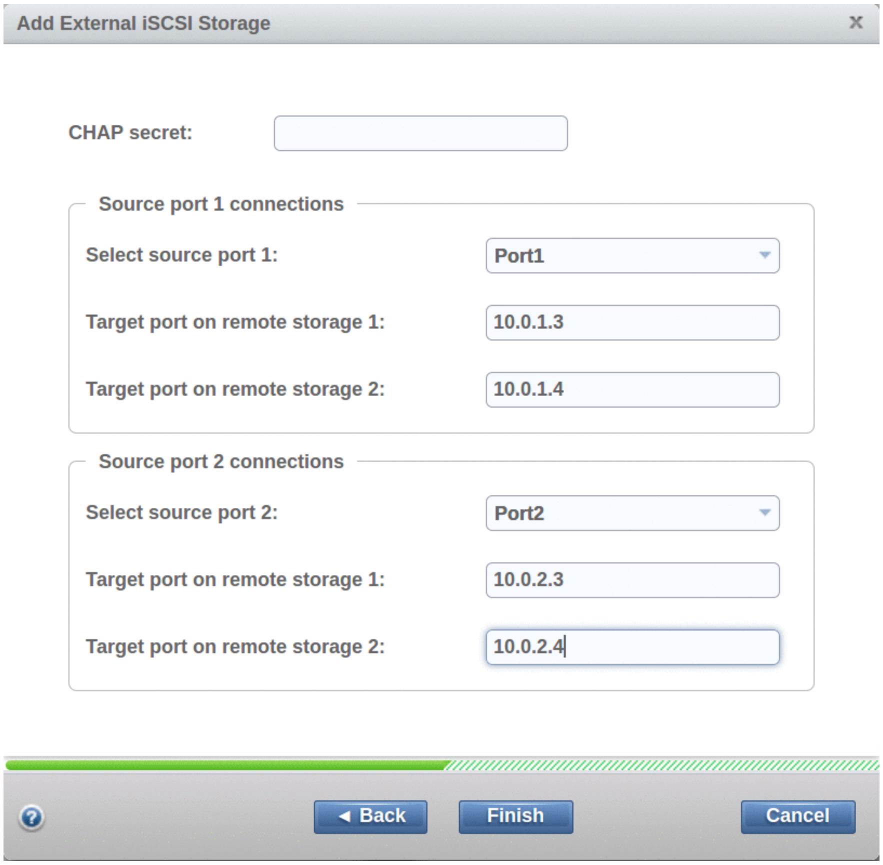
● IBM Storwize

○ IBM Spectrum Accelerate

○ Dell

*Figure 6-71 Adding an external iSCSI controller: controller type*

Enter the iSCSI connection details, as shown in Figure 6-72.



Add External iSCSI Storage

CHAP secret:

Source port 1 connections

Select source port 1: Port1

Target port on remote storage 1: 10.0.1.3

Target port on remote storage 2: 10.0.1.4

Source port 2 connections

Select source port 2: Port2

Target port on remote storage 1: 10.0.2.3

Target port on remote storage 2: 10.0.2.4

*Figure 6-72 Adding an external iSCSI controller: connection details*

Fill in each field as described below:

- **CHAP secret:** if the Challenge Handshake Authentication Protocol (CHAP) is used to secure iSCSI connections on the system, enter the current CHAP secret. This field is not required if you do not use CHAP.
- Source port 1 connections
  - **Select source port 1:** select one of the ports to be used as initiator for the iSCSI connection between the node and the external storage system.
  - **Target port on remote storage 1:** enter the IP address for one of the ports on the external storage system targeted by this source port.
  - **Target port on remote storage 2:** enter the IP address for the other port on the external storage system targeted by this source port.
- Source port 2 connections
  - **Select source port 2:** select the other port to be used as initiator for the iSCSI connection between the node and the external storage system.
  - **Target port on remote storage 1:** enter the IP address for one of the ports on the external storage system targeted by this source port.
  - **Target port on remote storage 2:** enter the IP address for the other port on the external storage system targeted by this source port.

The fields available vary depending on the configuration of your system and external controller type. However, the meaning of each field is always kept. The following fields can also be available:

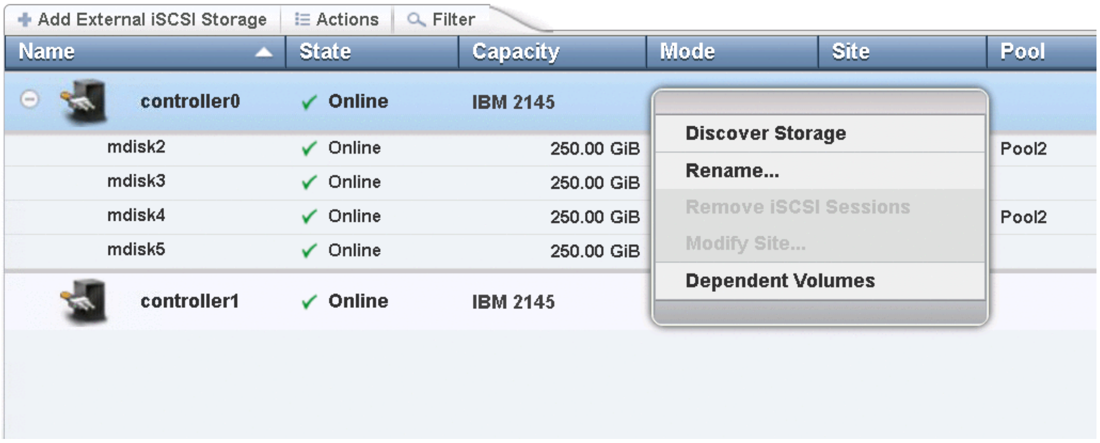
- **Site:** enter the site associated with the external storage system. This field is only shown for configurations using Hyperswap.
- **User name**: enter the user name associated with this connection. If the target storage system uses CHAP to authenticate connections, you must enter a user name. If you specify a user name, you must specify a CHAP secret. This field is not required if you do not use CHAP. This field is only shown for IBM Spectrum Accelerate and Dell EqualLogic controllers.

Click **Finish**. The system attempts to discover the target ports and establish iSCSI sessions between source and target. If the attempt is successful, the controller is added; otherwise the action fails.

## 6.4.3 Actions on external storage controllers

There are a number of actions that can be performed on external storage controllers. Some actions are available for external iSCSI controllers only.

To select any action right-click the controller, as shown in Figure 6-73. Alternatively, select the controller and click **Actions**.



| Name | State | Capacity | Mode | Site | Pool |
|---|---|---|---|---|---|
| controller0 | ✓ Online | IBM 2145 | | | |
| mdisk2 | ✓ Online | 250.00 GiB | | | Pool2 |
| mdisk3 | ✓ Online | 250.00 GiB | | | |
| mdisk4 | ✓ Online | 250.00 GiB | | | Pool2 |
| mdisk5 | ✓ Online | 250.00 GiB | | | |
| controller1 | ✓ Online | IBM 2145 | | | |

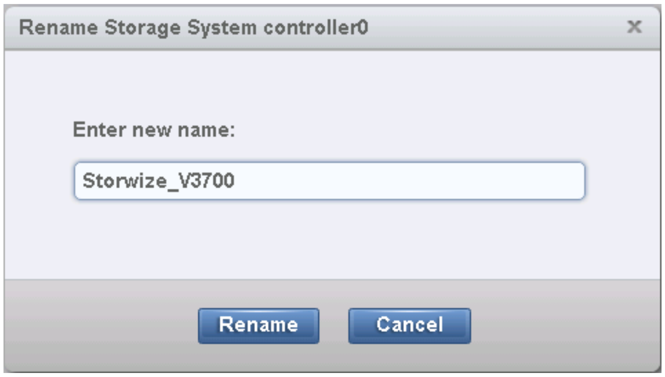
*Figure 6-73 Actions on external controllers*

## Discover storage

When you create or remove LUs on an external storage system, the change is not always automatically detected. If that is the case select **Discover Storage** for the system to rescan the Fibre Channel or iSCSI network. The rescan process discovers any new MDisks that were added to the system and rebalances MDisk access across the available ports. It also detects any loss of availability of the controller ports.

## Rename

Selecting **Rename** allows the user to modify the name of an external controller, as shown in Figure 6-74. Enter the new name and click **Rename**.

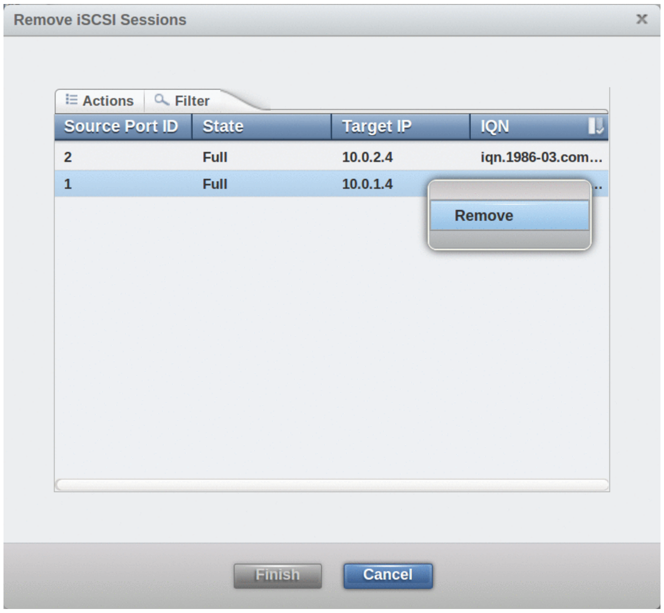


Enter new name: Storwize_V3700

*Figure 6-74 Renaming an external controller*

## Remove iSCSI sessions

This action is available only for external controllers attached with iSCSI. Right-click the session and select **Remove** to remove the iSCSI session established between the source and target port. Figure 6-75 on page 211.

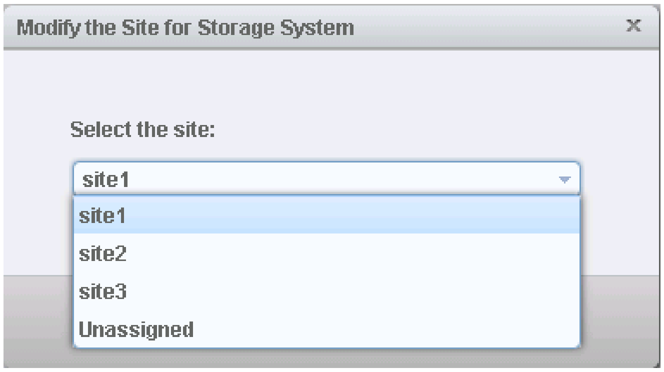


| Source Port ID | State | Target IP | IQN |
|---|---|---|---|
| 2 | Full | 10.0.2.4 | iqn.1986-03.com… |
| 1 | Full | 10.0.1.4 | .. |

Figure 6-75   *Removing iSCSI sessions from the external controller*

## Modify site

This action is available only for systems using Hyperswap. Selecting **Modify Site** allows the user to modify the site with which the external controller is associated, as shown Figure 6-76.



Select the site: site1

Figure 6-76   *Modifying the site of an external controller*

# 7

# Volumes

A volume is a logical disk provisioned out of a storage pool and is recognized by a host with a unique identifier (UID) field and a parameter list.

The first part of this chapter provides a brief overview of IBM Spectrum Virtualize volumes, the classes of volumes available, and the topologies that they are associated with. It also provides an overview of advanced customization available.

The second part describes how to create volumes using the GUI's *Quick* and *Advanced* volume creation menus, and shows you how to map these to defined hosts.

The third part provides an introduction to the new volume manipulation commands, designed to facilitate the creation and administration of volumes used for *IBM HyperSwap* topology.

> **Note:** Advanced host and volume administration, such as volume migration, creating volume copies, and so on, is described in Chapter 11, "Advanced Copy Services" on page 413.

This chapter includes the following topics:

- 7.1, "Introduction to volumes" on page 214
- 7.2, "Create Volumes menu" on page 224
- 7.3, "Creating volumes using the Quick Volume Creation" on page 228
- 7.4, "Mapping a volume to the host" on page 234
- 7.5, "Creating Custom volumes using the Advanced menu" on page 236
- 7.6, "HyperSwap and the mkvolume command" on page 244
- 7.7, "Mapping Volumes to Host after volume creation" on page 249
- 7.8, "Migrating a volume to another storage pool" on page 251
- 7.9, "Migrating volumes using the volume copy feature" on page 254
- 7.10, "Volume operations using CLI" on page 258

## 7.1 Introduction to volumes

A volume is a logical disk that the system presents to attached hosts, for an IBM Spectrum Virtualize system the volume presented is from a virtual disk (VDisk). A VDisk is a discrete area of usable storage that has been virtualized, using IBM Spectrum Virtualize code, from storage area network (SAN) storage that is managed by the IBM Spectrum Virtualize cluster. The term *virtual* is used because the volume presented does not necessarily exist on a single physical entity.

Volumes have the following characteristics or attributes:

- Volumes can be created and deleted.
- Volumes can be resized (expanded or shrunk).
- Volume extents can be migrated at run time to another MDisk or storage pool.
- Volumes can be created as fully allocated or thin-provisioned. A conversion from a fully allocated to a thin-provisioned volume and vice versa can be done at run time.
- Volumes can be stored in multiple storage pools (mirrored) to make them resistant to disk subsystem failures or to improve the read performance.
- Volumes can be mirrored synchronously or asynchronously for longer distances. An IBM Spectrum Virtualize system can run active volume mirrors to a maximum of three other IBM Spectrum Virtualize systems, but not from the same volume.
- Volumes can be copied by using FlashCopy. Multiple snapshots and quick restore from snapshots (reverse FlashCopy) are supported.
- Volumes can be compressed.
- Volumes can be virtual. The system supports VMware vSphere Virtual Volumes, sometimes referred to as VVols, which allow VMware vCenter to manage system objects, such as volumes and pools. The system administrator can create these objects and assign ownership to VMware administrators to simplify management of these objects.

> **Note:** A managed disk (MDisk) is a logical unit of physical storage. MDisks are either Redundant Arrays of Independent Disks (RAID) from internal storage, or external physical disks that are presented as a single logical disk on the SAN. Each MDisk is divided into several extents, which are numbered, from 0, sequentially from the start to the end of the MDisk. The extent size is a property of the storage pools that the MDisks are added to.
>
> **Attention:** MDisks are not visible to host systems.

Volumes have two major modes: Managed mode and image mode. Managed mode volumes have two policies: The sequential policy and the striped policy. Policies define how the extents of a volume are allocated from a storage pool.

The *type* attribute of a volume defines the allocation of extents that make up the volume copy:

- A *striped* volume contains a volume copy that has one extent allocated in turn from each MDisk that is in the storage pool. This is the default option, but you can also supply a list of MDisks to use as the stripe set as shown in Figure 7-1 on page 215.

**Attention:** By default, striped volume copies are striped across all MDisks in the storage pool. If some of the MDisks are smaller than others, the extents on the smaller MDisks are used up before the larger MDisks run out of extents. Manually specifying the stripe set in this case might result in the volume copy not being created.

If you are unsure if there is sufficient free space to create a striped volume copy, select one of the following options:

- Check the free space on each MDisk in the storage pool using the `lsfreeextents` command.
- Let the system automatically create the volume copy by not supplying a specific stripe set.

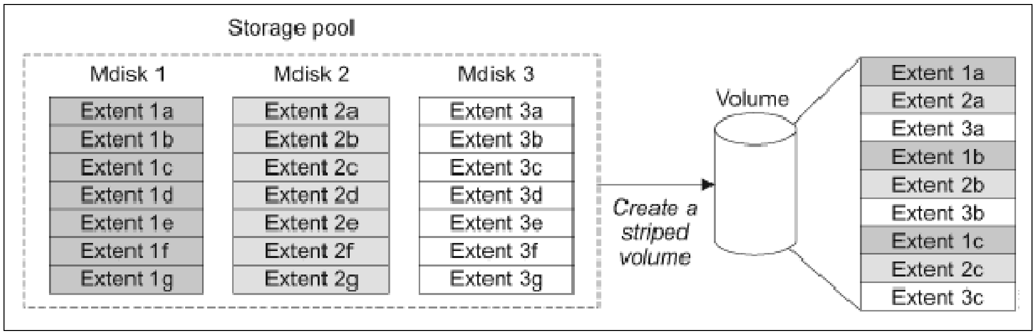


*Figure 7-1 Striped extent allocation*

- A *sequential* volume contains a volume copy that has extents allocated sequentially on one MDisk.
- *Image-mode* volumes are a special type of volume that have a direct relationship with one MDisk.

## 7.1.1 Image mode volumes

Image mode volumes are used to migrate LUNs that were previously mapped directly to host servers over to the control of the IBM Spectrum Virtualize system. Image mode provides a one-to-one mapping between the logical block addresses (LBAs) between a volume and an MDisk. Image mode volumes have a minimum size of one block (512 bytes) and always occupy at least one extent.

An image mode MDisk is mapped to one, and only one, image mode volume.

The volume capacity that is specified must be equal to the size of the image mode MDisk. When you create an image mode volume, the specified MDisk must be in unmanaged mode and must not be a member of a storage pool. The MDisk is made a member of the specified storage pool (`Storage Pool_IMG_`*`xxx`*) as a result of creating the image mode volume.

The Spectrum Virtualize also supports the reverse process, in which a managed mode volume can be migrated to an image mode volume. If a volume is migrated to another MDisk, it is represented as being in managed mode during the migration, and is only represented as an image mode volume after it reaches the state where it is a straight-through mapping.

An image mode MDisk is associated with exactly one volume. If the (image mode) MDisk is not a multiple of the MDisk Group’s extent size, the last extent is partial (not filled). An image mode volume is a pass-through one-to-one map of its MDisk. It cannot be a quorum disk and it does not have any metadata extents that are assigned to it from the IBM Spectrum Virtualize system. Managed or image mode MDisks are always members of a storage pool.

It is a preferred practice to put image mode MDisks in a dedicated storage pool and use a special name for it (for example, `Storage Pool_IMG_`*`xxx`*). The extent size that is chosen for this specific storage pool must be the same as the extent size into which you plan to migrate the data. All of the IBM Spectrum Virtualize copy services functions can be applied to image mode disks. See Figure 7-2.

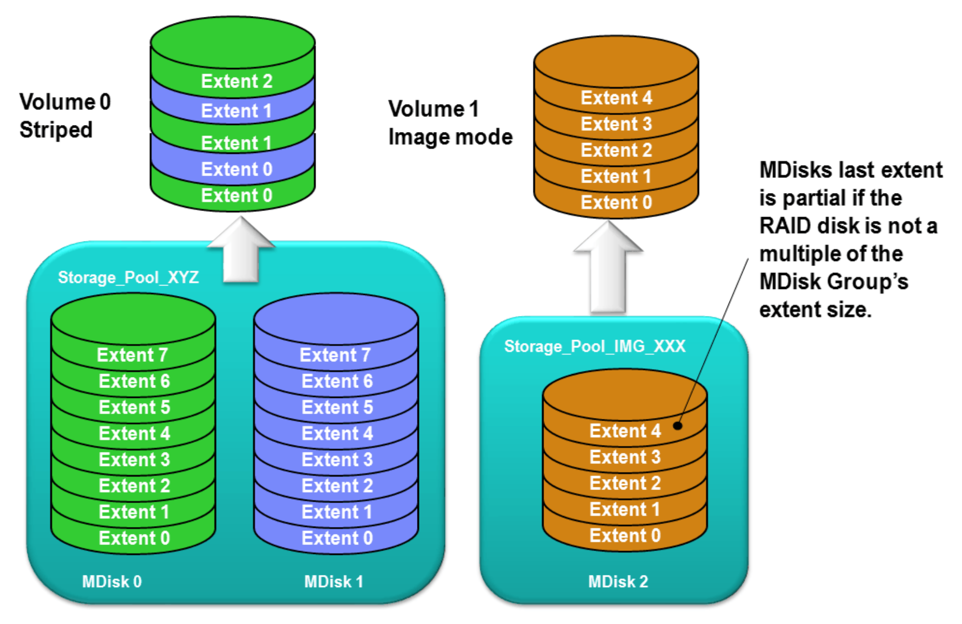


*Figure 7-2 Image mode volume versus striped volume*

### 7.1.2 Managed mode volumes

Volumes operating in managed mode provide a full set of virtualization functions. Within a storage pool, the IBM Spectrum Virtualize supports an arbitrary relationship between extents on (managed mode) volumes and extents on MDisks. Each volume extent maps to exactly one MDisk extent.

Figure 7-3 on page 217 shows this mapping. It also shows a volume that consists of several extents that are shown as V0 - V7. Each of these extents is mapped to an extent on one of the MDisks: A, B, or C. The mapping table stores the details of this indirection.

Several of the MDisk extents are unused. No volume extent maps to them. These unused extents are available for use in creating volumes, migration, expansion, and so on.

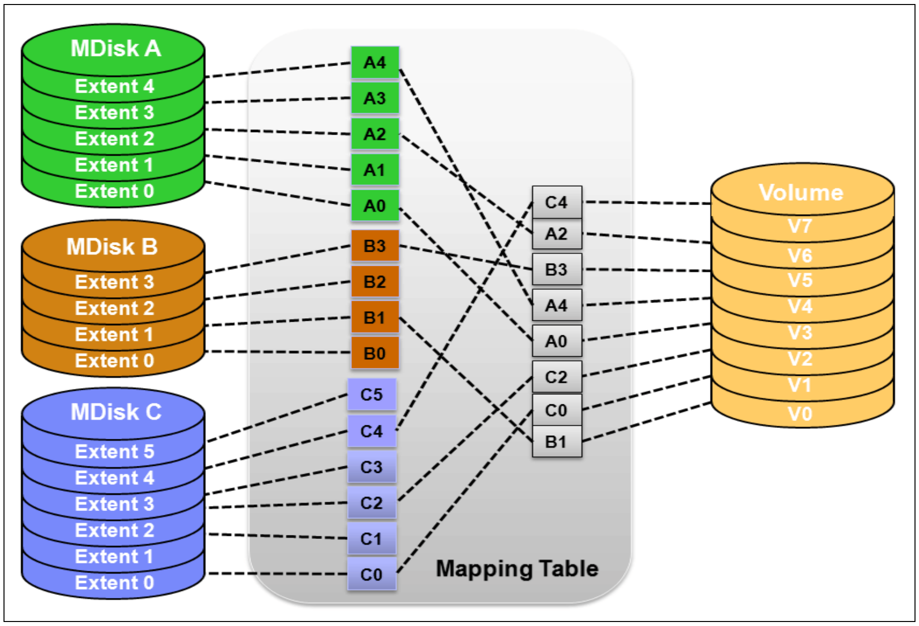


*Figure 7-3 Simple view of block virtualization*

The allocation of a specific number of extents from a specific set of MDisks is performed by the following algorithm: If the set of MDisks from which to allocate extents contains more than one MDisk, extents are allocated from MDisks in a round-robin fashion. If an MDisk has no free extents when its turn arrives, its turn is missed and the round-robin moves to the next MDisk in the set that has a free extent.

When a volume is created, the first MDisk from which to allocate an extent is chosen in a pseudo-random way rather than by choosing the next disk in a round-robin fashion. The pseudo-random algorithm avoids the situation where the *striping effect* that is inherent in a round-robin algorithm that places the first extent for many volumes on the same MDisk.

Placing the first extent of several volumes on the same MDisk can lead to poor performance for workloads that place a large I/O load on the first extent of each volume, or that create multiple sequential streams.

## 7.1.3 Cache mode for volumes

It is also possible to define the cache characteristics of a volume. Under normal conditions, a volume's read and write data is held in the cache of its preferred node, with a mirrored copy of write data that is held in the partner node of the same I/O Group. However, it is possible to create a volume with cache disabled, which means that the I/Os are passed directly through to the back-end storage controller rather than being held in the node's cache.

Having cache-disabled volumes makes it possible to use the native copy services in the underlying RAID array controller for MDisks (LUNs) that are used as the IBM Spectrum Virtualize image mode volumes. Using IBM Spectrum Virtualize Copy Services rather than the underlying disk controller copy services gives better results.

Cache characteristics of a volume can have any of the following settings:

- *readwrite*. All read and write I/O operations that are performed by the volume are stored in cache. This is the default cache mode for all volumes.
- *readonly*. All read I/O operations that are performed by the volume are stored in cache.

- *disabled*. All read and write I/O operations that are performed by the volume are not stored in cache. Under normal conditions, a volume's read and write data is held in the cache of its preferred node, with a mirrored copy of write data that is held in the partner node of the same I/O Group. With cache disabled volume, the I/Os are passed directly through to the back-end storage controller rather than being held in the node's cache.

> **Note:** Having cache-disabled volumes makes it possible to use the native copy services in the underlying RAID array controller for MDisks (LUNs) that are used as IBM Spectrum Virtualize image mode volumes. It is strongly recommended to consult IBM Support prior to turning off the cache for volumes in production environment to avoid any performance degradation.

## 7.1.4 Mirrored volumes

The mirrored volume feature provides a simple RAID 1 function; therefore, a volume has two physical copies of its data. This approach enables the volume to remain online and accessible even if one of the MDisks sustains a failure that causes it to become inaccessible.

The two copies of the volume often are allocated from separate storage pools or by using image-mode copies. The volume can participate in FlashCopy and remote copy relationships, it is serviced by an I/O Group, and it has a preferred node.

Each copy is not a separate object and cannot be created or manipulated except in the context of the volume. Copies are identified through the configuration interface with a copy ID of their parent volume. This copy ID can be 0 or 1.

This feature provides a point-in-time copy functionality that is achieved by “splitting” a copy from the volume. However, the mirrored volume feature does not address other forms of mirroring that are based on remote copy, which is sometimes called *IBM HyperSwap*, that mirrors volumes across I/O Groups or clustered systems. It is also not intended to manage mirroring or remote copy functions in back-end controllers.

Figure 7-4 provides an overview of volume mirroring.

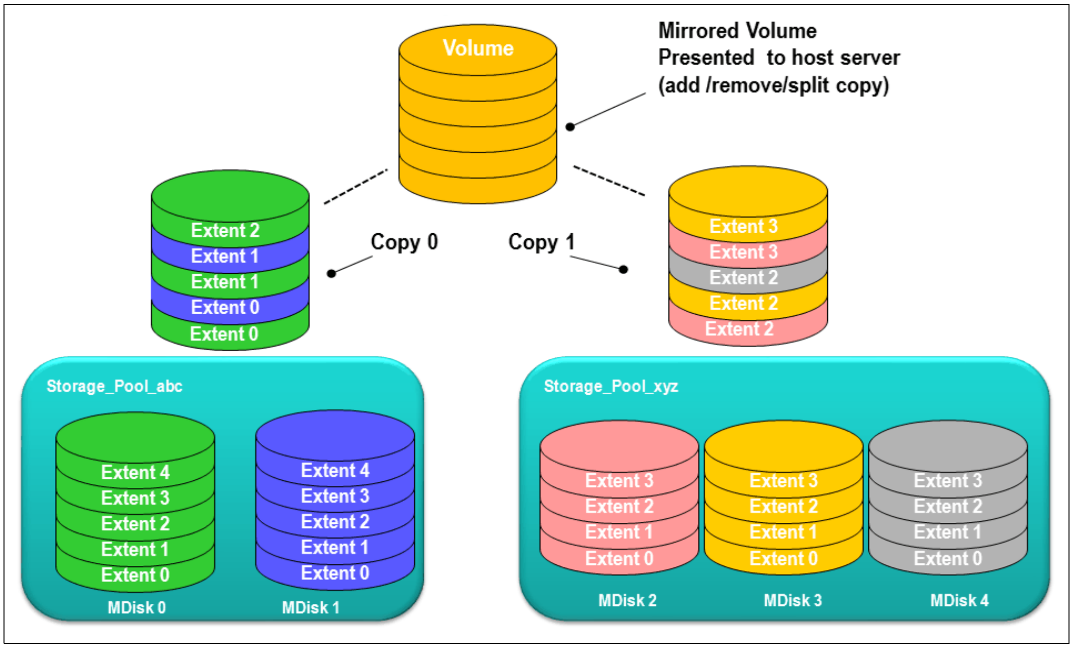


*Figure 7-4 Volume mirroring overview*

A second copy can be added to a volume with a single copy or removed from a volume with two copies. Checks prevent the accidental removal of the only remaining copy of a volume. A newly created, unformatted volume with two copies initially has the two copies in an out-of-synchronization state. The primary copy is defined as “fresh” and the secondary copy is defined as “stale”.

The synchronization process updates the secondary copy until it is fully synchronized. This update is done at the default *synchronization rate* or at a rate that is defined when the volume is created or modified. The synchronization status for mirrored volumes is recorded on the quorum disk.

If a two-copy mirrored volume is created with the **`format`** parameter, both copies are formatted in parallel, and the volume comes online when both operations are complete with the copies in sync.

If mirrored volumes are expanded or shrunk, all of their copies are also expanded or shrunk.

If it is known that MDisk space (which is used for creating copies) is already formatted or if the user does not require read stability, a `no synchronization` option can be selected that declares the copies as `synchronized` (even when they are not).

To minimize the time that is required to resynchronize a copy that is out of sync, only the 256 kibibyte (KiB) grains that were written to since the synchronization was lost are copied. This approach is known as an *incremental synchronization*. Only the changed grains must be copied to restore synchronization.

> **Important:** An unmirrored volume can be migrated from one location to another by adding a second copy to the wanted destination, waiting for the two copies to synchronize, and then removing the original copy 0. This operation can be stopped at any time. The two copies can be in separate storage pools with separate extent sizes.

Where there are two copies of a volume, one copy is known as the *primary copy*. If the primary is available and synchronized, reads from the volume are directed to it. The user can select the primary when the volume is created or can change it later.

Placing the primary copy on a high-performance controller maximizes the read performance of the volume.

### Write I/O operations data flow with a mirrored volume

For write I/O operations to a mirrored volume, the IBM Spectrum Virtualize preferred node definition, with the multipathing driver on the host, is used to determine the preferred path. The host routes the I/Os through the preferred path, and the corresponding node is responsible for further destaging written data from cache to both volume copies. Figure 7-5 shows the data flow for write I/O processing when volume mirroring is used.

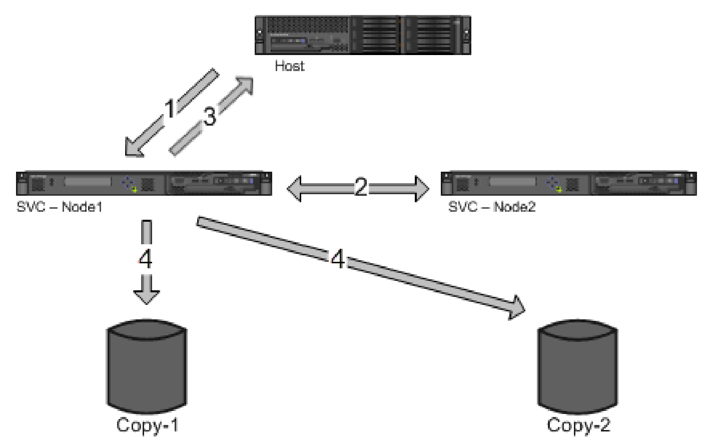


*Figure 7-5 Data flow for write I/O processing in a mirrored volume*

As shown in Figure 7-5, all the writes are sent by the host to the preferred node for each volume (1). Then, the data is mirrored to the cache of the partner node in the I/O Group (2), and acknowledgment of the write operation is sent to the host (3). The preferred node then destaged the written data to the two volume copies (4).

A volume with copies can be checked to see whether all of the copies are identical or consistent. If a medium error is encountered while it is reading from one copy, it is repaired by using data from the other copy. This consistency check is performed asynchronously with host I/O.

> **Important:** Mirrored volumes can be taken offline if there is no quorum disk available. This behavior occurs because the synchronization status for mirrored volumes is recorded on the quorum disk.

Mirrored volumes use bitmap space at a rate of 1 bit per 256 KiB grain, which translates to 1 MiB of bitmap space supporting 2 TiB of mirrored volumes. The default allocation of bitmap space is 20 MiB, which supports 40 TiB of mirrored volumes. If all 512 MiB of variable bitmap space is allocated to mirrored volumes, 1 PiB of mirrored volumes can be supported.

Table 7-1 shows you the bitmap space default configuration.

*Table 7-1 Bitmap space default configuration*

| Copy service | Minimum allocated bitmap space | Default allocated bitmap space | Maximum allocated bitmap space | Minimum[a] functionality when using the default values |
|---|---|---|---|---|
| Remote copy[b] | 0 | 20 MiB | 512 MiB | 40 TiB of remote mirroring volume capacity |
| FlashCopy[c] | 0 | 20 MiB | 2 GiB | ► 10 TiB of FlashCopy source volume capacity<br>► 5 TiB of incremental FlashCopy source volume capacity |
| Volume mirroring | 0 | 20 MiB | 512 MiB | 40 TiB of mirrored volumes |
| RAID | 0 | 40 MiB | 512 MiB | ► 80 TiB array capacity using RAID 0, 1, or 10<br>► 80 TiB array capacity in three-disk RAID 5 array<br>► Slightly less than 120 TiB array capacity in five-disk RAID 6 array |

[a] The actual amount of functionality might increase based on settings such as grain size and strip size. RAID is subject to a 15% margin or error.
[b] Remote copy includes Metro Mirror, Global Mirror, and active-active relationships.
[c] FlashCopy includes the FlashCopy function, Global Mirror with change volumes, and active-active relationships.

The sum of all bitmap memory allocation for all functions except FlashCopy must not exceed 552 MiB.

## 7.1.5 Thin-provisioned volumes

Volumes can be configured to be thin-provisioned or fully allocated. A *thin-provisioned* volume behaves as though application reads and writes were fully allocated. When a thin-provisioned volume is created, the user specifies two capacities:

- The real physical capacity that is allocated to the volume from the storage pool
- Its virtual capacity that is available to the host.

In a *fully allocated* volume, these two values are the same.

Therefore, the real capacity determines the quantity of MDisk extents that is initially allocated to the volume. The *virtual capacity* is the capacity of the volume that is reported to all other IBM Spectrum Virtualize components (for example, FlashCopy, cache, and remote copy), and to the host servers.

The *real capacity* is used to store the user data and the metadata for the thin-provisioned volume. The real capacity can be specified as an absolute value, or as a percentage of the virtual capacity.

Thin-provisioned volumes can be used as volumes that are assigned to the host, by FlashCopy to implement thin-provisioned FlashCopy targets, and with the mirrored volumes feature.

When a thin-provisioned volume is initially created, a small amount of the real capacity is used for initial metadata. I/Os are written to grains of the thin volume that were not previously written, which causes grains of the real capacity to be used to store metadata and the actual user data. I/Os are written to grains that were previously written, which updates the grain where data was previously written.

The grain size is defined when the volume is created. The grain size can be 32 KiB, 64 KiB, 128 KiB, or 256 KiB. The default grain size is 256 KiB, which is the recommended option. If you select 32 KiB for the grain size, the volume size cannot exceed 260 TiB. The grain size cannot be changed after the thin-provisioned volume is created. Generally, smaller grain sizes save space, but they require more metadata access, which can adversely affect performance.

When using thin-provisioned volume as a FlashCopy source or target volume, use 256 KiB to maximize performance. When using thin-provisioned volume as a FlashCopy source or target volume, specify the same grain size for the volume and for the FlashCopy function.

Figure 7-6 shows the thin-provisioning concept.

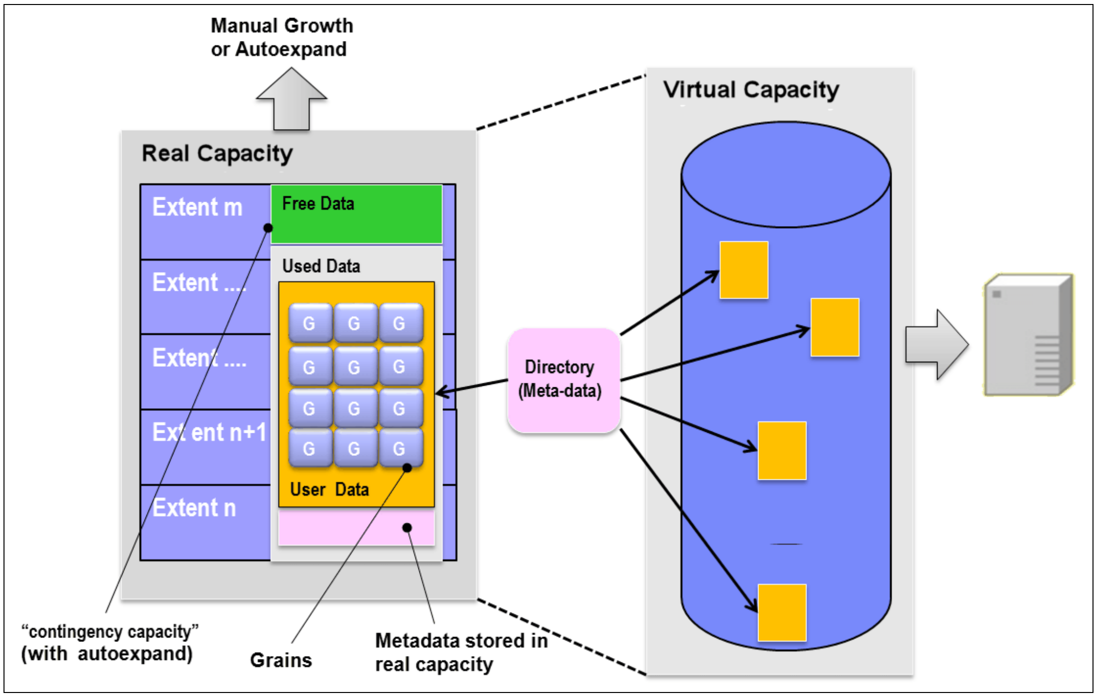


*Figure 7-6 Conceptual diagram of thin-provisioned volume*

Thin-provisioned volumes store user data and metadata. Each grain of data requires metadata to be stored. Therefore, the I/O rates that are obtained from thin-provisioned volumes are less than the I/O rates that are obtained from fully allocated volumes.

The metadata storage used is never greater than 0.1% of the user data. The resource usage is independent of the virtual capacity of the volume. If you are using the thin-provisioned volume directly with a host system, use a small grain size.

**Thin-provisioned volume format:** Thin-provisioned volumes do not need formatting. A read I/O, which requests data from deallocated data space, returns zeros. When a write I/O causes space to be allocated, the grain is “zeroed” before use.

The real capacity of a thin volume can be changed if the volume is not in image mode. Increasing the real capacity enables a larger amount of data and metadata to be stored on the volume. Thin-provisioned volumes use the real capacity that is provided in ascending order as new data is written to the volume. If the user initially assigns too much real capacity to the volume, the real capacity can be reduced to free storage for other uses.

A thin-provisioned volume can be configured to *autoexpand*. This feature causes the IBM Spectrum Virtualize to automatically add a fixed amount of more real capacity to the thin volume as required. Therefore, autoexpand attempts to maintain a fixed amount of unused real capacity for the volume, which is known as the *contingency capacity*.

The contingency capacity is initially set to the real capacity that is assigned when the volume is created. If the user modifies the real capacity, the contingency capacity is reset to be the difference between the used capacity and real capacity.

A volume that is created without the autoexpand feature, and therefore has a zero contingency capacity, goes offline when the real capacity is used and it must expand.

Autoexpand does not cause the real capacity to grow much beyond the virtual capacity. The real capacity can be manually expanded to more than the maximum that is required by the current virtual capacity, and the contingency capacity is recalculated.

To support the auto expansion of thin-provisioned volumes, the storage pools from which they are allocated have a configurable capacity warning. When the used capacity of the pool exceeds the warning capacity, a warning event is logged. For example, if a warning of 80% is specified, the event is logged when 20% of the free capacity remains.

A thin-provisioned volume can be converted nondisruptively to a fully allocated volume (or vice versa) by using the volume mirroring function. For example, the system allows an user to add a thin-provisioned copy to a fully allocated primary volume and then remove the fully allocated copy from the volume after they are synchronized.

The fully allocated-to-thin-provisioned migration procedure uses a zero-detection algorithm so that grains that contain all zeros do not cause any real capacity to be used.

## 7.1.6 Compressed Volumes

This is a custom type of volume where data is compressed as it is written to disk, saving additional space. To use the compression function, you must obtain the IBM Real-time Compression license.

## 7.1.7 Volumes for various topologies

A *Basic* volume is the simplest form of volume. It consists of a single volume copy, made up of extents *striped* across all MDisks in a storage pool. It services I/O using *readwrite* cache and is classified as *fully allocated* (reported real capacity and virtual capacity are equal). You can create other forms of volumes, depending on the type of *topology* that is configured on your system:

- With *standard topology*, which is a single-site configuration, you can create a *basic* volume or a *mirrored* volume.

  By using volume mirroring, a volume can have two physical copies. Each volume copy can belong to a different pool, and each copy has the same virtual capacity as the volume. In the management GUI, an asterisk indicates the primary copy of the mirrored volume. The primary copy indicates the preferred volume for read requests.

- With *HyperSwap topology*, which is a three-site HA configuration, you can create a *basic* volume or a *HyperSwap* volume.

  HyperSwap volumes create copies on separate sites for systems that are configured with HyperSwap topology. Data that is written to a HyperSwap volume is automatically sent to both copies, so that either site can provide access to the volume if the other site becomes unavailable.

  > **Note:** IBM Storwize V7000 product supports the IBM HyperSwap topology. However IBM Storwize V7000 does not support Stretched Cluster or Enhanced Stretched Cluster topologies. The Stretched Cluster or Enhanced Stretched Cluster topologies are only supported in IBM SAN Volume Controller environment.

- Virtual Volumes (VVols): The IBM Spectrum Virtualize V7.6 release also introduces *Virtual Volumes*. These are available in a system configuration that supports VMware vSphere Virtual Volumes, sometimes referred to as VVols. These allow VMware vCenter to manage system objects, such as volumes and pools. The Spectrum Virtualize system administrators can create volume objects of this class, and assign ownership to VMware administrators to simplify management.

> **Note:** For more information on configuring vVol with IBM Spectrum Virtualize, please refer to *Configuring VMware Virtual Volumes for Systems Powered by IBM Spectrum Virtualize*, SG24-8328 and

> **Note:** From V7.4 onwards, it is possible to prevent accidental deletion of volumes, if they have recently performed any I/O operations. This feature is called *Volume protection*, and it prevents active volumes, or host mappings from being deleted inadvertently. This is done using a global system setting. For more information, see the “Enabling volume protection” topic in the IBM Knowledge Center:
>
> `https://ibm.biz/BdsKhk`

## 7.2 Create Volumes menu

The GUI is the simplest means of volume creation, and presents different options in the Create Volumes menu depending on the topology of the system.

To start the process of creating a new volume, click the dynamic menu (function icons), open the Volumes menu and click the **Volumes** option of the IBM Spectrum Virtualize graphical user interface as shown in Figure 7-7.

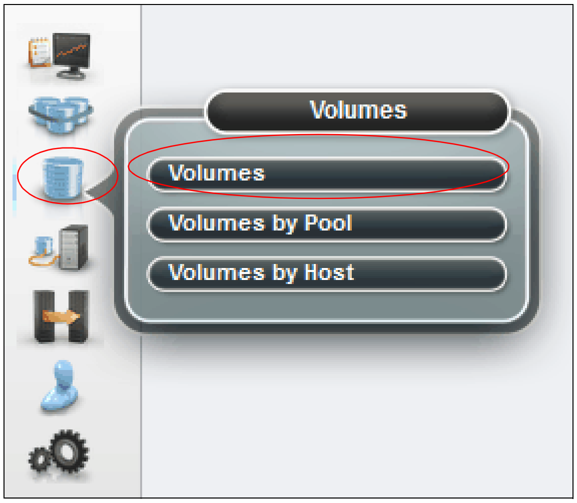
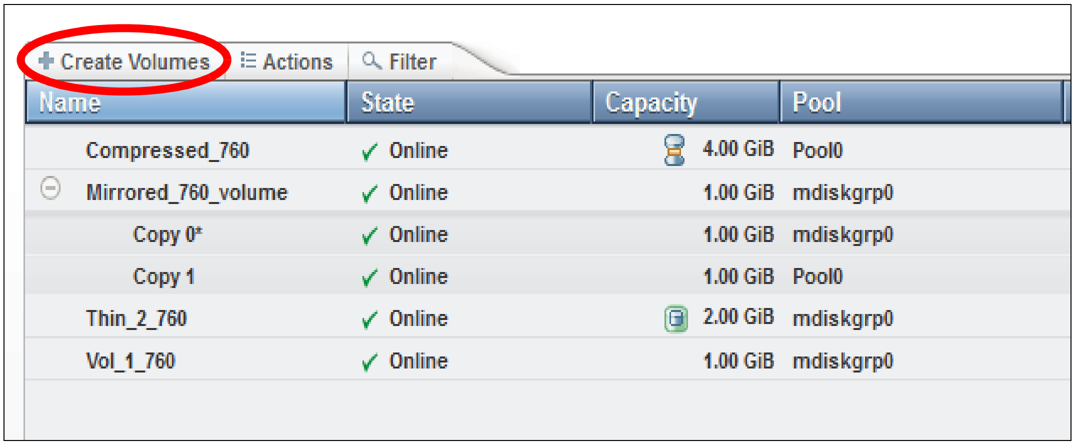


*Figure 7-7 Volumes menu*

A list of existing volumes, their state, capacity, and associated storage pools, is displayed.

To define a new *Basic* volume, click the **Create Volumes** option on the tab header as shown in Figure 7-8.



| Name | State | Capacity | Pool |
|---|---|---|---|
| Compressed_760 | ✓ Online | 4.00 GiB | Pool0 |
| Mirrored_760_volume | ✓ Online | 1.00 GiB | mdiskgrp0 |
| Copy 0* | ✓ Online | 1.00 GiB | mdiskgrp0 |
| Copy 1 | ✓ Online | 1.00 GiB | Pool0 |
| Thin_2_760 | ✓ Online | 2.00 GiB | mdiskgrp0 |
| Vol_1_760 | ✓ Online | 1.00 GiB | mdiskgrp0 |

*Figure 7-8 Create Volume window*

The **Create Volume** options tab opens the Create Volumes menu, which displays two potential creation methods, *Quick Volume Creation* and *Advanced* and volume classes.

> **Note:** The volume classes displayed on the Create Volume menu depend on the topology of the system.

Volumes can be created using the Quick Volume Creation submenu, or the Advanced submenu, as shown in Figure 7-9.

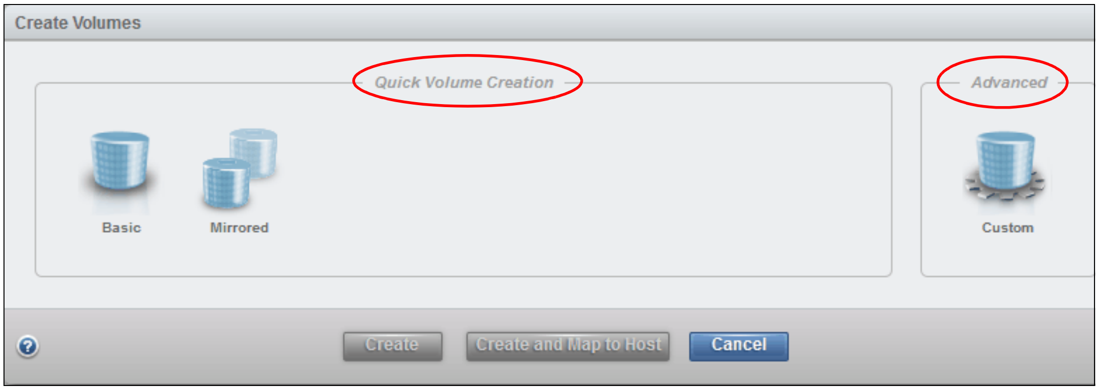


*Figure 7-9 Quick and Advanced Volume Creation options*

In the previous example, the Quick Volume Creation submenu shows icons that enable the quick creation of *Basic* and *Mirrored* volumes (in *standard topology*):

- For a *HyperSwap* topology, the Create Volumes menu displays as shown in Figure 7-10.

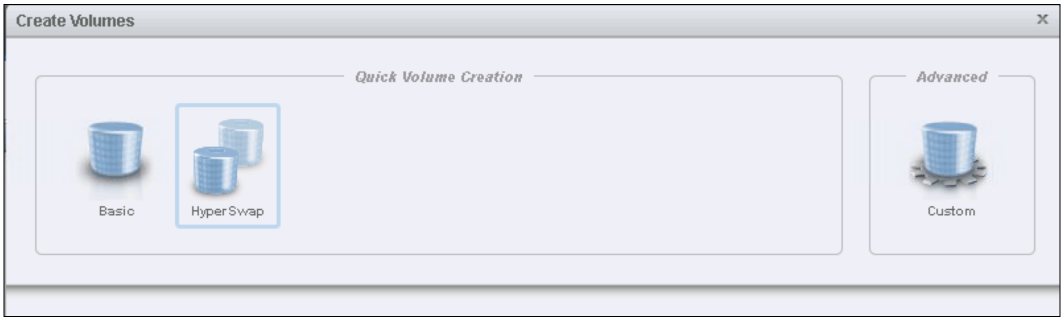
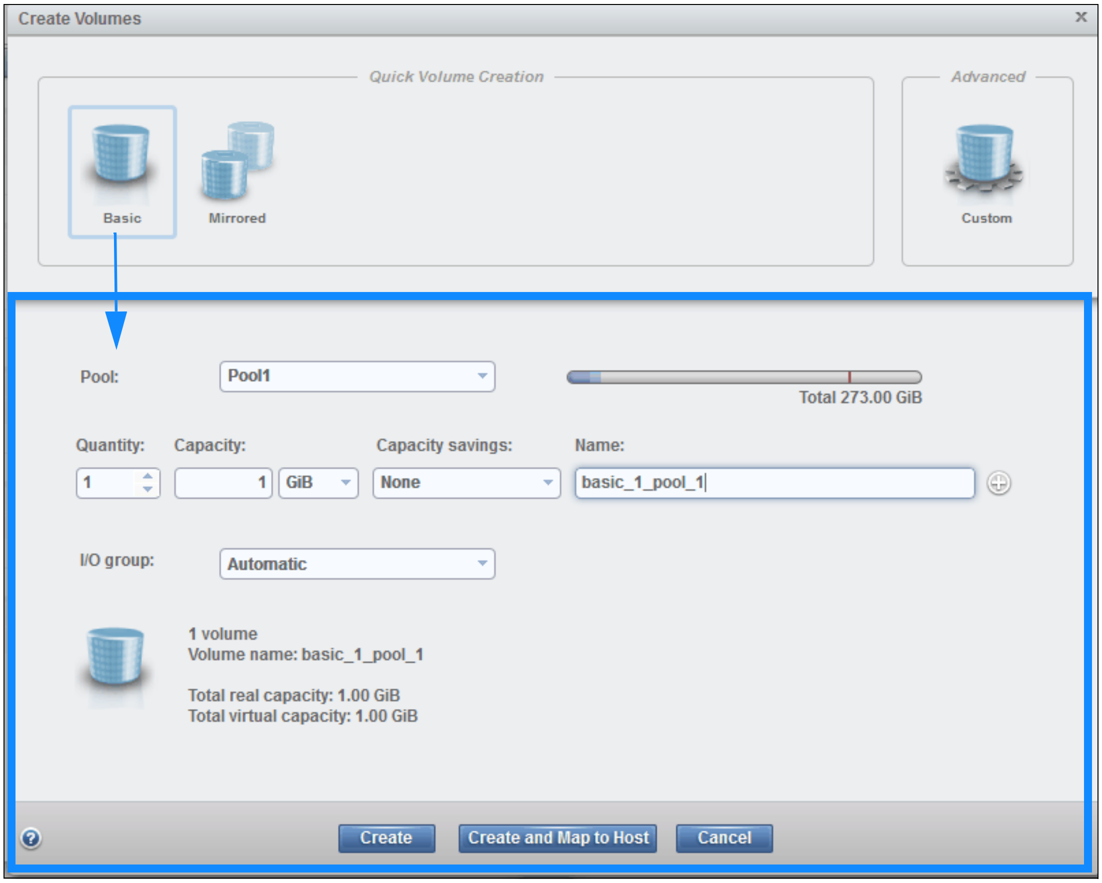


*Figure 7-10 Create Volumes menu with HyperSwap Topology*

Independent of the topology of the system, the Create Volume menu displays a *Basic* volume icon, in the Quick Volume Creation submenu, and always shows a *Custom* volume icon in the Advanced submenu. The Advanced submenu shows a *Custom* icon that can be used to customize parameters of volumes. *Custom* volumes are described in more detail later in this section.

Clicking any of the 3 x icons in the Create Volumes window opens a drop-down window where volume details can be entered. The example, as shown in Figure 7-11, uses a *Basic* volume to demonstrate this.



Pool: Pool1

Total 273.00 GiB

Quantity: 1
Capacity: 1 GiB
Capacity savings: None
Name: basic_1_pool_1

I/O group: Automatic

1 volume
Volume name: basic_1_pool_1

Total real capacity: 1.00 GiB
Total virtual capacity: 1.00 GiB

*Figure 7-11 Quick Volume Creation submenu*

**Notes:**

- A *Basic* volume is a volume whose data is striped across all available managed disks (MDisks) in one storage pool.
- A *Mirrored* volume is a volume with two physical copies, where each volume copy can belong to a different storage pool.
- A *Custom* volume, in the context of this menu, is either a *Basic* or *Mirrored* volume with customization from default parameters

***Quick Volume Creation*** also provides, using the **`Capacity Savings`** parameter, the ability to change the default provisioning of a *Basic* or *Mirrored* Volume to Thin-provisioned or Compressed. For more information, see “Quick Volume Creation with Capacity Saving options” on page 233.

**Note:** Volume migration is described in 7.8, “Migrating a volume to another storage pool” on page 251. Creating volume copies is described in 7.3.2, “Creating Mirrored volumes using Quick Volume Creation” on page 231.

# 7.3 Creating volumes using the Quick Volume Creation

This section focuses on using the Quick Volume Creation menu to create *Basic* and *Mirrored* volumes in a system with standard topology. It also covers creating host-to-volume mapping. As previously stated, **Quick Volume Creation** is available on four different volume classes:

- Basic
- Mirrored
- HyperSwap

> **Note:** The ability to create HyperSwap volumes using the GUI simplifies creation and configuration. This simplification is enhanced by the GUI using the command: **`mkvolume`**.

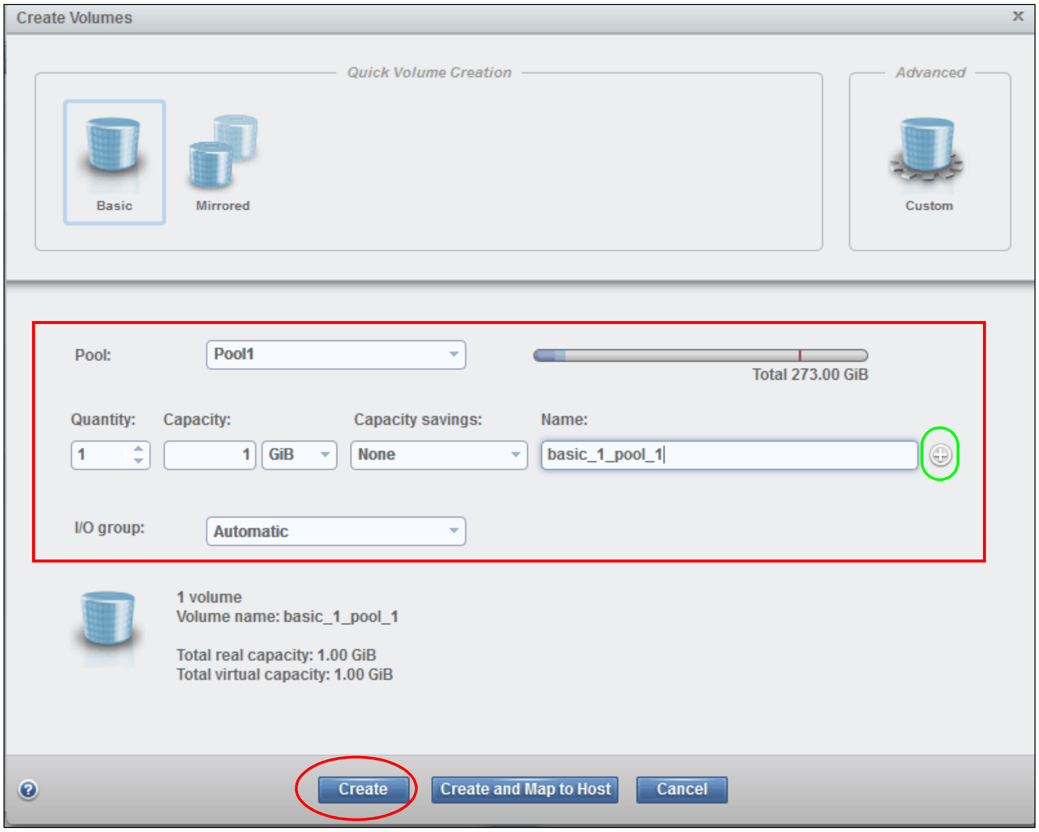
## 7.3.1 Creating Basic volumes using Quick Volume Creation

The most commonly used type of volume is the *Basic* volume. This type of volume is fully provisioned, with the entire size dedicated to the defined volume. The host and the IBM Spectrum Virtualize system see the fully allocated space.

Create a *Basic* volume by clicking the *Basic* icon as shown in Figure 7-9 on page 226. This opens an additional input window where you can define the following information:

- Pool: The Pool in which the volume will be created (drop-down)
- Quantity: Number of volumes to be created (numeric up/down)
- Capacity: Size of the volume in Units (drop-down)
- Capacity Savings (drop-down):
  - None
  - Thin-provisioned
  - Compressed
- Name: Name of the Volume (cannot start with a numeric)
- I/O group

*Basic* Volume creation process is shown in Figure 7-12.



Pool: Pool1

Total 273.00 GiB

Quantity: 1

Capacity: 1 GiB

Capacity savings: None

Name: basic_1_pool_1

I/O group: Automatic

1 volume
Volume name: basic_1_pool_1

Total real capacity: 1.00 GiB
Total virtual capacity: 1.00 GiB

*Figure 7-12 Creating Basic volume*

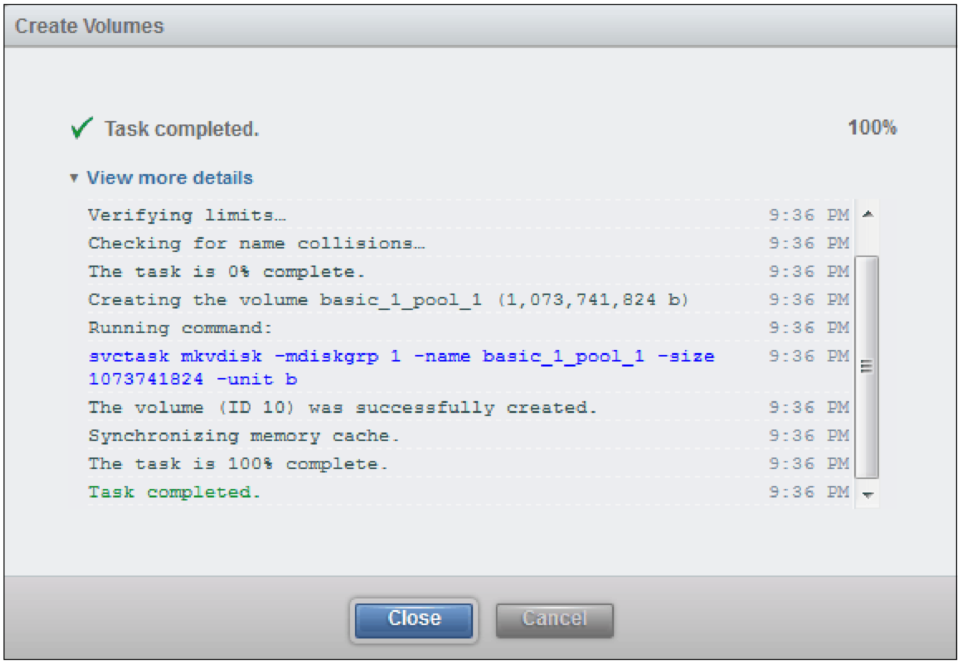
An appropriate naming convention is recommended for volumes for easy identification of their association with the host or host cluster. At a minimum, it should contain the name of the pool or some tag that identifies the underlying storage subsystem. It can also contain the host name that the volume will be mapped to, or perhaps the content of this volume, for example, name of applications to be installed.

When all of the characteristics of the *Basic* volume have been defined, it can be created by selecting one of the following options:

- Create
- Create and Map to Host

**Note:** The Plus sign (+) icon highlighted in green as shown in Figure 7-12, can be used to create more volumes in the same instance of the volume creation wizard.

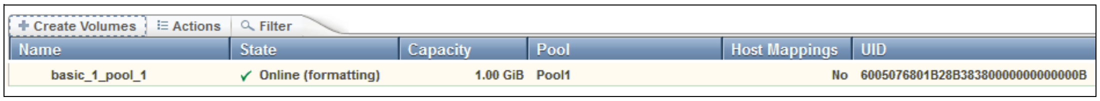
In this example, **Create** option has been selected (the volume-to-host mapping can be performed later). At the end of the volume creation, following confirmation window will appear as shown in Figure 7-13.



```
Verifying limits…                                                  9:36 PM
Checking for name collisions…                                      9:36 PM
The task is 0% complete.                                           9:36 PM
Creating the volume basic_1_pool_1 (1,073,741,824 b)               9:36 PM
Running command:                                                   9:36 PM
svctask mkvdisk -mdiskgrp 1 -name basic_1_pool_1 -size             9:36 PM
1073741824 -unit b
The volume (ID 10) was successfully created.                       9:36 PM
Synchronizing memory cache.                                        9:36 PM
The task is 100% complete.                                         9:36 PM
Task completed.                                                    9:36 PM
```

*Figure 7-13 Create Volume Task Completion window: Success*

Success is also indicated by the state of the *Basic* volume being reported as formatting in the **Volumes** pane as shown in Figure 7-14.



| Name | State | Capacity | Pool | Host Mappings | UID |
|---|---|---|---|---|---|
| basic_1_pool_1 | ✓ Online (formatting) | 1.00 GiB | Pool1 | No | 6005076801B28B38380000000000000B |

*Figure 7-14 Basic Volume Fast-Format*

**Notes:**

- Fully allocated volumes are automatically formatted through the quick initialization process after the volume is created. This process makes fully allocated volumes available for use immediately.
- Quick initialization requires a small amount of I/O to complete, and limits the number of volumes that can be initialized at the same time. Some volume actions, such as moving, expanding, shrinking, or adding a volume copy, are disabled when the specified volume is initializing. Those actions are available after the initialization process completes.
- The quick initialization process can be disabled in circumstances where it is not necessary. For example, if the volume is the target of a Copy Services function, the Copy Services operation formats the volume.The quick initialization process can also be disabled for performance testing, so that the measurements of the raw system capabilities can take place without waiting for the process to complete.

For more information, see the Fully allocated volumes topic in the IBM Knowledge Center:

`https://ibm.biz/BdsKht`

### 7.3.2 Creating Mirrored volumes using Quick Volume Creation

IBM Spectrum Virtualize offers the capability to mirror volumes, which means a single volume, presented to a host, can have two physical copies. Each volume copy can belong to a different pool, and each copy has the same virtual capacity as the volume. When a server writes to a mirrored volume, the system writes the data to both copies. When a server reads a mirrored volume, the system picks one of the copies to read.

Normally this is the primary copy (as indicated in the management GUI by an asterisk (*)). If one of the mirrored volume copies is temporarily unavailable (for example, because the storage system that provides the pool is unavailable), the volume remains accessible to servers. The system remembers which areas of the volume are written and resynchronizes these areas when both copies are available.

The use of mirrored volumes results in the following outcomes:

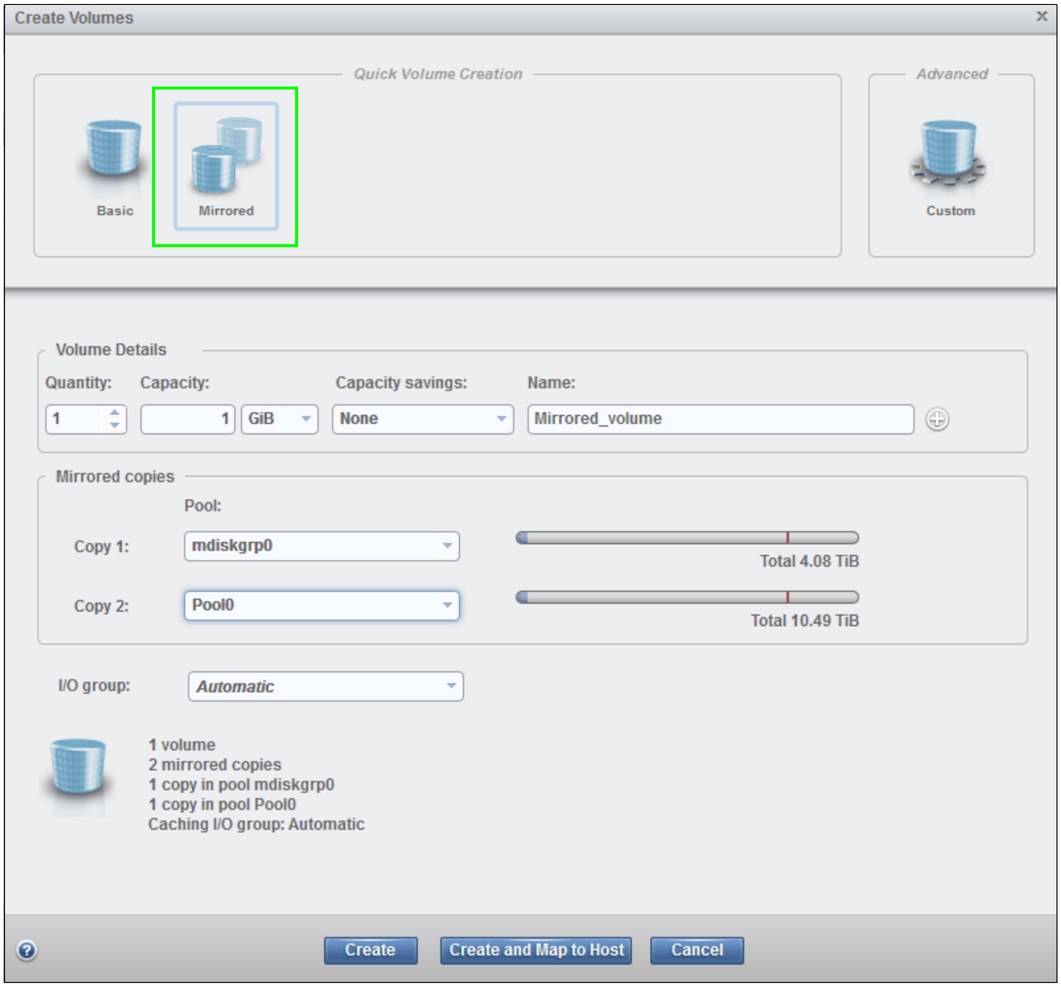
- Improves availability of volumes by protecting them from a single storage system failure
- Provides concurrent maintenance of a storage system that does not natively support concurrent maintenance
- Provides an alternative method of data migration with better availability characteristics
- Converts between fully allocated volumes and thin-provisioned volumes

**Note:** Volume mirroring is not a true disaster recovery (DR) solution, because both copies are accessed by the same node pair and addressable by only a single cluster, but it can improve availability.

To create a mirrored volume, complete the following steps:

1. In the Create Volumes window, click **Mirrored** as shown inFigure 7-15 on page 232 and enter the **Volume Details: Quantity**, **Capacity**, **Capacity savings**, and **Name**. Next, in the **Mirrored copies** subsection, choose the **Pool** of **Copy1** and **Copy2** using the drop-down menu. Although the mirrored volume can be created in the same pool, this is not typical.

   We suggest that you keep mirrored volumes on a separate set of physical disks (Pools). Leave the **I/O group** option at its default setting of **Automatic** (see Figure 7-15).



Volume Details

Quantity: 1
Capacity: 1 GiB
Capacity savings: None
Name: Mirrored_volume

Mirrored copies

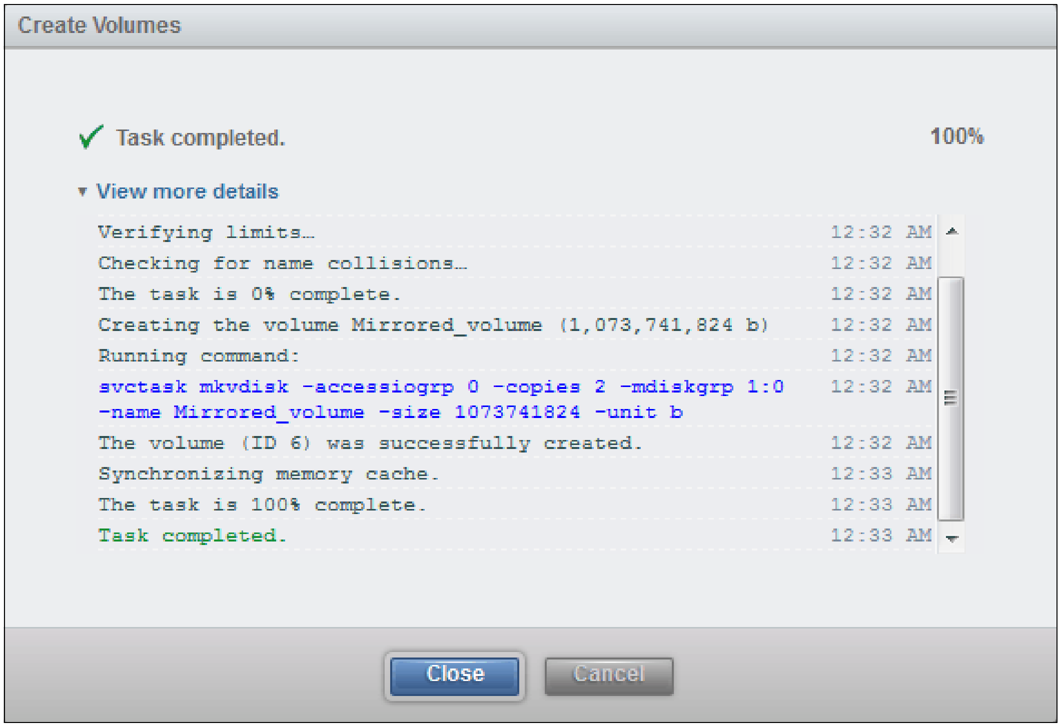
Pool:
Copy 1: mdiskgrp0 Total 4.08 TiB
Copy 2: Pool0 Total 10.49 TiB

I/O group: *Automatic*

1 volume
2 mirrored copies
1 copy in pool mdiskgrp0
1 copy in pool Pool0
Caching I/O group: Automatic

*Figure 7-15 Mirrored Volume creation*

2. Click **Create** (or **Create and Map to Host**)
3. Next, the GUI displays the underlying CLI commands being run to create the mirrored volume and indicates completion as shown in Figure 7-16.



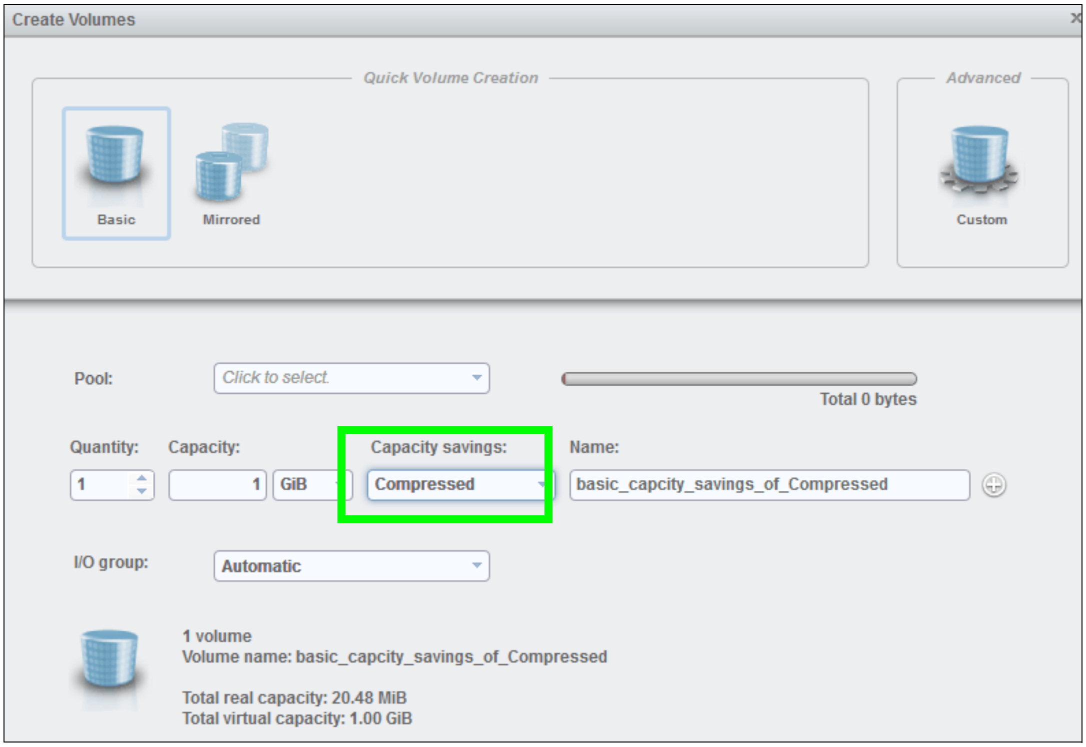
```
Verifying limits…                                                    12:32 AM
Checking for name collisions…                                        12:32 AM
The task is 0% complete.                                             12:32 AM
Creating the volume Mirrored_volume (1,073,741,824 b)                12:32 AM
Running command:                                                     12:32 AM
svctask mkvdisk -accessiogrp 0 -copies 2 -mdiskgrp 1:0               12:32 AM
-name Mirrored_volume -size 1073741824 -unit b
The volume (ID 6) was successfully created.                          12:32 AM
Synchronizing memory cache.                                          12:33 AM
The task is 100% complete.                                           12:33 AM
Task completed.                                                      12:33 AM
```

*Figure 7-16 Task complete - created Mirrored Volume*

> **Note:** When creating a Mirrored Volume using this menu, you are not required to specify the Mirrored Sync rate. It defaults to 2 MB/s. Customization of this synchronization rate can be achieved using the **Advanced** → **Custom** menu.

### Quick Volume Creation with Capacity Saving options

The Quick Volume Creation menu also provides, using the **`Capacity Savings`** parameter, the ability to alter the provisioning of a *Basic* or *Mirrored* volume into Thin-provisioned or Compressed. This is achieved by selecting either Thin-provisioned or Compressed from the drop-down menu as shown in Figure 7-17.



Pool: Click to select.

Total 0 bytes

Quantity: 1

Capacity: 1 GiB

Capacity savings: Compressed

Name: basic_capcity_savings_of_Compressed

I/O group: Automatic

1 volume
Volume name: basic_capcity_savings_of_Compressed

Total real capacity: 20.48 MiB
Total virtual capacity: 1.00 GiB

*Figure 7-17 Quick Volume Creation with Capacity Saving option set to Compressed*

Alternatively, select Thin-provisioned from the menu to define a Thin-provisioned volume.

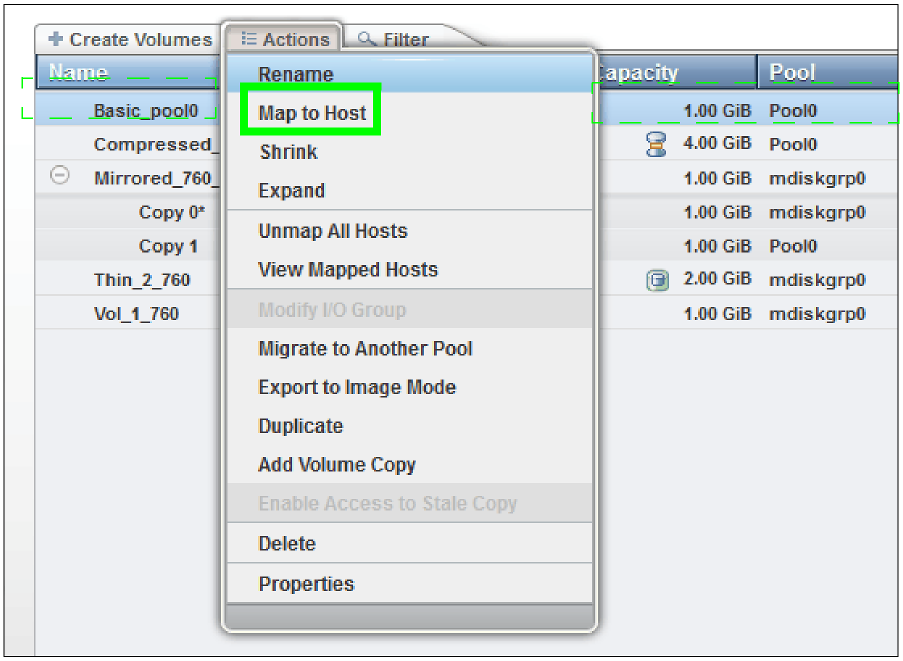
## 7.4 Mapping a volume to the host

After a volume is created, it can be mapped to a host:

1. From the Volumes menu, highlight the volume that you want to create a mapping for and then select **Actions** from the menu bar.

> **Tip:** An alternative way of opening the **Actions** menu is to highlight (select) a volume and use the right mouse button.

2. From the Actions menu, select the **Map to Host** option as shown in Figure 7-18.



| Name | Capacity | Pool |
|---|---|---|
| Basic_pool0 | 1.00 GiB | Pool0 |
| Compressed_ | 4.00 GiB | Pool0 |
| Mirrored_760_ | 1.00 GiB | mdiskgrp0 |
| Copy 0* | 1.00 GiB | mdiskgrp0 |
| Copy 1 | 1.00 GiB | Pool0 |
| Thin_2_760 | 2.00 GiB | mdiskgrp0 |
| Vol_1_760 | 1.00 GiB | mdiskgrp0 |

*Figure 7-18 Map to Host*

3. This opens a Map to Host window. In this window, use the **Select the Host** menu to select a host to map the volume to as shown in Figure 7-19.

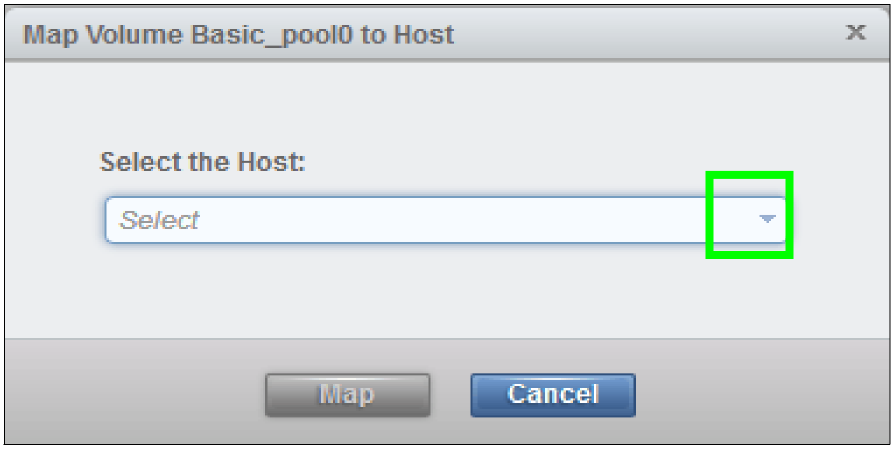


*Figure 7-19 Mapping a Volume to Host*

4. Select the Host from the drop-down, select the check box for your choice, and then click **Map** as shown in Figure 7-20.

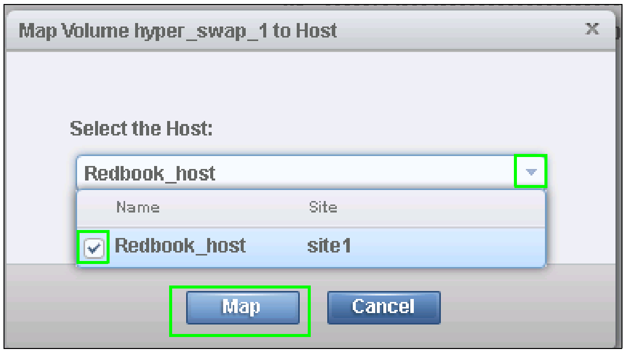
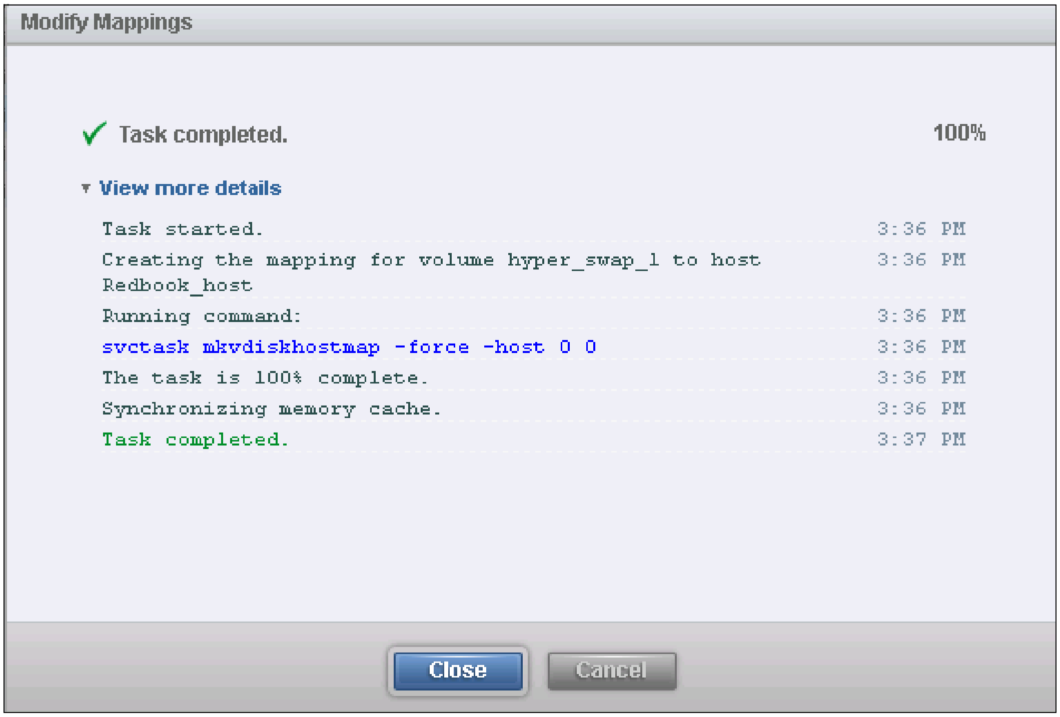


Figure 7-20 Selected host from the drop-down

5. The Modify Mappings window displays the command details, and then a Task completed message as shown in Figure 7-21.



✓ Task completed. 100%

```
Task started.                                               3:36 PM
Creating the mapping for volume hyper_swap_1 to host        3:36 PM
Redbook_host
Running command:                                            3:36 PM
svctask mkvdiskhostmap -force -host 0 0                     3:36 PM
The task is 100% complete.                                  3:36 PM
Synchronizing memory cache.                                 3:36 PM
Task completed.                                             3:37 PM
```

Figure 7-21 Successful completion of Host to Volume mapping

## 7.5 Creating Custom volumes using the Advanced menu

The Advanced menu of the Create Volumes window enables *Custom* volume creation. It provides an alternative method of defining Capacity savings options, such as Thin-provisioning and Compression, but also expands on the base level default options for available *Basic* and *Mirrored* volumes. A *Custom* volume can be customized regarding Mirror sync rate, Cache mode, and Fast-Format.

The **Advanced** menu consists of several submenus:

- Volume Details (Mandatory - defines the *Capacity savings* option)
- Volume Location (Mandatory - defines Pools to be used)
- Thin Provisioning

- Compression
- General - for changing default options for Cache mode and Formatting
- Summary

Work through these submenus to customize your *Custom* volume as wanted, and then commit these changes using **Create** as shown in Figure 7-22.

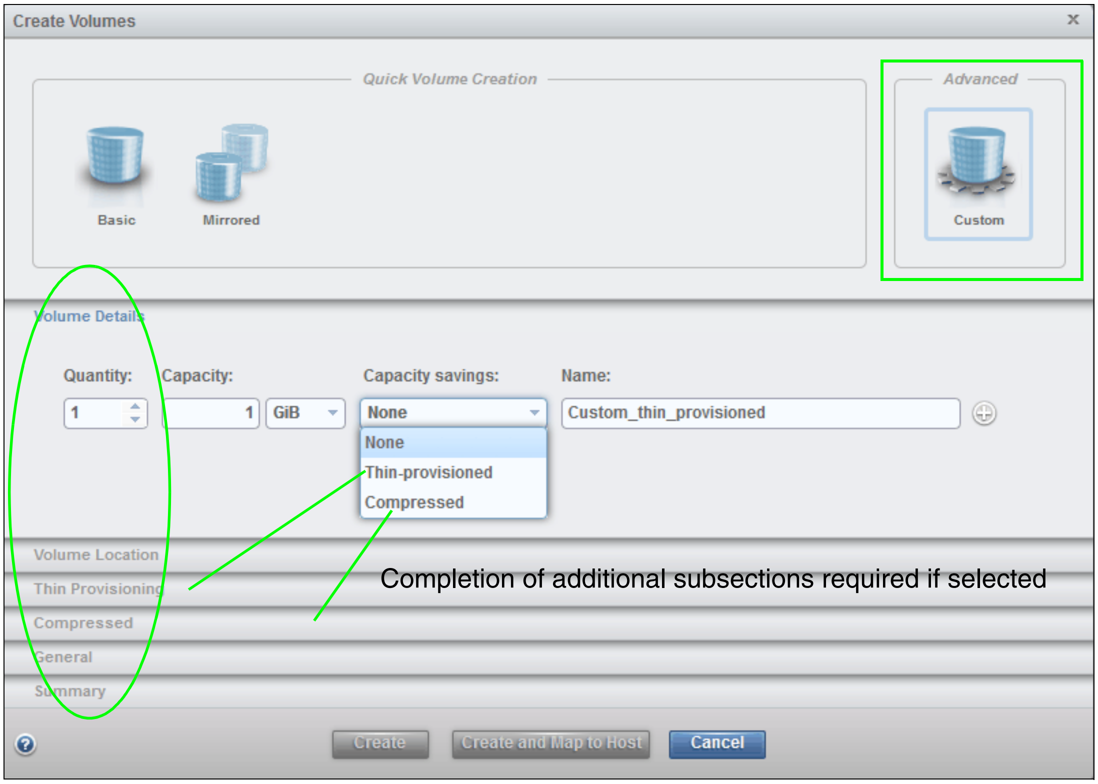


*Figure 7-22 Customization submenus*

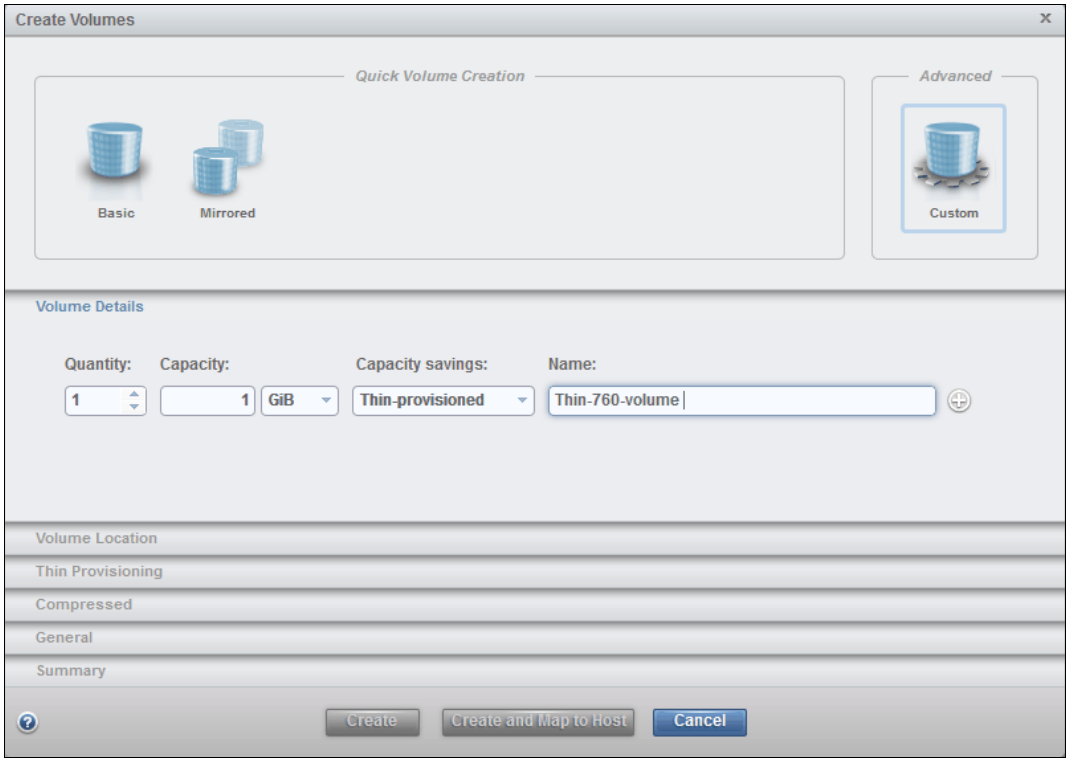
## 7.5.1 Creating a custom thin-provisioned volume

A thin-provisioned volume can be defined and created using the Advanced menu. Regarding application reads and writes, thin-provisioned volumes behave as though they were fully allocated. When creating a thin-provisioned volume, you can specify two capacities:

- The real physical capacity allocated to the volume from the storage pool. The real capacity determines the quantity of extents that are initially allocated to the volume.
- Its virtual capacity available to the host. The virtual capacity is the capacity of the volume that is reported to all other components (for example, FlashCopy, cache, and remote copy) and to the hosts.

To create a thin-provisioned volume, complete the following steps:

1. From the Create Volumes window, select the **Advanced** option. This opens the subsection **Volume Details** where you can input the **Quantity**, **Capacity** (virtual), **Capacity Savings** (choose **Thin-provisioned** from the drop-down), and **Name** of the volume being created (Figure 7-23).



Quantity: 1

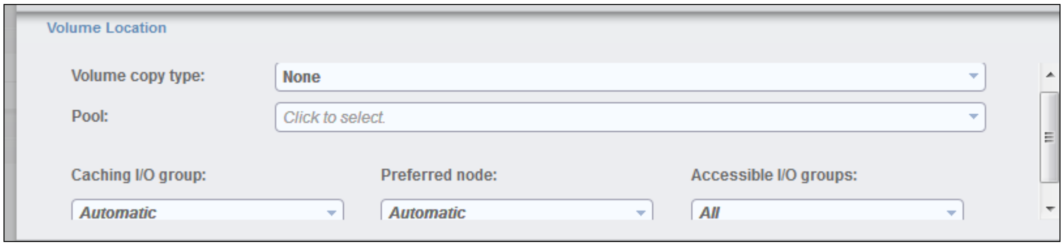
Capacity: 1 GiB

Capacity savings: Thin-provisioned

Name: Thin-760-volume

*Figure 7-23 Create a thin-provisioned volume*

2. Next, click the **Volume Location** subsection to define the pool in which the volume will be created. Use the drop-down in the **Pool** option to choose the pool. All other options, such as **Volume copy type, Caching I/O group, Preferred node,** and **Accessible I/O groups**, can be left with their default options (Figure 7-24).



Volume copy type: None

Pool: Click to select

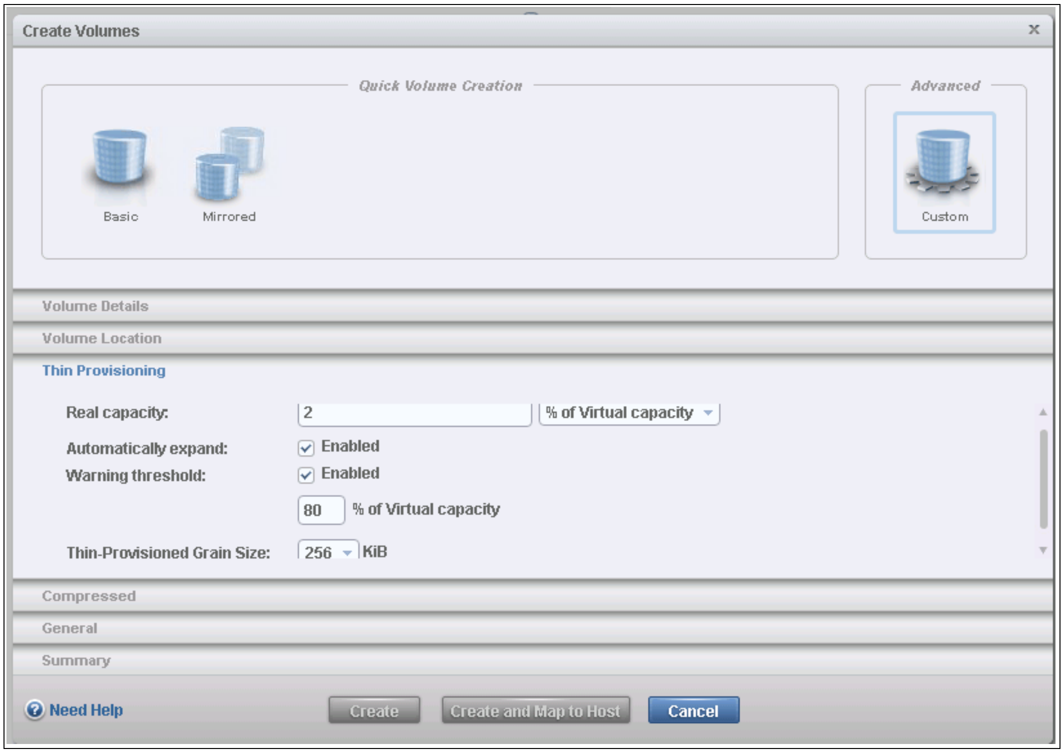
Caching I/O group: Automatic

Preferred node: Automatic

Accessible I/O groups: All

*Figure 7-24 Volume Location*

3. Next, click the **Thin Provisioning** subsection to manage the real and virtual capacity, expansion criteria, and grain size (Figure 7-25).



Real capacity: 2 % of Virtual capacity

Automatically expand: ☑ Enabled

Warning threshold: ☑ Enabled

80 % of Virtual capacity

Thin-Provisioned Grain Size: 256 KiB

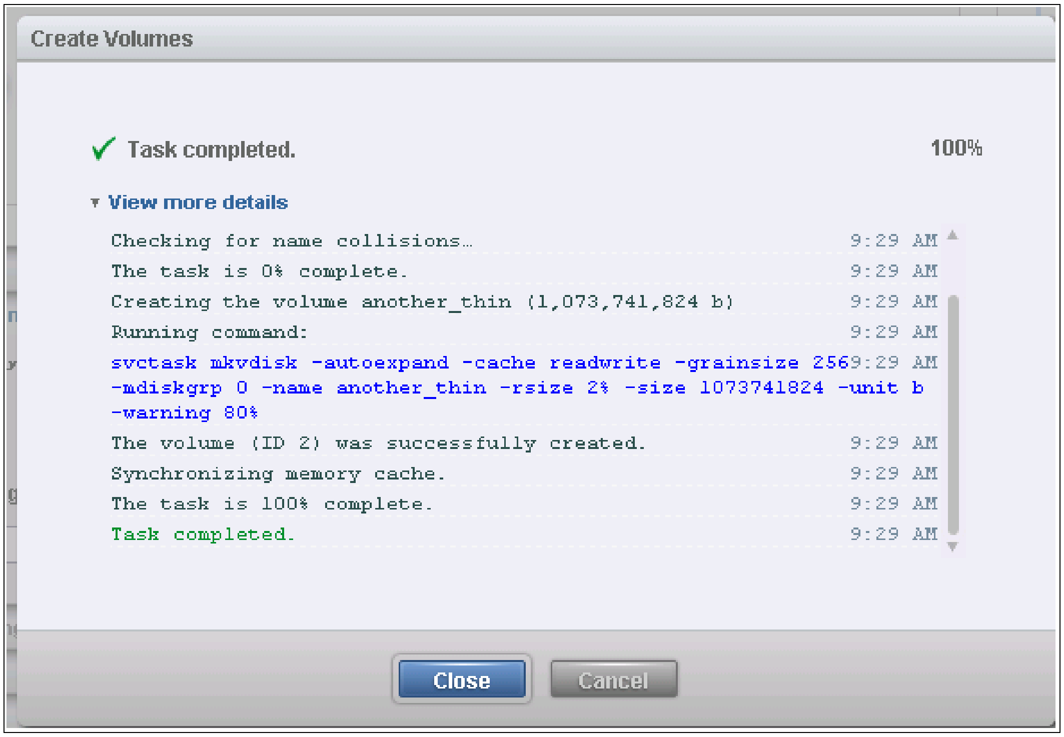
*Figure 7-25 Thin Provisioning*

The Thin Provisioning options are as follows (defaults displayed in brackets):

- **Real capacity**: (2%). Specify the size of the real capacity space used during creation.
- **Automatically Expand**: (Enabled). This option enables the automatic expansion of real capacity, if more capacity is to be allocated.
- **Warning threshold:** (Enabled). Enter a threshold for receiving capacity alerts.
- **Grain Size**: 256 kibibytes (KiB). Specify the grain size for real capacity. This describes the size of chunk of storage to be added to used capacity. For example, when the host writes 1 megabyte (MB) of new data, the capacity is increased by adding four chunks of 256 KiB each.

> **Important:** If you do not use the **`autoexpand`** feature, the volume will go offline after reaching real capacity.
>
> The default grain size is 256 KiB. The optimum choice of grain size is dependent upon volume use type. For more information, see the “Performance Problem When Using EasyTier With Thin Provisioned Volumes” topic:
>
> `http://www.ibm.com/support/docview.wss?uid=ssg1S1003982`
>
> - If you are *not* going to use the thin-provisioned volume as a FlashCopy source or target volume, use 256 KiB to maximize performance.
> - If you *are* going to use the thin-provisioned volume as a FlashCopy source or target volume, specify the same grain size for the volume and for the FlashCopy function.

4. Apply all required changes and click **Create** to define the volume (Figure 7-26).



```
Checking for name collisions…                                  9:29 AM
The task is 0% complete.                                       9:29 AM
Creating the volume another_thin (1,073,741,824 b)             9:29 AM
Running command:                                               9:29 AM
svctask mkvdisk -autoexpand -cache readwrite -grainsize 2569:29 AM
-mdiskgrp 0 -name another_thin -rsize 2% -size 1073741824 -unit b
-warning 80%
The volume (ID 2) was successfully created.                    9:29 AM
Synchronizing memory cache.                                    9:29 AM
The task is 100% complete.                                     9:29 AM
Task completed.                                                9:29 AM
```

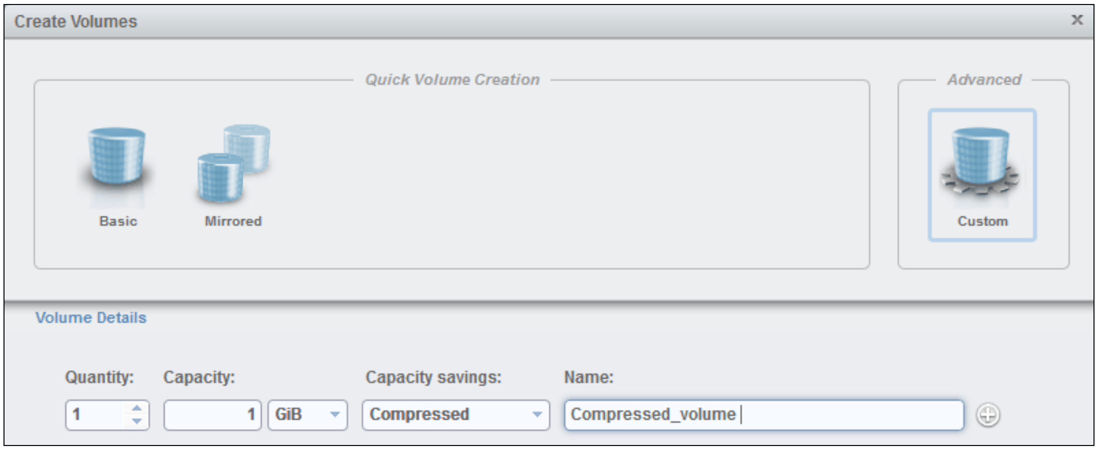
*Figure 7-26 Task completed, the thin-provisioned volume is created*

5. Again, you can directly start the wizard for mapping this volume to the host by clicking **Create and Map to Host**.

## 7.5.2 Creating Custom Compressed volumes

The configuration of compressed volumes is similar to thin-provisioned volumes.To create a Compressed volume, complete the following steps:

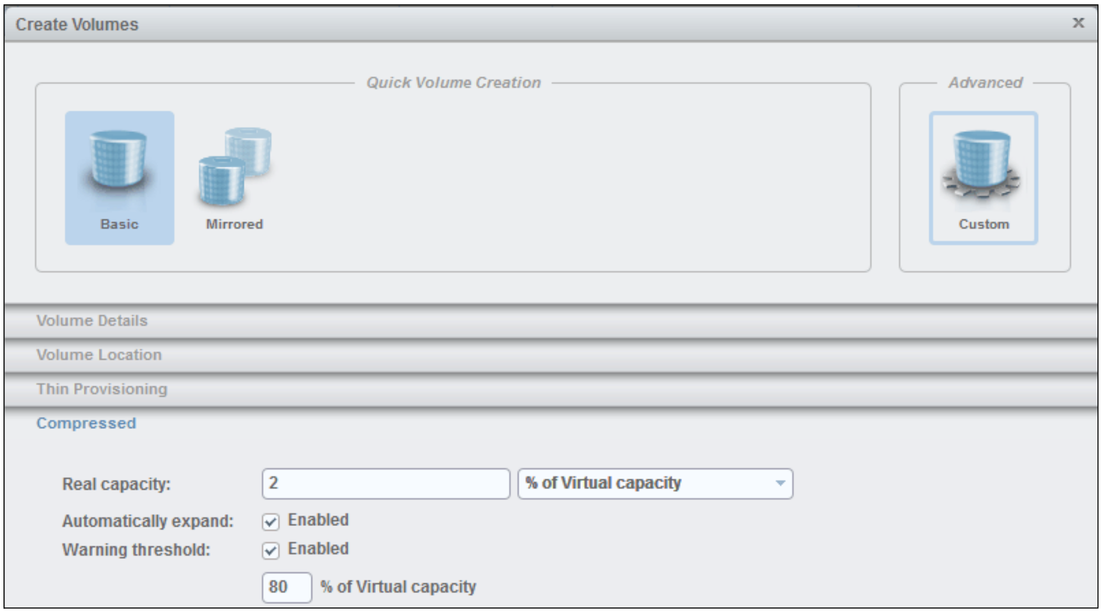
1. From the Create Volumes window, select the **Advanced** option. This opens the subsection **Volume Details** where **Quantity**, **Capacity** (virtual), **Capacity Savings** (choose **Compressed** from the drop-down), and **Name** can be input (Figure 7-27).



Quantity: 1
Capacity: 1 GiB
Capacity savings: Compressed
Name: Compressed_volume

*Figure 7-27 Defining a volume as compressed using the Capacity savings option*

2. Open the **Volume Location** subsection and select, from the drop-down, the Pool in which the compressed volume will be created (use all other parameter defaults as defined).
3. Open the **Compression** subsection and verify that **Real Capacity** is set to a minimum of the default value of 2% (use all other parameter defaults as defined). See Figure 7-28.



Real capacity: 2 % of Virtual capacity

Automatically expand: ☑ Enabled

Warning threshold: ☑ Enabled

80 % of Virtual capacity

*Figure 7-28 Checking Compressed volume Custom / Advanced settings*

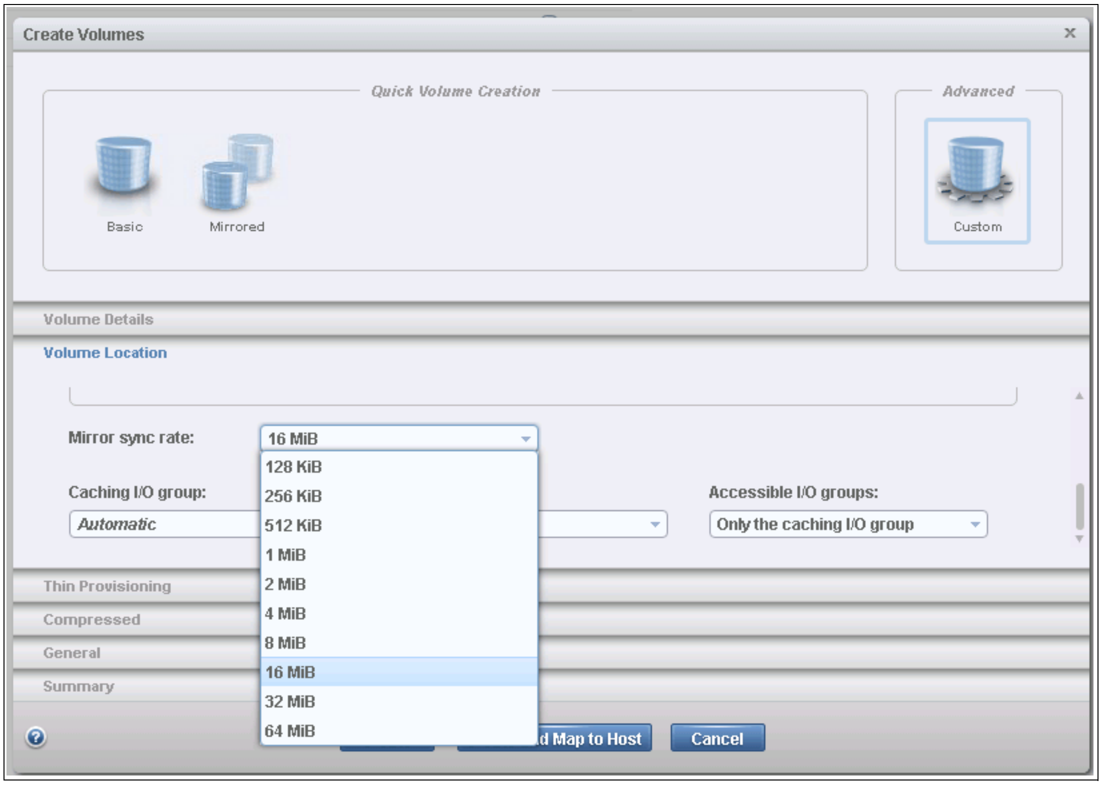
4. Confirm and commit the selection by clicking **Create**.

## 7.5.3 Custom Mirrored Volumes

The **Advanced** option in the Create Volumes window is used to customize volume creation. Using this feature, the default options can be overridden and volume creation can be tailored to the specifics of the clients environment.

### Modifying the Mirror sync rate

The **Mirror sync rate** can be changed from the default setting using the **Advanced** option, subsection **Volume Location**, of the Create Volumes window. This option sets the priority of copy synchronization progress, enabling a preferential rate to be set for more important volumes (Figure 7-29).



Mirror sync rate: 16 MiB

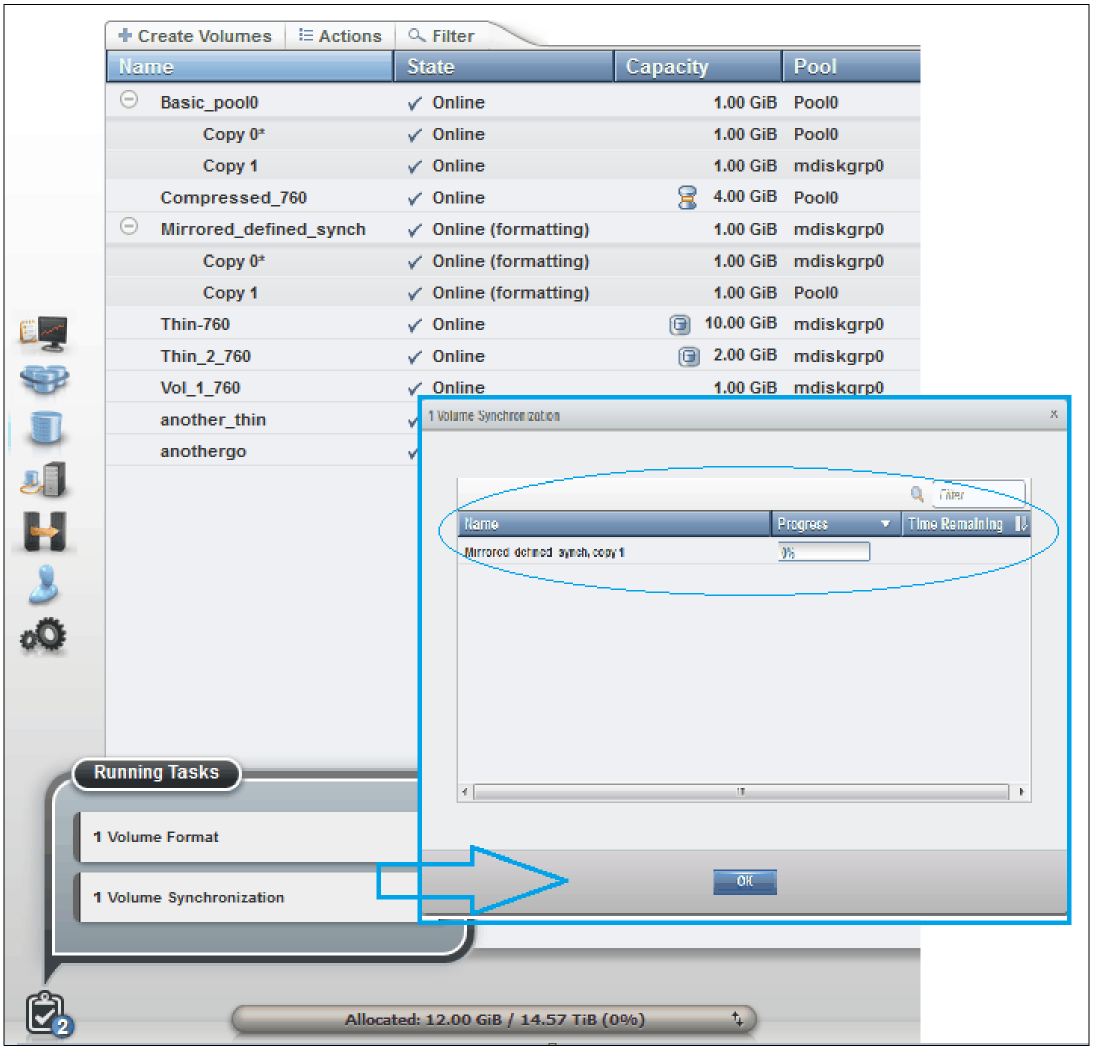
128 KiB
256 KiB
512 KiB
1 MiB
2 MiB
4 MiB
8 MiB
16 MiB
32 MiB
64 MiB

Caching I/O group: *Automatic*

Accessible I/O groups: Only the caching I/O group

*Figure 7-29 Customization of Mirrored sync rate*

The progress of formatting and synchronization of a newly created Mirrored Volume can be checked from the **Running Tasks** menu. This menu reports the progress of all currently running tasks; including **Volume Format** and **Volume Synchronization** (Figure 7-30).



| Name | State | Capacity | Pool |
|---|---|---|---|
| Basic_pool0 | ✓ Online | 1.00 GiB | Pool0 |
| Copy 0* | ✓ Online | 1.00 GiB | Pool0 |
| Copy 1 | ✓ Online | 1.00 GiB | mdiskgrp0 |
| Compressed_760 | ✓ Online | 4.00 GiB | Pool0 |
| Mirrored_defined_synch | ✓ Online (formatting) | 1.00 GiB | mdiskgrp0 |
| Copy 0* | ✓ Online (formatting) | 1.00 GiB | mdiskgrp0 |
| Copy 1 | ✓ Online (formatting) | 1.00 GiB | Pool0 |
| Thin-760 | ✓ Online | 10.00 GiB | mdiskgrp0 |
| Thin_2_760 | ✓ Online | 2.00 GiB | mdiskgrp0 |
| Vol_1_760 | ✓ Online | 1.00 GiB | mdiskgrp0 |
| another_thin | ✓ | | |
| anothergo | ✓ | | |

| Name | Progress | Time Remaining |
|---|---|---|
| Mirrored_defined_synch, copy 1 | 0% | |

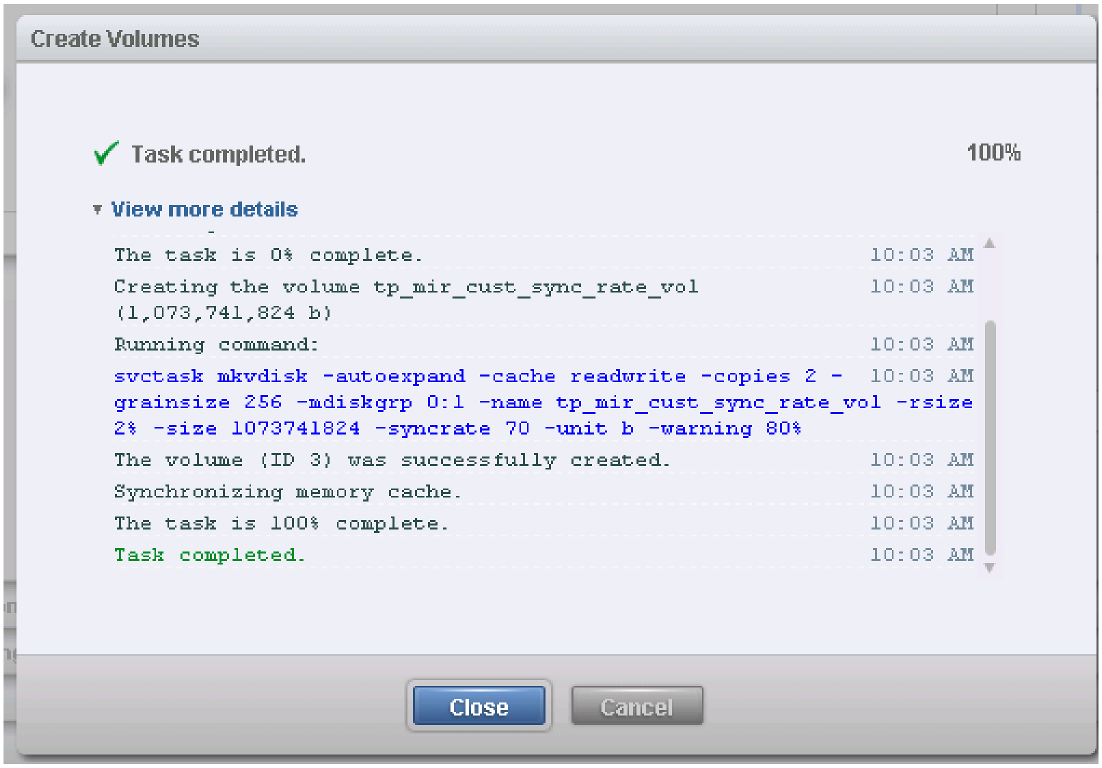
*Figure 7-30*

## Creating a Custom Thin-provisioned Mirrored volume

The **Advanced** option in the Create Volumes window is used to customize volume creation. Using this feature, the default options can be overridden and volume creation can be tailored to the specifics of the clients environment.

The **Mirror Sync rate** can be changed from the default setting using the **Advanced** option, subsection **Volume Location**, of the Create Volume window. This option sets the priority of copy synchronization progress, enabling a preferential rate to be set for more important mirrored volumes.

The summary shows you the capacity information and the allocated space. You can click **Advanced** and customize the thin-provision settings or the mirror synchronization rate. After you create the volume, the confirmation window opens as shown in Figure 7-31.



```
The task is 0% complete.                                    10:03 AM
Creating the volume tp_mir_cust_sync_rate_vol               10:03 AM
(1,073,741,824 b)
Running command:                                            10:03 AM
svctask mkvdisk -autoexpand -cache readwrite -copies 2 -    10:03 AM
grainsize 256 -mdiskgrp 0:1 -name tp_mir_cust_sync_rate_vol -rsize
2% -size 1073741824 -syncrate 70 -unit b -warning 80%
The volume (ID 3) was successfully created.                 10:03 AM
Synchronizing memory cache.                                 10:03 AM
The task is 100% complete.                                  10:03 AM
Task completed.                                             10:03 AM
```

*Figure 7-31 Confirmation window*

The initial synchronization of thin-mirrored volumes is fast when a small amount of real and virtual capacity is used.

## 7.6 HyperSwap and the mkvolume command

HyperSwap volume configuration is not possible until site awareness has been configured.

When the HyperSwap topology is configured, the GUI uses **mkvolume** command to create volumes instead of traditional **mkvdisk** command. This section describes the **mkvolume** command that is used in HyperSwap topology. The GUI continues to use **mkvdisk** when all other classes of volumes are created.

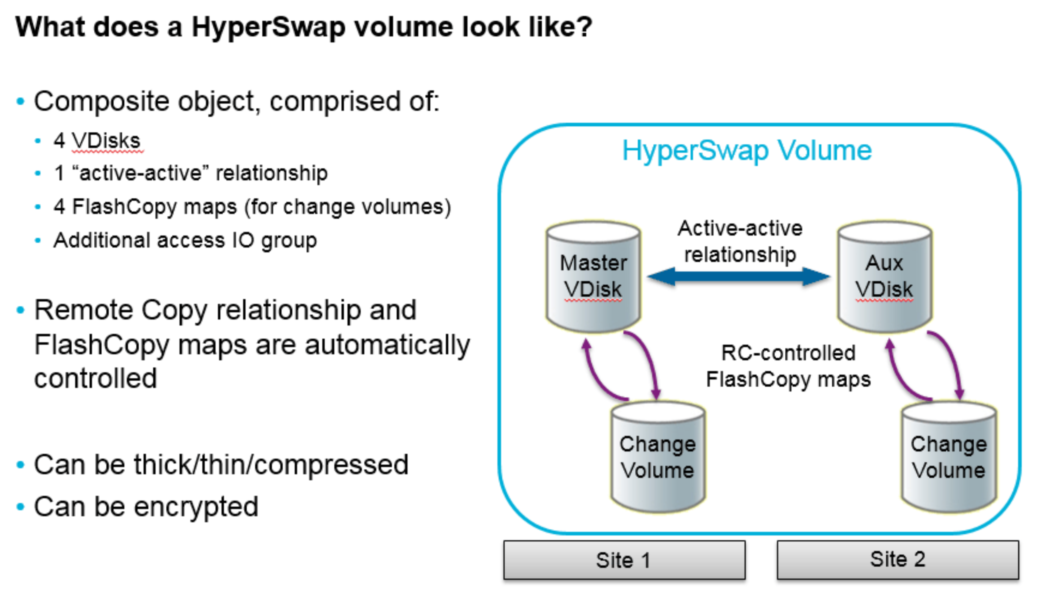
In this section, new **mkvolume** command is described, and how the GUI uses this command, when HyperSwap topology has been configured, rather than the “traditional” **mkvdisk** command.

> **Note:** It is still possible to create HyperSwap volumes as in the V7.5 release, as described in the following white paper:
>
> https://www.ibm.com/support/techdocs/atsmastr.nsf/WebIndex/WP102538
>
> You can also get more information in the IBM Redbooks publication *IBM Storwize V7000, Spectrum Virtualize, HyperSwap, and VMware Implementation*, SG24-8317:
>
> https://www.redbooks.ibm.com/redbooks.nsf/RedpieceAbstracts/sg248317.html

*HyperSwap* volumes are a new type of HA volumes to be supported by IBM Spectrum Virtualize. They are built off two existing IBM Spectrum Virtualize technologies:

- Metro Mirror
- (VDisk) Volume Mirroring

These technologies have been combined in an active-active configuration deployed using Change Volumes (as used in the Global Mirror with Change Volumes) to create a Single Volume (from a host perspective) in an HA form. The volume presented is a combination of 4x “traditional” volumes, but is a single entity from a host (and administrative) perspective a as shown in Figure 7-32.



**What does a HyperSwap volume look like?**

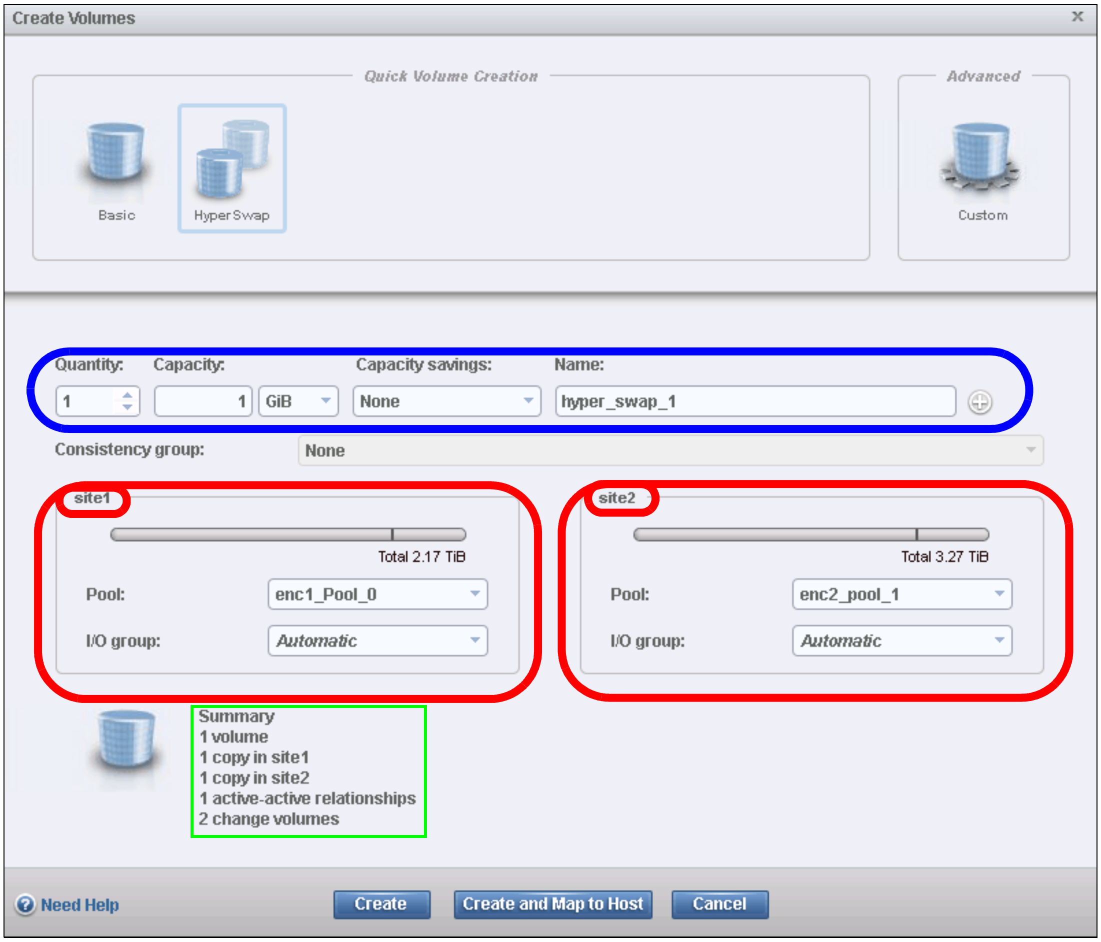
- Composite object, comprised of:
  - 4 VDisks
  - 1 “active-active” relationship
  - 4 FlashCopy maps (for change volumes)
  - Additional access IO group
- Remote Copy relationship and FlashCopy maps are automatically controlled
- Can be thick/thin/compressed
- Can be encrypted

*Figure 7-32 What makes up a HyperSwap Volume*

The GUI simplifies the complexity of HyperSwap volume creation by only presenting the volume class of *HyperSwap* as a **Quick Volume Creation** option after *HyperSwap* topology has been configured.

In the following example, *HyperSwap* topology has been configured and the Quick Volume Creation window is being used to define a *HyperSwap* Volume as shown in Figure 7-33 on page 246.

The *capacity* and *name* characteristics are defined as for a *Basic* volume (highlighted in blue in the example) and the mirroring characteristics are defined by the Site parameters (highlighted in red).



Quantity: 1
Capacity: 1 GiB
Capacity savings: None
Name: hyper_swap_1
Consistency group: None

site1
Total 2.17 TiB
Pool: enc1_Pool_0
I/O group: Automatic

site2
Total 3.27 TiB
Pool: enc2_pool_1
I/O group: Automatic

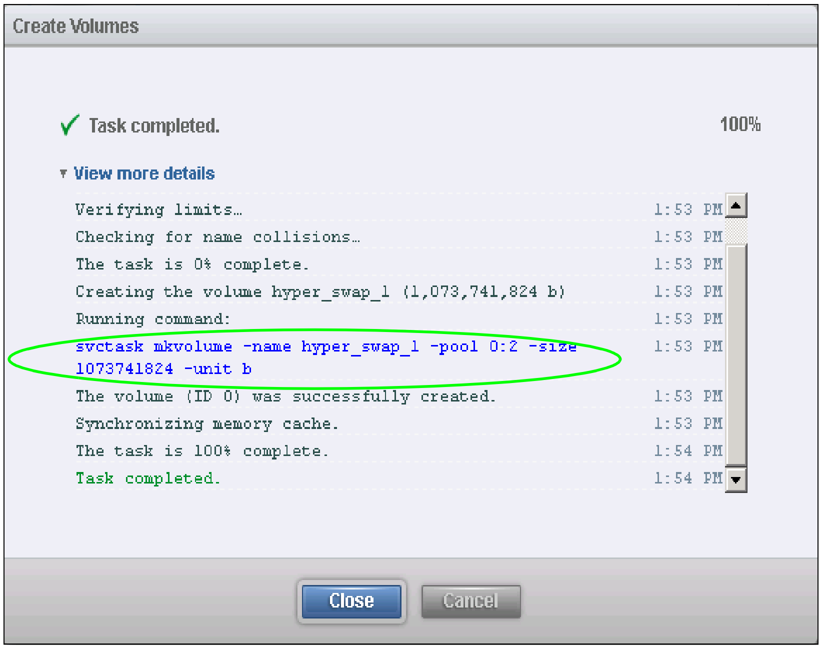
Summary
1 volume
1 copy in site1
1 copy in site2
1 active-active relationships
2 change volumes

*Figure 7-33 HyperSwap Volume creation with Summary of actions*

The drop-downs assist in creation, and the *Summary* (lower left of the creation window) indicates the actions that will be carried out when the **Create** option is selected. As shown in Figure 7-33, a single volume will be created, with volume copies in `site1` and `site2`. This volume will be in an active-active (Metro-Mirror) relationship with additional resilience provided by 2x *change* volumes.

The command run to create this volume is shown in Figure 7-34, and can be summarized as follows:

```
svctask mkvolume -name <name_of_volume> -pool <X:Y> -size <Size_of_volume> -unit
<units>
```



```
Verifying limits…                                          1:53 PM
Checking for name collisions…                              1:53 PM
The task is 0% complete.                                   1:53 PM
Creating the volume hyper_swap_1 (1,073,741,824 b)         1:53 PM
Running command:                                           1:53 PM
svctask mkvolume -name hyper_swap_1 -pool 0:2 -size        1:53 PM
1073741824 -unit b
The volume (ID 0) was successfully created.                1:53 PM
Synchronizing memory cache.                                1:53 PM
The task is 100% complete.                                 1:54 PM
Task completed.                                            1:54 PM
```

*Figure 7-34 Example mkvolume command*

With a single **mkvolume** command, HyperSwap volume is created. Previously (using IBM Spectrum Virtualize V7.5) this was only possible with careful planning and issuing multiple commands listed below:

1. **mkvdisk master_vdisk**
2. **mkvdisk aux_vdisk**
3. **mkvdisk master_change_volume**
4. **mkvdisk aux_change_volume**
5. **mkrcrelationship –activeactive**
6. **chrcrelationship -masterchange**
7. **chrcrelationship -auxchange**
8. **addvdiskacces**

## 7.6.1 Volume manipulation commands

Five CLI commands for administering Volumes were released in IBM Spectrum Virtualize V7.6, but the GUI continues to use existing commands, for all volume administration, with the exception of HyperSwap volume creation (**mkvolume**) and deletion (**rmvolume**).

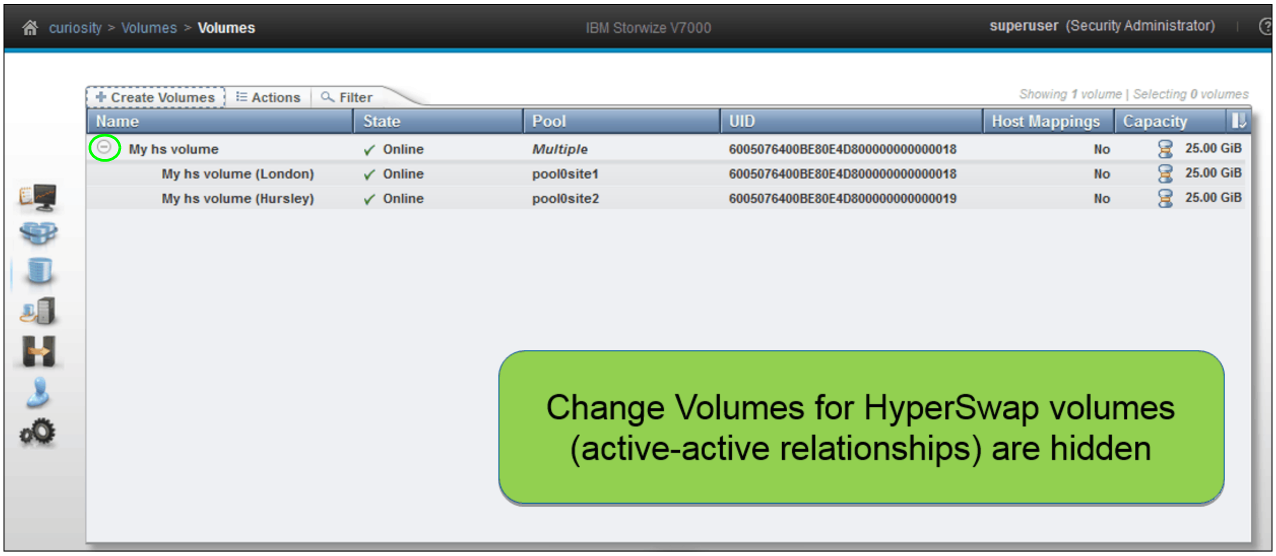
The five new CLI commands for administering Volumes are:

- **mkvolume**
- **mkimagevolume**
- **addvolumecopy**
- **rmvolumecopy**
- **rmvolume**

In addition, the following new **lsvdisk** and GUI functionality is available:

The **lsvdisk** command now includes volume_id, volume_name, and function fields to easily identify the individual VDisk that make up a HyperSwap volume. These views are “rolled-up” in the GUI to provide views that reflect the client’s view of the *HyperSwap* volume and its site-dependent copies, as opposed to the “low-level” VDisks and VDisk-change-volumes.

As shown in the Figure 7-35, **Volumes** → **Volumes** shows the *HyperSwap* Volume `My hs volume` with an expanded view opened using the Plus sign (+) to reveal 2x volume copies: `My hs volume (London)` (Master VDisk) and `My hs volume (Hursley)` (Auxiliary VDisk). We do not show the VDisk-Change-Volumes.

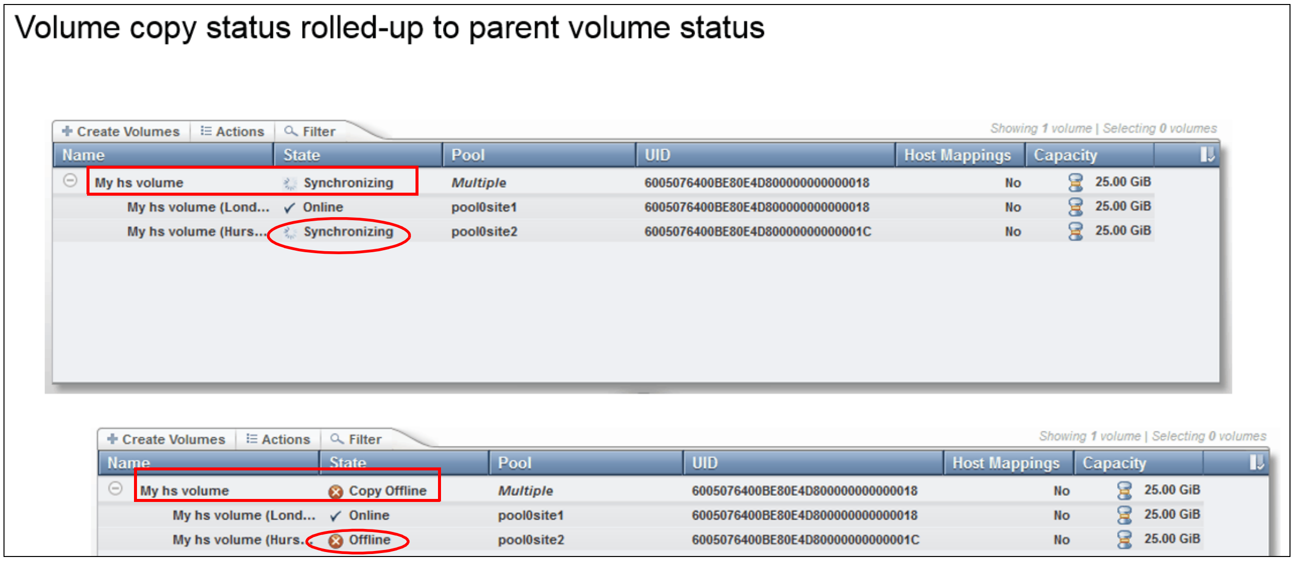

curiosity > Volumes > Volumes
IBM Storwize V7000
superuser (Security Administrator)
Create Volumes
Actions
Filter
Showing 1 volume | Selecting 0 volumes

| Name | State | Pool | UID | Host Mappings | Capacity |
|---|---|---|---|---|---|
| My hs volume | ✓ Online | *Multiple* | 6005076400BE80E4D800000000000018 | No | 25.00 GiB |
| My hs volume (London) | ✓ Online | pool0site1 | 6005076400BE80E4D800000000000018 | No | 25.00 GiB |
| My hs volume (Hursley) | ✓ Online | pool0site2 | 6005076400BE80E4D800000000000019 | No | 25.00 GiB |

Change Volumes for HyperSwap volumes (active-active relationships) are hidden

*Figure 7-35 Hidden Change Volumes*

Likewise, the status of the *HyperSwap* volume is reported at a “parent” level. If one of the copies is syncing or offline, the parent *HyperSwap* volume reflects this state as shown in Figure 7-36.


Volume copy status rolled-up to parent volume status

| Name | State | Pool | UID | Host Mappings | Capacity |
|---|---|---|---|---|---|
| My hs volume | Synchronizing | *Multiple* | 6005076400BE80E4D800000000000018 | No | 25.00 GiB |
| My hs volume (Lond... | ✓ Online | pool0site1 | 6005076400BE80E4D800000000000018 | No | 25.00 GiB |
| My hs volume (Hurs... | Synchronizing | pool0site2 | 6005076400BE80E4D80000000000001C | No | 25.00 GiB |

| Name | State | Pool | UID | Host Mappings | Capacity |
|---|---|---|---|---|---|
| My hs volume | ⊗ Copy Offline | *Multiple* | 6005076400BE80E4D800000000000018 | No | 25.00 GiB |
| My hs volume (Lond... | ✓ Online | pool0site1 | 6005076400BE80E4D800000000000018 | No | 25.00 GiB |
| My hs volume (Hurs... | ⊗ Offline | pool0site2 | 6005076400BE80E4D80000000000001C | No | 25.00 GiB |

*Figure 7-36 Parent volume reflects state of copy volume*

HyperSwap related Individual commands are briefly described here.

- **`mkvolume`**

  Create a new empty volume using storage from existing storage pools. The type of volume created is determined by the system topology and the number of storage pools specified. Volume is always formatted (zeroed). This command can be used to create:

  – Basic volume- any topology

  - Mirrored volume- standard topology
  - HyperSwap volume- HyperSwap topology
- **`rmvolume`**

  Remove a volume. For a HyperSwap volume, this includes deleting the active-active relationship and the change volumes.

  The **`-force`** parameter with **`rmvdisk`** is replaced by individual override parameters, making it clearer to the user exactly what protection they are bypassing.
- **`mkimagevolume`**

  Create a new image mode volume. This command can be used to import a volume, preserving existing data. Implemented as a separate command to provide greater differentiation between the action of creating a new empty volume and creating a volume by importing data on an existing MDisk.
- **`addvolumecopy`**

  Add a new copy to an existing volume. The new copy is always synchronized from the existing copy. For HyperSwap topology systems, this creates a highly available volume. This command can be used to create the following volume types:

  - Mirrored volume (standard topology)
  - HyperSwap volume (HyperSwap topology)
- **`rmvolumecopy`**

  Remove a copy of a volume. Leaves the volume intact. Converts a Mirrored, or HyperSwap volume into a basic volume. For a HyperSwap volume, this includes deleting the active-active relationship and the change volumes.

  This enables a copy to be identified simply by its site.

  The **`-force`** parameter with **`rmvdiskcopy`** is replaced by individual override parameters, making it clearer to the user exactly what protection they are bypassing.

  Refer to IBM Knowledge center at `https://ibm.biz/BdsKgy` for more details.

# 7.7 Mapping Volumes to Host after volume creation

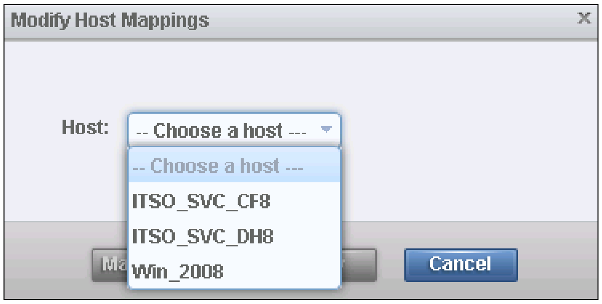
Newly created volume can be mapped to the host at creation time, or later. If the volume was not mapped to a host during creation, then to map it to a host, follow the steps in 7.7.1, “Mapping newly created volumes to the host using the wizard” on page 249.

## 7.7.1 Mapping newly created volumes to the host using the wizard

We continue to map the volume that was created in 7.3, “Creating volumes using the Quick Volume Creation” on page 228. We assume that you followed that procedure and clicked **Continue** as, for example, shown in Figure 7-18 on page 235.

To map the volumes, complete the following steps:

1. Select a host to which the new volume should be attached as shown in Figure 7-37.



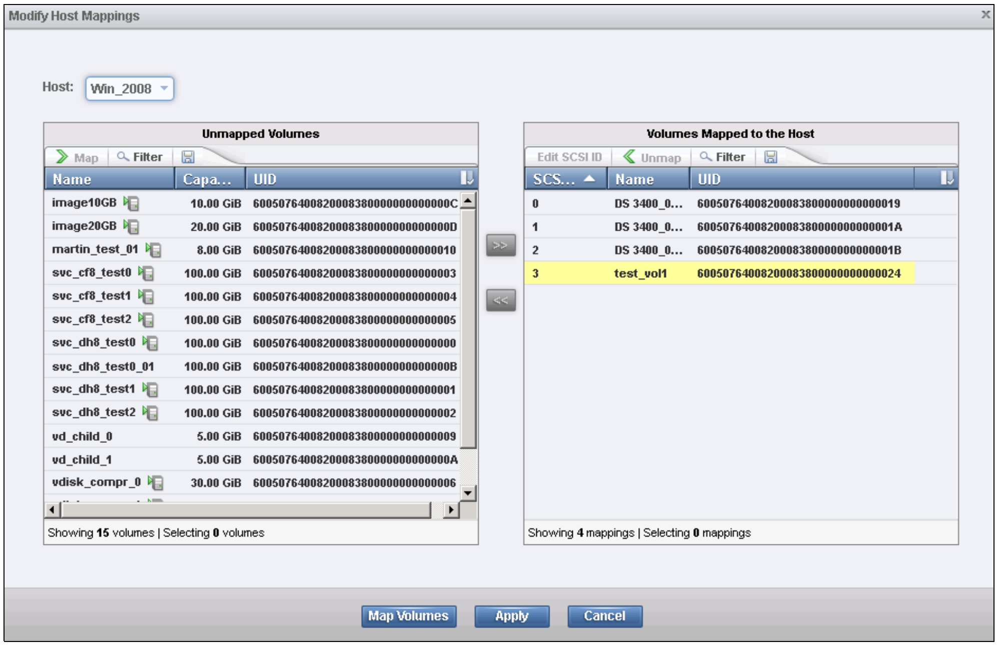
Modify Host Mappings

Host: -- Choose a host ---

- -- Choose a host ---
- ITSO_SVC_CF8
- ITSO_SVC_DH8
- Win_2008

*Figure 7-37 Choose a host*

2. The Modify Host Mappings window opens, and your host and the newly created volume are already selected. Click **Map Volumes** to map the volume to the host as shown in Figure 7-38.



Modify Host Mappings

Host: Win_2008

**Unmapped Volumes**

| Name | Capa... | UID |
|---|---|---|
| image10GB | 10.00 GiB | 6005076400820008380000000000000C |
| image20GB | 20.00 GiB | 6005076400820008380000000000000D |
| martin_test_01 | 8.00 GiB | 60050764008200083800000000000010 |
| svc_cf8_test0 | 100.00 GiB | 60050764008200083800000000000003 |
| svc_cf8_test1 | 100.00 GiB | 60050764008200083800000000000004 |
| svc_cf8_test2 | 100.00 GiB | 60050764008200083800000000000005 |
| svc_dh8_test0 | 100.00 GiB | 60050764008200083800000000000000 |
| svc_dh8_test0_01 | 100.00 GiB | 6005076400820008380000000000000B |
| svc_dh8_test1 | 100.00 GiB | 60050764008200083800000000000001 |
| svc_dh8_test2 | 100.00 GiB | 60050764008200083800000000000002 |
| vd_child_0 | 5.00 GiB | 60050764008200083800000000000009 |
| vd_child_1 | 5.00 GiB | 6005076400820008380000000000000A |
| vdisk_compr_0 | 30.00 GiB | 60050764008200083800000000000006 |

Showing **15** volumes | Selecting **0** volumes

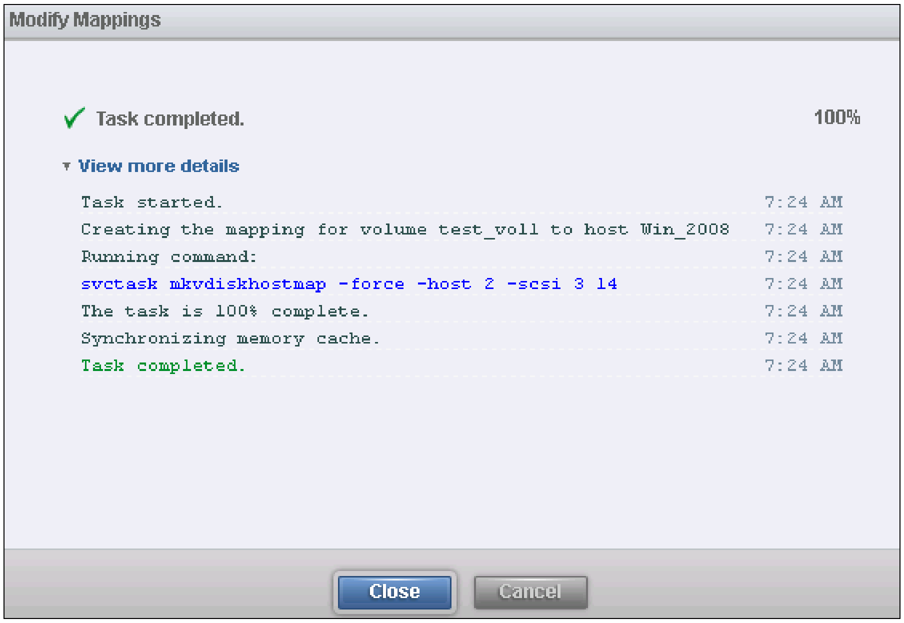
**Volumes Mapped to the Host**

| SCS... | Name | UID |
|---|---|---|
| 0 | DS 3400_0... | 60050764008200083800000000000019 |
| 1 | DS 3400_0... | 6005076400820008380000000000001A |
| 2 | DS 3400_0... | 6005076400820008380000000000001B |
| 3 | test_vol1 | 60050764008200083800000000000024 |

Showing **4** mappings | Selecting **0** mappings

*Figure 7-38 Modify mappings*

3. The confirmation window shows the result of mapping the volume task as shown in Figure 7-39.



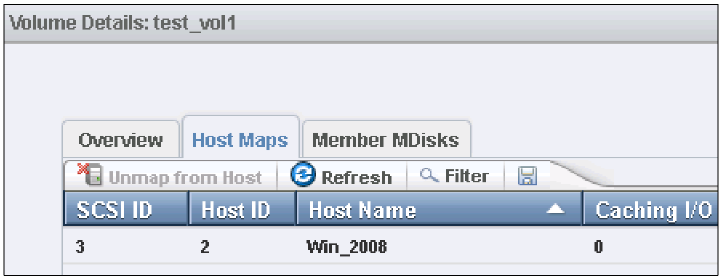
✓ Task completed. 100%

▾ View more details

```
Task started.                                              7:24 AM
Creating the mapping for volume test_vol1 to host Win_2008 7:24 AM
Running command:                                           7:24 AM
svctask mkvdiskhostmap -force -host 2 -scsi 3 14           7:24 AM
The task is 100% complete.                                 7:24 AM
Synchronizing memory cache.                                7:24 AM
Task completed.                                            7:24 AM
```

*Figure 7-39 Confirmation of volume to host mapping*

4. After the task completes, the wizard returns to the Volumes window. By double-clicking the volume, you can see the host maps as shown in Figure 7-40.



| SCSI ID | Host ID | Host Name ▲ | Caching I/O |
|---|---|---|---|
| 3 | 2 | Win_2008 | 0 |

*Figure 7-40 Host maps*

The host is now able to access the volumes and store data on them. See 7.8, “Migrating a volume to another storage pool” on page 251 for information about discovering the volumes on the host and making additional host settings, if required.

Multiple volumes can also be created in preparation for discovering them later, and customize mappings.

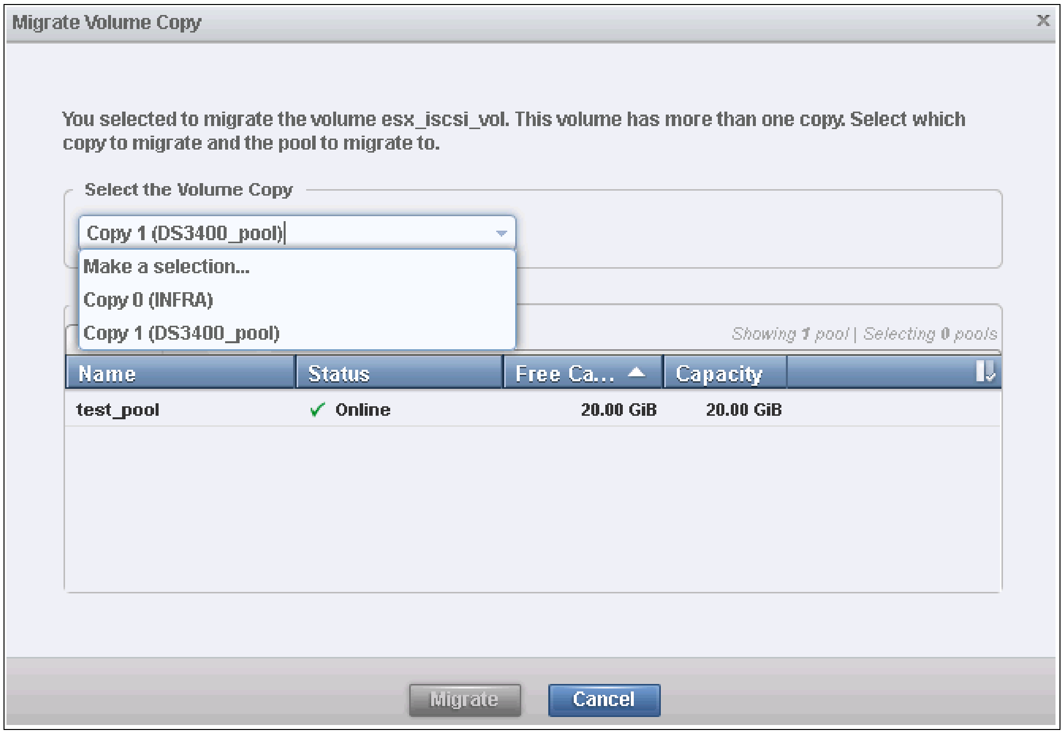
## 7.8 Migrating a volume to another storage pool

IBM Spectrum Virtualize provides online volume migration while applications are running. Storage pools are managed disk groups, as described in Chapter 6, “Storage pools” on page 165. With volume migration, data can be moved between storage pools, regardless of whether the pool is an internal pool, or a pool on another external storage system. This migration is done without the server and application knowing that it even occurred.

The migration process itself is a low priority process that does not affect the performance of the IBM Spectrum Virtualize system. However, it moves one extent after another to the new storage pool, so the performance of the volume is affected by the performance of the new storage pool after the migration process.

To migrate a volume to another storage pool, complete the following steps:

1. Click **Migrate to Another Pool** as shown in Figure 7-41. The Migrate Volume Copy window opens. If your volume consists of more than one copy, select the copy (from the menu shown in Figure 7-41) that you want to migrate to another storage pool. If the selected volume consists of one copy, this selection menu is not available.



Migrate Volume Copy

You selected to migrate the volume esx_iscsi_vol. This volume has more than one copy. Select which copy to migrate and the pool to migrate to.

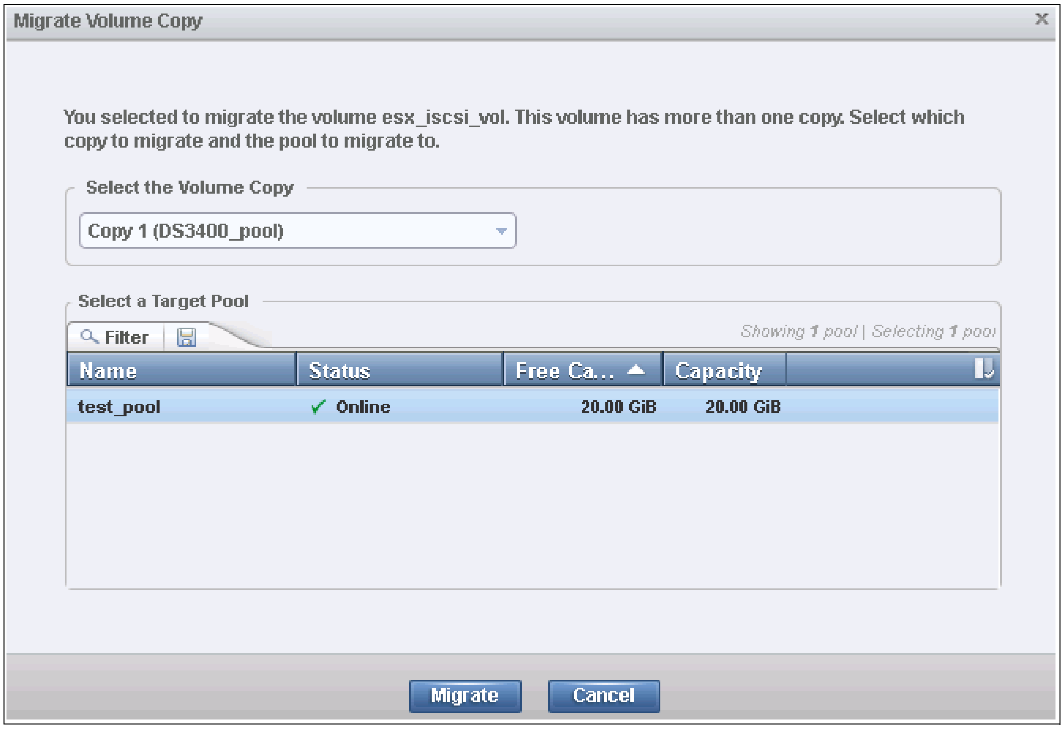
Select the Volume Copy

Copy 1 (DS3400_pool)

| Name | Status | Free Ca... | Capacity |
|---|---|---|---|
| test_pool | ✓ Online | 20.00 GiB | 20.00 GiB |

*Figure 7-41 Migrate Volume Copy window: Select copy*

2. Select the new target storage pool and click **Migrate** as shown in Figure 7-42.


Migrate Volume Copy

You selected to migrate the volume esx_iscsi_vol. This volume has more than one copy. Select which copy to migrate and the pool to migrate to.

Select the Volume Copy: Copy 1 (DS3400_pool)

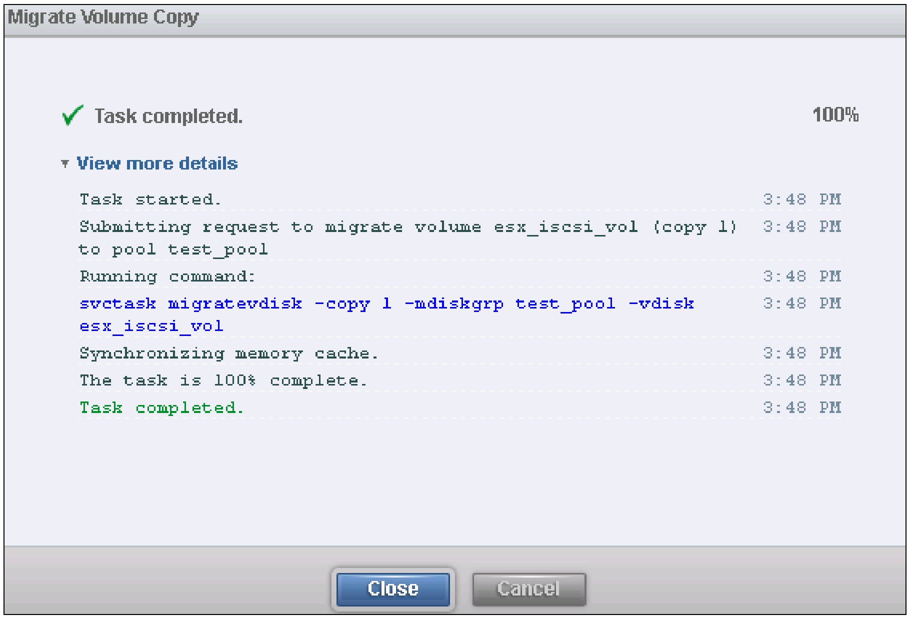
Select a Target Pool

Showing 1 pool | Selecting 1 pool

| Name | Status | Free Ca... | Capacity |
|---|---|---|---|
| test_pool | ✓ Online | 20.00 GiB | 20.00 GiB |

*Figure 7-42 Migrate Volume Copy: select the target pool*

3. The volume copy migration starts as shown in Figure 7-43. Click **Close** to return to the Volumes pane.


Migrate Volume Copy

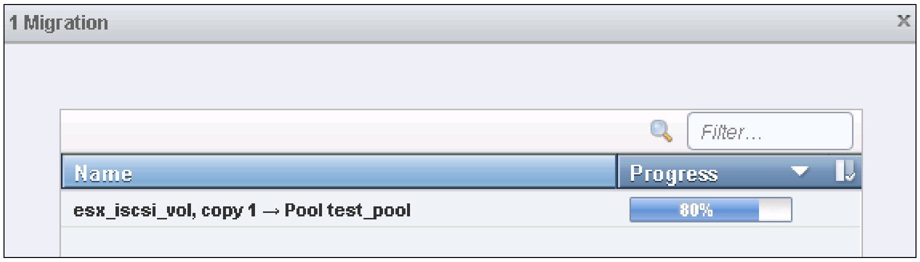
✓ Task completed. 100%

View more details

```
Task started.                                                     3:48 PM
Submitting request to migrate volume esx_iscsi_vol (copy 1)       3:48 PM
to pool test_pool
Running command:                                                  3:48 PM
svctask migratevdisk -copy 1 -mdiskgrp test_pool -vdisk           3:48 PM
esx_iscsi_vol
Synchronizing memory cache.                                       3:48 PM
The task is 100% complete.                                        3:48 PM
Task completed.                                                   3:48 PM
```

*Figure 7-43 Migrate Volume Copy started*

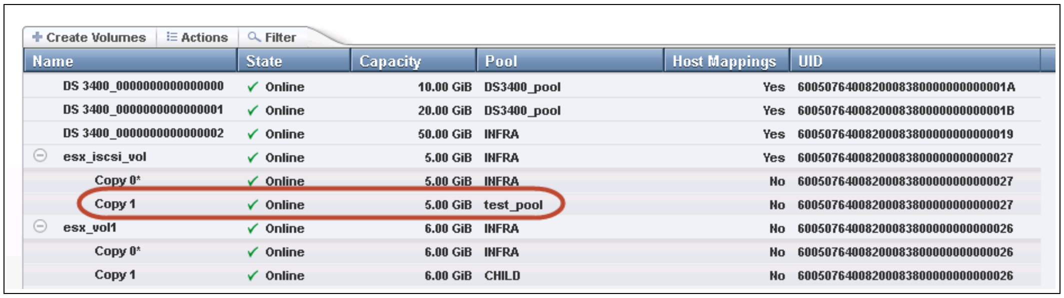
Depending on the size of the volume, the migration process takes some time, but you can monitor the status of the migration in the running tasks bar at the bottom of the GUI as shown in Figure 7-44.



| Name | Progress |
|---|---|
| esx_iscsi_vol, copy 1 → Pool test_pool | 80% |

*Figure 7-44 Migration progress*

After the migration is completed, the volume is shown in the new storage pool. Figure 7-45 shows that it was moved from the `DS3400_pool` to the `test_pool`.



| Name | State | Capacity | Pool | Host Mappings | UID |
|---|---|---|---|---|---|
| DS 3400_0000000000000000 | ✓ Online | 10.00 GiB | DS3400_pool | Yes | 6005076400820008380000000000001A |
| DS 3400_0000000000000001 | ✓ Online | 20.00 GiB | DS3400_pool | Yes | 6005076400820008380000000000001B |
| DS 3400_0000000000000002 | ✓ Online | 50.00 GiB | INFRA | Yes | 60050764008200083800000000000019 |
| esx_iscsi_vol | ✓ Online | 5.00 GiB | INFRA | Yes | 60050764008200083800000000000027 |
| Copy 0* | ✓ Online | 5.00 GiB | INFRA | No | 60050764008200083800000000000027 |
| Copy 1 | ✓ Online | 5.00 GiB | test_pool | No | 60050764008200083800000000000027 |
| esx_vol1 | ✓ Online | 6.00 GiB | INFRA | No | 60050764008200083800000000000026 |
| Copy 0* | ✓ Online | 6.00 GiB | INFRA | No | 60050764008200083800000000000026 |
| Copy 1 | ✓ Online | 6.00 GiB | CHILD | No | 60050764008200083800000000000026 |

*Figure 7-45 Migration complete*

The volume copy has now been migrated without any host or application downtime to the new storage pool. It is also possible to migrate both volume copies to other pools online.

Another way to migrate volumes to another pool is by performing the migration using the volume copies, as described in 7.9, “Migrating volumes using the volume copy feature”.

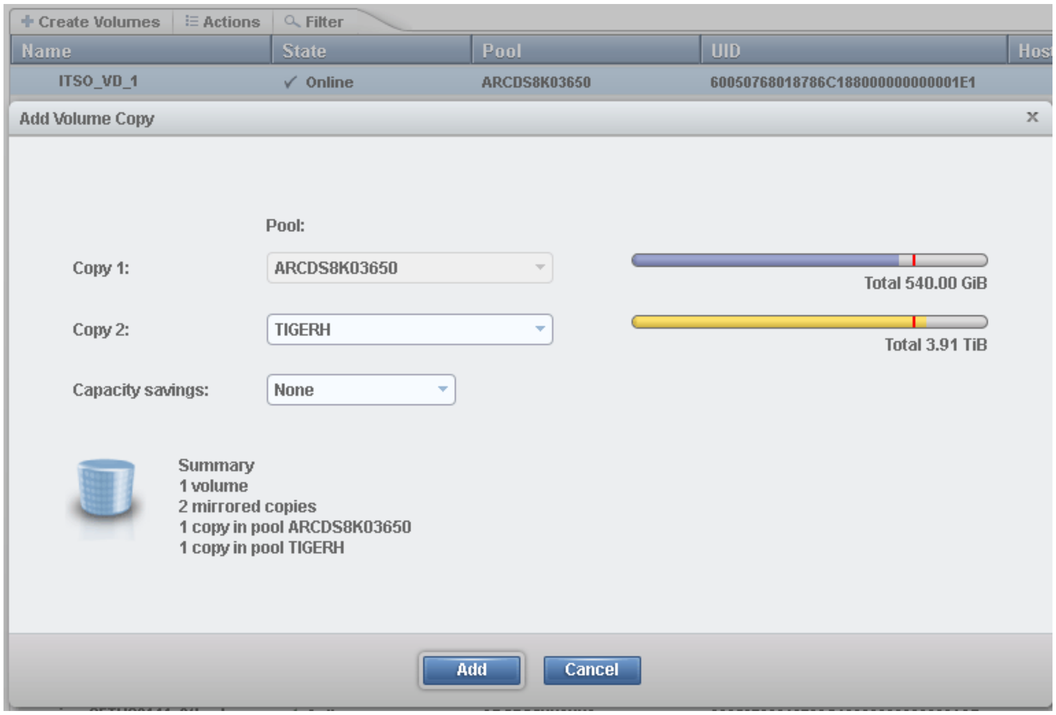
> **Note:** Migrating a volume between storage pools with different extent sizes is *not* supported. If you need to migrate a volume to a storage pool with a different extent size, use volume copy features instead.

## 7.9 Migrating volumes using the volume copy feature

IBM Spectrum Virtualize supports creation, synchronizing, splitting, and deleting volume copies. A combination of these tasks can be used to migrate volumes to other storage pools. The easiest way to migrate volume copies is to use the migration feature described in 7.8, “Migrating a volume to another storage pool”. If you use this feature, one extent after another is migrated to the new storage pool. However, using volume copies provides another possibility to migrate volumes.

To migrate a volume, complete the following steps:

1. Create a second copy of your volume in the target storage pool as shown in Figure 7-46 on page 255.



| Name | State | Pool | UID | Hos |
|---|---|---|---|---|
| ITSO_VD_1 | ✓ Online | ARCDS8K03650 | 60050768018786C188000000000001E1 | |

Pool:

Copy 1: ARCDS8K03650 Total 540.00 GiB

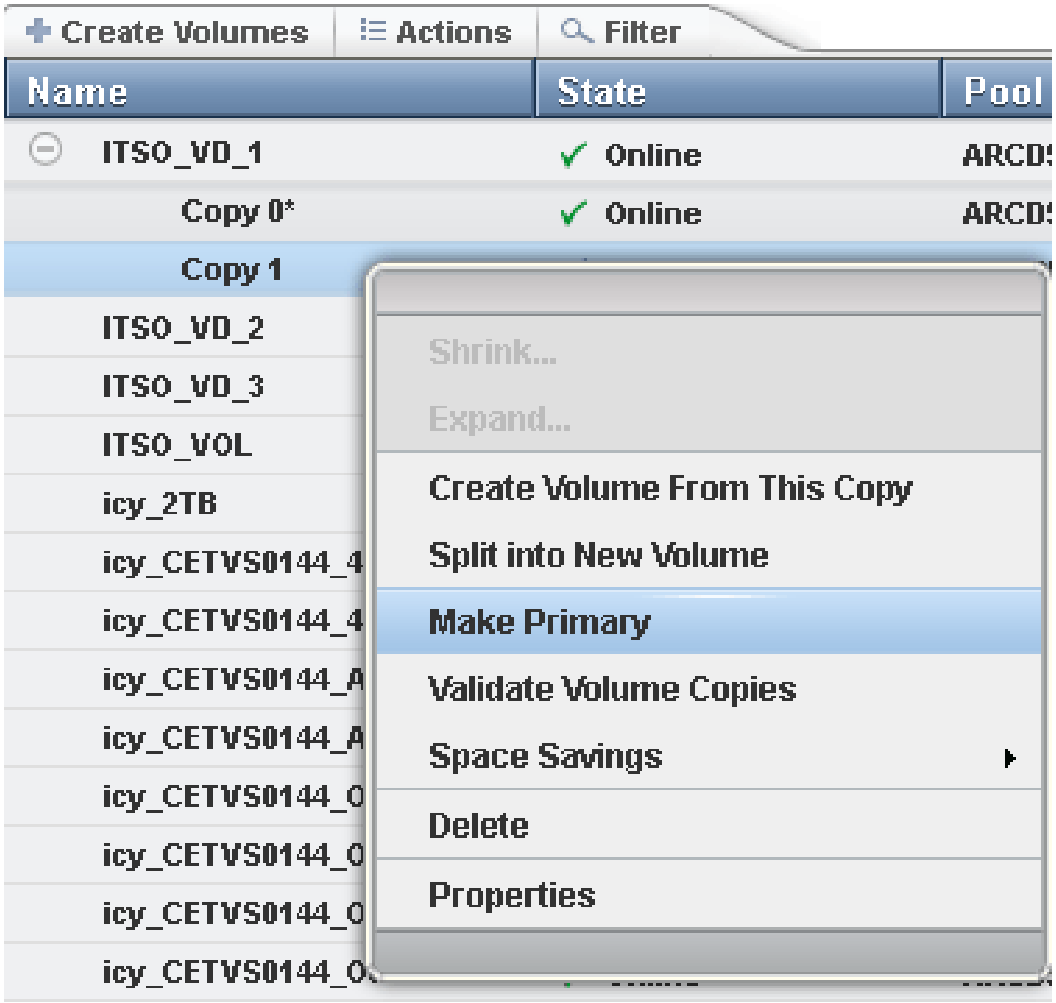
Copy 2: TIGERH Total 3.91 TiB

Capacity savings: None

Summary
1 volume
2 mirrored copies
1 copy in pool ARCDS8K03650
1 copy in pool TIGERH

*Figure 7-46 Adding the volume copy to another pool*

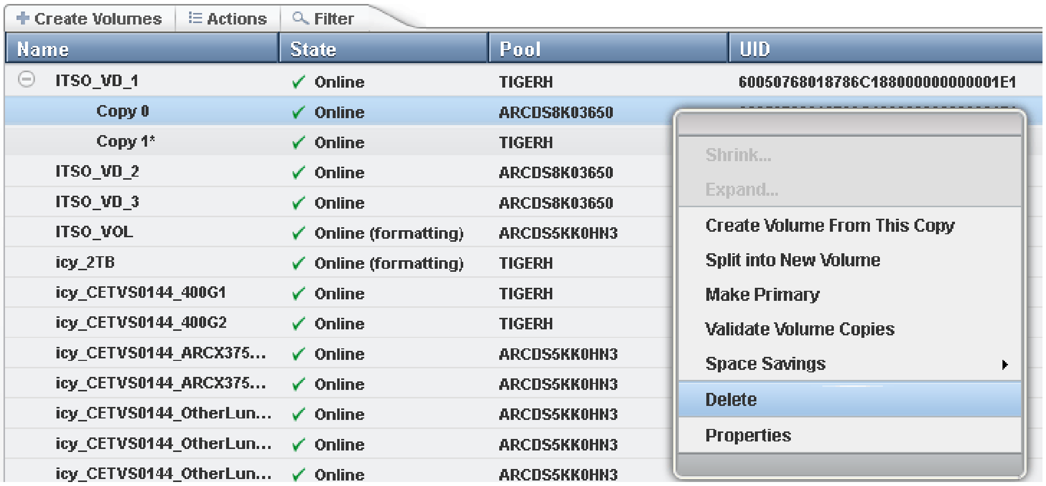
2. Wait until the copies are synchronized.
3. Change the role of the copies and make the new copy of the primary copy as shown in Figure 7-47 on page 256.



| Name | State | Pool |
|---|---|---|
| ITSO_VD_1 | ✓ Online | ARCD |
| Copy 0* | ✓ Online | ARCD |
| Copy 1 | | |
| ITSO_VD_2 | | |
| ITSO_VD_3 | | |
| ITSO_VOL | | |
| icy_2TB | | |
| icy_CETVS0144_4 | | |
| icy_CETVS0144_4 | | |
| icy_CETVS0144_A | | |
| icy_CETVS0144_A | | |
| icy_CETVS0144_O | | |
| icy_CETVS0144_O | | |
| icy_CETVS0144_O | | |
| icy_CETVS0144_O | | |

Figure 7-47 Making the new copy in a different storage pool as primary

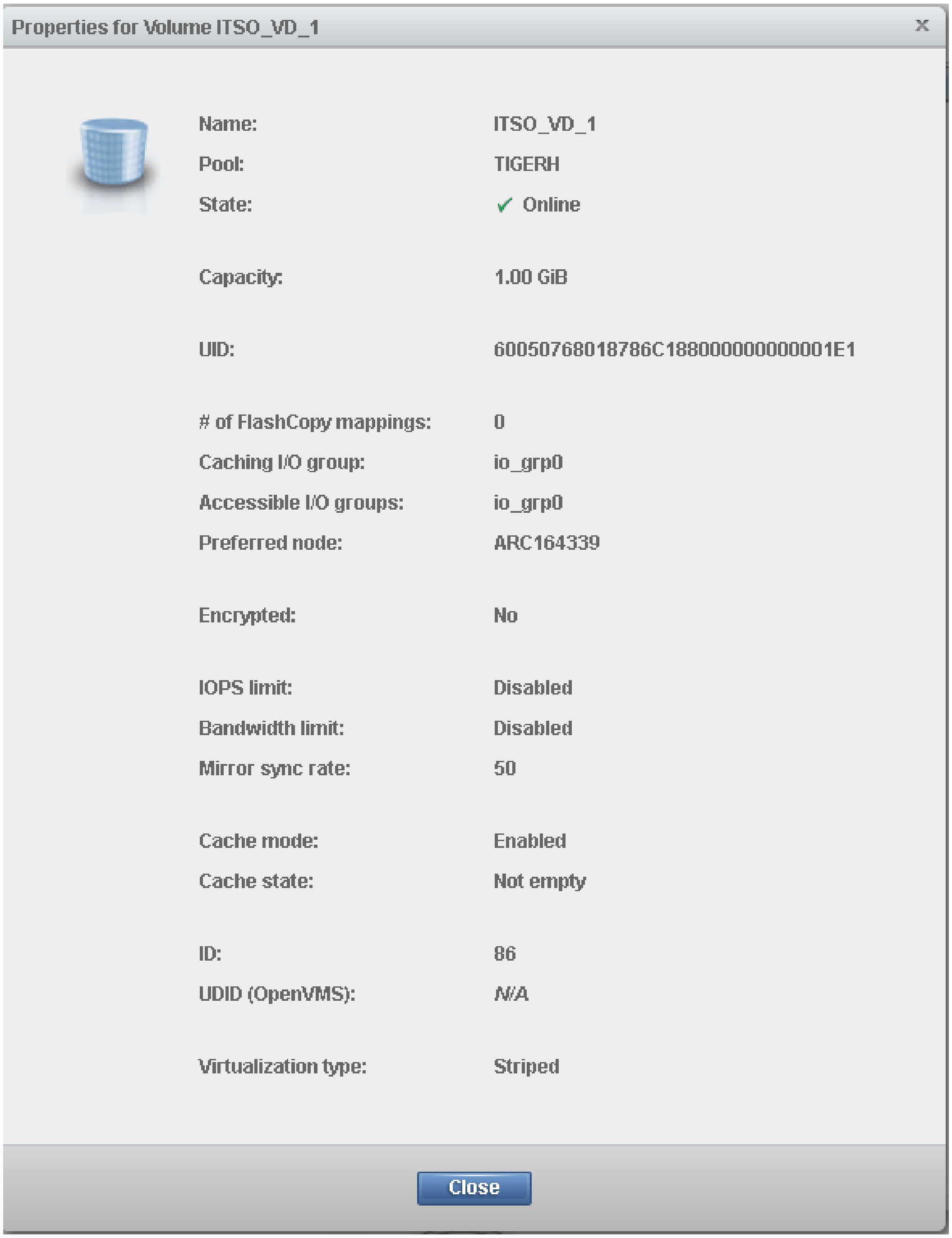
4. Split or delete the old copy from the volume as shown in Figure 7-48.



| Name | State | Pool | UID |
|---|---|---|---|
| ITSO_VD_1 | ✓ Online | TIGERH | 60050768018786C188000000000001E1 |
| Copy 0 | ✓ Online | ARCDS8K03650 | |
| Copy 1* | ✓ Online | TIGERH | |
| ITSO_VD_2 | ✓ Online | ARCDS8K03650 | |
| ITSO_VD_3 | ✓ Online | ARCDS8K03650 | |
| ITSO_VOL | ✓ Online (formatting) | ARCDS5KK0HN3 | |
| icy_2TB | ✓ Online (formatting) | TIGERH | |
| icy_CETVS0144_400G1 | ✓ Online | TIGERH | |
| icy_CETVS0144_400G2 | ✓ Online | TIGERH | |
| icy_CETVS0144_ARCX375... | ✓ Online | ARCDS5KK0HN3 | |
| icy_CETVS0144_ARCX375... | ✓ Online | ARCDS5KK0HN3 | |
| icy_CETVS0144_OtherLun... | ✓ Online | ARCDS5KK0HN3 | |
| icy_CETVS0144_OtherLun... | ✓ Online | ARCDS5KK0HN3 | |
| icy_CETVS0144_OtherLun... | ✓ Online | ARCDS5KK0HN3 | |

Figure 7-48 Deleting the older copy

5. Ensure that the new copy is in the target storage pool as shown in Figure 7-49 on page 257



Properties for Volume ITSO_VD_1

Name: ITSO_VD_1
Pool: TIGERH
State: ✓ Online

Capacity: 1.00 GiB

UID: 60050768018786C18800000000000001E1

# of FlashCopy mappings: 0
Caching I/O group: io_grp0
Accessible I/O groups: io_grp0
Preferred node: ARC164339

Encrypted: No

IOPS limit: Disabled
Bandwidth limit: Disabled
Mirror sync rate: 50

Cache mode: Enabled
Cache state: Not empty

ID: 86
UDID (OpenVMS): N/A

Virtualization type: Striped

*Figure 7-49 Verifying the new copy in the target storage pool*

This migration process requires more user interaction, but it offers some benefits, for example, if you migrate a volume from a tier 1 storage pool to a lower performance tier 2 storage pool. In step 1, you create the copy on the tier 2 pool; all reads are still performed in the tier 1 pool to the primary copy. After the synchronization, all writes are destaged to both pools, but the reads are still only done from the primary copy.

Now you can switch the role of the copies online (step 3), and test the performance of the new pool. If you are done testing your lower performance pool, you can split or delete the old copy in tier 1, or switch back to tier 1 in seconds, in case tier 2 pool did not meet your performance requirements.

# 7.10 Volume operations using CLI

This section describes various configuration and administrative tasks that can be performed on volumes that can be carried out on volumes using the command line interface (CLI). For more information, see the command-line interface section in IBM Knowledge Center at:

https://ibm.biz/BdsKgv

Refer to Appendix B, "CLI setup and SAN boot" on page 675 for how to implement SAN Boot.

## 7.10.1 Creating a volume

The **mkvdisk** command creates sequential, striped, or image mode volume objects. When they are mapped to a host object, these objects are seen as disk drives with which the host can perform I/O operations. When a volume is created, you must enter several parameters at the CLI. Mandatory and optional parameters are available.

> **Creating an image mode disk:** If you do not specify the **-size** parameter when you create an image mode disk, the entire MDisk capacity is used.

You must know the following information before you start to create the volume:

- In which storage pool the volume has its extents
- From which I/O Group the volume is accessed
- Which IBM Spectrum Virtualize node is the preferred node for the volume
- Size of the volume
- Name of the volume
- Type of the volume
- Whether this volume is managed by IBM Easy Tier to optimize its performance

When you are ready to create your striped volume, use the **mkvdisk** command. In Example 7-1, this command creates a 10 gigabyte (GB) striped volume with volume ID 20 within the storage pool STGPool_DS3500-2 and assigns it to the io_grp0 I/O Group. Its preferred node is node 1.

*Example 7-1 The mkvdisk command*

```
IBM_Storwize:ITSO:superuser>mkvdisk -mdiskgrp Site1_Pool -iogrp io_grp0 -size 10 -unit gb
-name Tiger
Virtual Disk, id [6], successfully created
```

To verify the results, use the **lsvdisk** command, as shown in Example 7-2.

*Example 7-2 The lsvdisk command*

```
IBM_Storwize:ITSO:superuser>lsvdisk 6
id 6
name Tiger
IO_group_id 0
IO_group_name io_grp0
status online
mdisk_grp_id 3
```

```
mdisk_grp_name Site1_Pool
capacity 10.00GB
type striped
formatted no
formatting yes
mdisk_id
mdisk_name
FC_id
FC_name
RC_id
RC_name
vdisk_UID 6005076801FE80840800000000000020
throttling 0
preferred_node_id 4
fast_write_state not_empty
cache readwrite
udid
fc_map_count 0
sync_rate 50
copy_count 1
se_copy_count 0
filesystem
mirror_write_priority latency
RC_change no
compressed_copy_count 0
access_IO_group_count 1
last_access_time
parent_mdisk_grp_id 3
parent_mdisk_grp_name Site1_Pool
owner_type none
owner_id
owner_name
encrypt no
volume_id 6
volume_name Tiger
function

copy_id 0
status online
sync yes
auto_delete no
primary yes
mdisk_grp_id 3
mdisk_grp_name Site1_Pool
type striped
mdisk_id
mdisk_name
fast_write_state not_empty
used_capacity 10.00GB
real_capacity 10.00GB
free_capacity 0.00MB
overallocation 100
autoexpand
warning
grainsize
se_copy no
easy_tier on
easy_tier_status balanced
tier ssd
tier_capacity 0.00MB
tier enterprise
```

```
tier_capacity 10.00GB
tier nearline
tier_capacity 0.00MB
compressed_copy no
uncompressed_used_capacity 10.00GB
parent_mdisk_grp_id 3
parent_mdisk_grp_name Site1_Pool
encrypt no
```

The required tasks to create a volume are complete.

## 7.10.2 Volume information

Use the **`lsvdisk`** command to display summary information about all volumes that are defined within the IBM Spectrum Virtualize environment as shown in Example 7-3. To display more detailed information about a specific volume, run the command again and append the volume name parameter or the volume ID.

*Example 7-3 The lsvdisk command*

```
IBM_2145:ITSO_DH8:superuser>lsvdisk -delim " "
id name IO_group_id IO_group_name status mdisk_grp_id mdisk_grp_name capacity type FC_id
FC_name RC_id RC_name vdisk_UID fc_map_count copy_count fast_write_state se_copy_count
RC_change compressed_copy_count parent_mdisk_grp_id parent_mdisk_grp_name formatting
encrypt volume_id volume_name function
0 SQL_Data0 0 io_grp0 online 3 Site1_Pool 300.00GB striped
6005076801FE80840800000000000000 0 1 empty 0 no 0 3 Site1_Pool no no 0 SQL_Data0
1 vdisk0 0 io_grp0 online 0 Site2_Pool 2.00TB striped     6005076801FE8084080000000000001B
0 1 empty 1 no 0 0 Site2_Pool no no 1 vdisk0
2 SQL_test 0 io_grp0 online many many 10.00GB many     6005076801FE8084080000000000001C 0 2
empty 0 no 0 many many no no 2 SQL_TEST striped
3 test_basic 0 io_grp0 online many many 5.00GB many     6005076801FE8084080000000000001D 0
2 empty 0 no 0 many many no no 3 test_basic
4 SQL_Data4 0 io_grp0 online 3 Site1_Pool 300.00GB striped
6005076801FE80840800000000000004 0 1 empty 0 no 0 3 Site1_Pool no no 4 SQL_Data4
5 VM_Datastore0 0 io_grp0 online 0 Site2_Pool 250.00GB striped
6005076801FE80840800000000000005 0 1 empty 0 no 0 0 Site2_Pool no no 5 VM_Datastore0
9 VM_Datastore4 0 io_grp0 online 0 Site2_Pool 250.00GB striped
6005076801FE80840800000000000009 0 1 empty 0 no 0 0 Site2_Pool no no 9 VM_Datastore4
15 child5 0 io_grp0 online 0 Site2_Pool 10.00GB striped
6005076801FE80840800000000000010 0 1 empty 0 no 0 0 Site2_Pool no no 15 child5
```

## 7.10.3 Creating a thin-provisioned volume

Example 7-4 shows how to create a thin-provisioned volume. In addition to the normal parameters, you must use the following parameters:

**`-rsize`** This parameter makes the volume a thin-provisioned volume; otherwise, the volume is fully allocated.

**`-autoexpand`** This parameter specifies that thin-provisioned volume copies automatically expand their real capacities by allocating new extents from their storage pool.

**`-grainsize`** This parameter sets the grain size in kilobytes (KB) for a thin-provisioned volume.

*Example 7-4 Usage of the command mkvdisk*

```
IBM_Storwize:ITSO:superuser>mkvdisk -mdiskgrp Site1_Pool -iogrp 0 -vtype striped -size 10
-unit gb -rsize 50% -autoexpand -grainsize 256
Virtual Disk, id [21], successfully created
```

This command creates a space-efficient 10 GB volume. The volume belongs to the storage pool that is named `Site1_Pool` and is owned by input output (I/O) Group `io_grp0`. The real capacity automatically expands until the volume size of 10 GB is reached. The grain size is set to 256 KB, which is the default.

> **Disk size:** When the **`-rsize`** parameter is used, you have the following options: `disk_size`, `disk_size_percentage`, and **`auto`**.
>
> Specify the `disk_size_percentage` value by using an integer, or an integer that is immediately followed by the percent (%) symbol.
>
> Specify the units for a `disk_size` integer by using the **`-unit`** parameter; the default is MB. The **`-rsize`** value can be greater than, equal to, or less than the size of the volume.
>
> The **`auto`** option creates a volume copy that uses the entire size of the MDisk. If you specify the **`-rsize auto`** option, you must also specify the **`-vtype image`** option.
>
> An entry of 1 GB uses 1,024 MB.

### 7.10.4 Creating a volume in image mode

This virtualization type enables an image mode volume to be created when an MDisk has data on it, perhaps from a pre-virtualized subsystem. When an image mode volume is created, it directly corresponds to the previously unmanaged MDisk from which it was created. Therefore, except for a thin-provisioned image mode volume, the volume's logical block address (LBA) *x* equals MDisk LBA *x*.

You can use this command to bring a non-virtualized disk under the control of the clustered system. After it is under the control of the clustered system, you can migrate the volume from the single managed disk.

When the first MDisk extent is migrated, the volume is no longer an image mode volume. You can add an image mode volume to an already populated storage pool with other types of volumes, such as striped or sequential volumes.

> **Size:** An image mode volume must be at least 512 bytes (the capacity cannot be 0). That is, the minimum size that can be specified for an image mode volume must be the same as the storage pool extent size to which it is added, with a minimum of 16 MiB.

You must use the **`-mdisk`** parameter to specify an MDisk that has a mode of `unmanaged`. The **`-fmtdisk`** parameter cannot be used to create an image mode volume.

> **Capacity:** If you create a mirrored volume from two image mode MDisks without specifying a **`-capacity`** value, the capacity of the resulting volume is the smaller of the two MDisks and the remaining space on the larger MDisk is inaccessible.
>
> If you do not specify the **`-size`** parameter when you create an image mode disk, the entire MDisk capacity is used.

Use the **mkvdisk** command to create an image mode volume, as shown in Example 7-5.

*Example 7-5 The mkvdisk (image mode) command*

```
IBM_Storwize:ITSO:superuser>mkvdisk -mdiskgrp ITSO_Pool1 -iogrp 0 -mdisk mdisk25 -vtype
image -name Image_Volume_A
Virtual Disk, id [6], successfully created
```

This command creates an image mode volume that is called `Image_Volume_A` that uses the `mdisk10` MDisk. The volume belongs to the storage pool `STGPool_DS3500-1` and the volume is owned by the `io_grp0` I/O Group.

If we run the **lsvdisk** command again, the volume that is named `Image_Volume_A` has a status of image, as shown in Example 7-6.

*Example 7-6 The lsvdisk command*

```
IBM_Storwize:ITSO:superuser>lsvdisk -filtervalue type=image
id name           IO_group_id IO_group_name status mdisk_grp_id mdisk_grp_name capacity
type  FC_id FC_name RC_id RC_name vdisk_UID                        fc_map_count copy_count
fast_write_state se_copy_count RC_change compressed_copy_count parent_mdisk_grp_id
parent_mdisk_grp_name formatting encrypt volume_id volume_name    function
6  Image_Volume_A 0           io_grp0       online 5            ITSO_Pool1     1.00GB
image                            6005076801FE80840800000000000021 0            1
empty            0             no        0                     5
ITSO_Pool1            no          no       6         Image_Volume_A
```

## 7.10.5 Adding a mirrored volume copy

You can create a mirrored copy of a volume, which keeps a volume accessible even when the MDisk on which it depends becomes unavailable. You can create a copy of a volume on separate storage pools or by creating an image mode copy of the volume. Copies increase the availability of data; however, they are not separate objects. You can create or change mirrored copies from the volume only.

In addition, you can use volume mirroring as an alternative method of migrating volumes between storage pools.

For example, if you have a non-mirrored volume in one storage pool and want to migrate that volume to another storage pool, you can add a copy of the volume and specify the second storage pool. After the copies are synchronized, you can delete the copy on the first storage pool. The volume is copied to the second storage pool while remaining online during the copy.

To create a mirrored copy of a volume, use the **addvdiskcopy** command. This command adds a copy of the chosen volume to the selected storage pool, which changes a non-mirrored volume into a mirrored volume.

In the following scenario, we show creating a mirrored volume from one storage pool to another storage pool.

As you can see in Example 7-7, the volume has a copy with `copy_id 0`.

*Example 7-7 The lsvdisk command*

```
IBM_Storwize:ITSO:superuser>lsvdisk Volume_no_mirror
id 23
name Volume_no_mirror
IO_group_id 0
```

```
IO_group_name io_grp0
status online
mdisk_grp_id 0
mdisk_grp_name STGPool_DS3500-1
capacity 1.00GB
type striped
formatted no
mdisk_id
mdisk_name
FC_id
FC_name
RC_id
RC_name
vdisk_UID 6005076801AF813F1000000000000019
throttling 0
preferred_node_id 1
fast_write_state empty
cache readwrite
udid
fc_map_count 0
sync_rate 50
copy_count 1
se_copy_count 0
filesystem
mirror_write_priority latency

copy_id 0
status online
sync yes
primary yes
mdisk_grp_id 0
mdisk_grp_name STGPool_DS3500-1
type striped
mdisk_id
mdisk_name
fast_write_state empty
used_capacity 1.00GB
real_capacity 1.00GB
free_capacity 0.00MB
overallocation 100
autoexpand
warning
grainsize
se_copy no
easy_tier on
easy_tier_status inactive
tier generic_ssd
tier_capacity 0.00MB
tier generic_hdd
tier_capacity
1.00GB
```

In Example 7-8, we add the volume copy mirror by using the **addvdiskcopy** command.

*Example 7-8 The addvdiskcopy command*

```
IBM_Storwize:ITSO:superuser>addvdiskcopy -mdiskgrp STGPool_DS5000-1 -vtype striped -unit gb
Volume_no_mirror
Vdisk [23] copy [1] successfully created
```

During the synchronization process, you can see the status by using the **lsvdisksyncprogress** command. As shown in Example 7-9, the first time that the status is checked, the synchronization progress is at 48%, and the estimated completion time is 151026203918 (Estimated completion time is displayed in the YYMMDDHHMMSS format. In our example it is 2016, Oct-26 20:39:18). The second time that the command is run, the progress status is at 100%, and the synchronization is complete.

*Example 7-9 Synchronization*

```
IBM_Storwize:ITSO:superuser>lsvdisksyncprogress
vdisk_id vdisk_name       copy_id progress estimated_completion_time
23       Volume_no_mirror 1       48       161026203918
IBM_Storwize:ITSO:superuser>lsvdisksyncprogress
vdisk_id vdisk_name       copy_id progress estimated_completion_time
23       Volume_no_mirror 1       100
```

As you can see in Example 7-10, the new mirrored volume copy (`copy_id 1`) was added and can be seen by using the **lsvdisk** command.

*Example 7-10 The lsvdisk command*

```
IBM_Storwize:ITSO:superuser>lsvdisk 23
id 23
name Volume_no_mirror
IO_group_id 0
IO_group_name io_grp0
status online
mdisk_grp_id many
mdisk_grp_name many
capacity 1.00GB
type many
formatted no
mdisk_id many
mdisk_name many
FC_id
FC_name
RC_id
RC_name
vdisk_UID 6005076801AF813F1000000000000019
throttling 0
preferred_node_id 1
fast_write_state empty
cache readwrite
udid
fc_map_count 0
sync_rate 50
copy_count 2
se_copy_count 0
filesystem
mirror_write_priority latency

copy_id 0
status online
sync yes
primary yes
mdisk_grp_id 0
mdisk_grp_name STGPool_DS3500-1
type striped
mdisk_id
mdisk_name
```

```
fast_write_state empty
used_capacity 1.00GB
real_capacity 1.00GB
free_capacity 0.00MB
overallocation 100
autoexpand
warning
grainsize
se_copy no
easy_tier on
easy_tier_status inactive
tier generic_ssd
tier_capacity 0.00MB
tier generic_hdd
tier_capacity 1.00GB

copy_id 1
status online
sync yes
primary no
mdisk_grp_id 2
mdisk_grp_name STGPool_DS5000-1
type striped
mdisk_id
mdisk_name
fast_write_state empty
used_capacity 1.00GB
real_capacity 1.00GB
free_capacity 0.00MB
overallocation 100
autoexpand
warning
grainsize
se_copy no
easy_tier on
easy_tier_status inactive
tier generic_ssd
tier_capacity 0.00MB
tier generic_hdd
tier_capacity
1.00GB
```

While you are adding a volume copy mirror, you can define a mirror with different parameters to the volume copy. Therefore, you can define a thin-provisioned volume copy for a non-volume copy volume and vice versa, which is one way to migrate a non-thin-provisioned volume to a thin-provisioned volume.

> **Volume copy mirror parameters:** To change the parameters of a volume copy mirror, you must delete the volume copy and redefine it with the new values.

Now, we can change the name of the volume that was mirrored from `Volume_no_mirror` to `Volume_mirrored`, as shown in Example 7-11.

*Example 7-11 Volume name changes*

```
IBM_Storwize:ITSO:superuser>chvdisk -name Volume_mirrored Volume_no_mirror
```

## 7.10.6 Adding a compressed volume copy

Use the **`addvdiskcopy`** command to add a compressed copy to an existing volume.

We show the usage of **`addvdiskcopy`** with the **`-autodelete`** flag set. The **`-autodelete`** flag specifies the primary copy is deleted after the secondary copy is synchronized.

Example 7-12 shows a shortened **`lsvdisk`** output of an uncompressed volume.

*Example 7-12 An uncompressed volume*

```
IBM_Storwize:ITSO_Gen2_SiteB:superuser>lsvdisk 0
id 0
name Uncompressed_Volume
IO_group_id 0
IO_group_name io_grp0
status online
..
compressed_copy_count 0
..

copy_id 0
status online
sync yes
auto_delete no
primary yes
..
compressed_copy no
uncompressed_used_capacity 2.00GB
..
```

In Example 7-13 we add a compressed copy with the **`-autodelete`** flag set.

*Example 7-13 Compressed copy*

```
IBM_Storwize:ITSO_Gen2_SiteB:superuser>addvdiskcopy -autodelete -rsize 2 -mdiskgrp 0
-compressed 0
Vdisk [0] copy [1] successfully created
```

Example 7-14 shows the **`lsvdisk`** output with an additional compressed volume (copy 1) and volume copy 0 being set to **`auto_delete yes`**.

*Example 7-14 The lsvdisk command output*

```
IBM_Storwize:ITSO_Gen2_SiteB:superuser>lsvdisk 0
id 0
name Uncompressed_Volume
..
copy_count 2
..
compressed_copy_count 1
..

copy_id 0
status online
sync yes
auto_delete yes
primary yes
..
compressed_copy no
```

```
uncompressed_used_capacity 2.00GB
..

copy_id 1
status online
sync no
auto_delete no
primary no
..
compressed_copy yes
uncompressed_used_capacity 0.00MB
..
```

When copy 1 is synchronized, copy 0 is deleted. You can monitor the progress of volume copy synchronization by using **lsvdisksyncprogress**.

> **Note:** Consider the compression leading practices before adding the first compressed copy to a system.

### 7.10.7 Splitting a mirrored volume

The **splitvdiskcopy** command creates a volume in the specified I/O Group from a copy of the specified volume. If the copy that you are splitting is not synchronized, you must use the **-force** parameter. If you are attempting to remove the only synchronized copy, the command fails. To avoid this failure, wait for the copy to synchronize or split the unsynchronized copy from the volume by using the **-force** parameter. You can run the command when either volume copy is offline.

Example 7-15 shows the **splitvdiskcopy** command, which is used to split a mirrored volume. It creates a volume that is named `Volume_new` from the volume that is named `Volume_mirrored`.

*Example 7-15 Split volume*

```
IBM_Storwize:ITSO:superuser>splitvdiskcopy -copy 1 -iogrp 0 -name Volume_new
Volume_mirrored
Virtual Disk, id [24], successfully created
```

As you can see in Example 7-16, the new volume that is named `Volume_new` was created as an independent volume.

*Example 7-16 The lsvdisk command*

```
IBM_Storwize:ITSO:superuser>lsvdisk Volume_new
id 24
name Volume_new
IO_group_id 0
IO_group_name io_grp0
status online
mdisk_grp_id 2
mdisk_grp_name STGPool_DS5000-1
capacity 1.00GB
type striped
formatted no
mdisk_id
mdisk_name
FC_id
```

```
FC_name
RC_id
RC_name
vdisk_UID 6005076801AF813F100000000000001A
throttling 0
preferred_node_id 2
fast_write_state empty
cache readwrite
udid
fc_map_count 0
sync_rate 50
copy_count 1
se_copy_count 0
filesystem
mirror_write_priority latency

copy_id 0
status online
sync yes
primary yes
mdisk_grp_id 2
mdisk_grp_name STGPool_DS5000-1
type striped
mdisk_id
mdisk_name
fast_write_state empty
used_capacity 1.00GB
real_capacity 1.00GB
free_capacity 0.00MB
overallocation 100
autoexpand
warning
grainsize
se_copy no
easy_tier on
easy_tier_status inactive
tier generic_ssd
tier_capacity 0.00MB
tier generic_hdd
tier_capacity 1.00GB
```

By issuing the command that is shown in Example 7-15 on page 267, Volume_mirrored does not have its mirrored copy and a volume is created automatically.

## 7.10.8 Modifying a volume

Running the **chvdisk** command modifies a single property of a volume. Only one property can be modified at a time. Therefore, changing the name and modifying the I/O Group require two invocations of the command.

You can specify a new name or label. The new name can be used to reference the volume. The I/O Group with which this volume is associated can be changed. Changing the I/O Group with which this volume is associated requires a flush of the cache within the nodes in the current I/O Group to ensure that all data is written to disk. I/O must be suspended at the host level before you perform this operation.

**Tips:** If the volume has a mapping to any hosts, it is impossible to move the volume to an I/O Group that does not include any of those hosts.

This operation fails if insufficient space exists to allocate bitmaps for a mirrored volume in the target I/O Group.

If the **`-force`** parameter is used and the system is unable to destage all write data from the cache, the contents of the volume are corrupted by the loss of the cached data.

If the **`-force`** parameter is used to move a volume that has out-of-sync copies, a full resynchronization is required.

### 7.10.9 I/O governing

You can set a limit on the number of I/O operations that are accepted for a volume. The limit is set in terms of I/O operations per second (IOPS) or MBps. By default, no I/O governing rate is set when a volume is created.

Base the choice between I/O and MB as the I/O governing throttle on the disk access profile of the application. Database applications generally issue large amounts of I/O, but they transfer only a relatively small amount of data. In this case, setting an I/O governing throttle that is based on MBps does not achieve much. It is better to use an IOPS as a second throttle.

At the other extreme, a streaming video application generally issues a small amount of I/O, but it transfers large amounts of data. In contrast to the database example, setting an I/O governing throttle that is based on IOPS does not achieve much, so it is better to use an MBps throttle.

**I/O governing rate:** An I/O governing rate of `0` (displayed as throttling in the CLI output of the **`lsvdisk`** command) does not mean that zero IOPS (or MBps) can be achieved. It means that no throttle is set.

An example of the **`chvdisk`** command is shown in Example 7-17.

*Example 7-17 The chvdisk command*

```
IBM_Storwize:ITSO:superuser>chvdisk -rate 20 -unitmb volume_7
IBM_Storwize:ITSO:superuser>chvdisk -warning 85% volume_7
```

**New name first:** The **`chvdisk`** command specifies the new name first. The name can consist of letters A - Z and a - z, numbers 0 - 9, the dash (-), and the underscore (_). It can be 1 - 63 characters. However, it cannot start with a number, dash, or the word "`vdisk`," because this prefix is reserved for IBM Spectrum Virtualize system assignment only.

The first command changes the volume throttling of `volume_7` to `20` MBps. The second command changes the thin-provisioned volume warning to `85%`. To verify the changes, issue the **`lsvdisk`** command, as shown in Example 7-18.

*Example 7-18 The lsvdisk command: Verifying throttling*

```
IBM_Storwize:ITSO:superuser>lsvdisk volume_7
id 1
name volume_7
```

```
IO_group_id 0
IO_group_name io_grp0
status online
mdisk_grp_id 0
mdisk_grp_name STGPool_DS3500-1
capacity 10.00GB
type striped
formatted no
mdisk_id
mdisk_name
FC_id
FC_name
RC_id
RC_name
vdisk_UID 6005076801AF813F100000000000001F
virtual_disk_throttling (MB) 20
preferred_node_id 2
fast_write_state empty
cache readwrite
udid
fc_map_count 0
sync_rate 50
copy_count 1
se_copy_count 1
filesystem
mirror_write_priority latency

copy_id 0
status online
sync yes
primary yes
mdisk_grp_id 0
mdisk_grp_name STGPool_DS3500-1
type striped
mdisk_id
mdisk_name
fast_write_state empty
used_capacity 0.41MB
real_capacity 2.02GB
free_capacity 2.02GB
overallocation 496
autoexpand on
warning 85
grainsize 32
se_copy yes
easy_tier on
easy_tier_status inactive
tier generic_ssd
tier_capacity 0.00MB
tier generic_hdd
tier_capacity
2.02GB
```

### 7.10.10 Deleting a volume

When this command is run on an existing fully managed mode volume, any data that remained on it is lost. The extents that made up this volume are returned to the pool of free extents that are available in the storage pool.

If any remote copy, IBM FlashCopy, or host mappings still exist for this volume, the delete fails unless the **-force** flag is specified. This flag ensures the deletion of the volume and any volume to host mappings and copy mappings.

If the volume is the subject of a “migrate to image mode” process, the delete fails unless the **-force** flag is specified. This flag halts the migration and then deletes the volume.

If the command succeeds (without the **-force** flag) for an image mode volume, the underlying back-end controller logical unit is consistent with the data that a host might previously read from the image mode volume. That is, all fast write data was flushed to the underlying LUN. If the **-force** flag is used, consistency is not guaranteed.

If any non-destaged data exists in the fast write cache for this volume, the deletion of the volume fails unless the **-force** flag is specified. Now, any non-destaged data in the fast write cache is deleted.

Use the **rmvdisk** command to delete a volume from your IBM Spectrum Virtualize system configuration, as shown in Example 7-19.

*Example 7-19 rmvdisk command*

```
IBM_Storwize:ITSO:superuser>rmvdisk volume_A
```

This command deletes the `volume_A` volume from the IBM Spectrum Virtualize configuration. If the volume is assigned to a host, you must use the **-force** flag to delete the volume, as shown in Example 7-20.

*Example 7-20 rmvdisk -force command*

```
IBM_Storwize:ITSO:superuser>rmvdisk -force volume_A
```

### 7.10.11 Using volume protection

To prevent active volumes or host mappings from being deleted inadvertently, the system supports a global setting that prevents these objects from being deleted if the system detects that they have had recent I/O activity.

Use the **chsystem** command to set an interval at which the volume must be idle before it can be deleted from the system. Any command that is affected by this setting fails. If the **-force** parameter is used, and the volume has not been idle for the specified interval, the deletion fails.

The following commands are affected by this setting:

- **rmvdisk**
- **rmvolume**
- **rmvdiskcopy**
- **rmvdiskhostmap**
- **rmmdiskgrp**
- **rmhostiogrp**
- **rmhost**

- **rmhostport**

To enable volume protection by setting an inactive interval to prevent deletion of volumes, complete these steps:

Issue **svctask chsystem -vdiskprotectionenabled yes -vdiskprotectiontime 60**. The parameter **-vdiskprotectionenabled yes** enables volume protection and the **-vdiskprotectiontime** parameter indicates how long a volume must be inactive before it can be deleted. Volumes can only be deleted if they have been inactive for over 60 minutes.

To disable volume protection, complete the following step:

Issue **svctask chsystem -vdiskprotectionenabled no**.

## 7.10.12 Expanding a volume

Expanding a volume presents a larger capacity disk to your operating system. Although this expansion can be easily performed by using the IBM Spectrum Virtualize system, you must ensure that your operating systems support expansion before this function is used.

Assuming that your operating systems support expansion, you can use the **expandvdisksize** command to increase the capacity of a volume, as shown in Example 7-21.

*Example 7-21 expandvdisksize command*

```
IBM_Storwize:ITSO:superuser>expandvdisksize -size 5 -unit gb volume_C
```

This command expands the `volume_C` volume (which was 35 GB) by another 5 GB to give it a total size of 40 GB.

To expand a thin-provisioned volume, you can use the **-rsize** option, as shown in Example 7-22. This command changes the real size of the `volume_B` volume to a real capacity of 55 GB. The capacity of the volume is unchanged.

*Example 7-22 The* **lsvdisk** *command*

```
IBM_Storwize:ITSO:superuser>lsvdisk volume_B
id 26
capacity 100.00GB
type striped
.
.
copy_id 0
status online
used_capacity 0.41MB
real_capacity 50.02GB
free_capacity 50.02GB
overallocation 199
autoexpand on
warning 80
grainsize 32
se_copy yes
IBM_Storwize:ITSO:superuser>expandvdisksize -rsize 5 -unit gb volume_B
IBM_Storwize:ITSO:superuser>lsvdisk volume_B
id 26
name volume_B
capacity 100.00GB
type striped
```

```
.
.
copy_id 0
status online
used_capacity 0.41MB
real_capacity 55.02GB
free_capacity 55.02GB
overallocation 181
autoexpand on
warning 80
grainsize 32
se_copy yes
```

**Important:** If a volume is expanded, its type becomes striped even if it was previously sequential or in image mode. If there are not enough extents to expand your volume to the specified size, you receive the following error message:

```
CMMVC5860E Ic_failed_vg_insufficient_virtual_extents
```

### 7.10.13 Assigning a volume to a host

Use the **mkvdiskhostmap** command to map a volume to a host. When run, this command creates a mapping between the volume and the specified host, which presents this volume to the host as though the disk was directly attached to the host. It is only after this command is run that the host can perform I/O to the volume. Optionally, a SCSI LUN ID can be assigned to the mapping.

When the HBA on the host scans for devices that are attached to it, the HBA discovers all of the volumes that are mapped to its FC ports. When the devices are found, each one is allocated an identifier (SCSI LUN ID).

For example, the first disk that is found is generally SCSI LUN 1. You can control the order in which the HBA discovers volumes by assigning the SCSI LUN ID, as required. If you do not specify a SCSI LUN ID, the system automatically assigns the next available SCSI LUN ID, based on any mappings that exist with that host.

By using the volume and host definition that we created in the previous sections, we assign volumes to hosts that are ready for their use. We use the **mkvdiskhostmap** command, as shown in Example 7-23.

*Example 7-23 The mkvdiskhostmap command*

```
IBM_Storwize:ITSO:superuser>mkvdiskhostmap -host Almaden  volume_B
Virtual Disk to Host map, id [0], successfully created
IBM_Storwize:ITSO:superuser>mkvdiskhostmap -host Almaden volume_C
Virtual Disk to Host map, id [1], successfully created
```

This command displays `volume_B` and `volume_C` that are assigned to host `Almaden`, as shown in Example 7-24.

*Example 7-24 The lshostvdiskmap -delim command*

```
IBM_Storwize:ITSO:superuser>lshostvdiskmap -delim :
id:name:SCSI_id:vdisk_id:vdisk_name:vdisk_UID
2:Almaden:0:26:volume_B:6005076801AF813F1000000000000020
2:Almaden:1:27:volume_C:6005076801AF813F1000000000000021
```

> **Assigning a specific LUN ID to a volume:** The optional **`-scsi scsi_num`** parameter can help assign a specific LUN ID to a volume that is to be associated with a host. The default (if nothing is specified) is to increment based on what is already assigned to the host.

Certain HBA device drivers stop when they find a gap in the SCSI LUN IDs, as shown in the following examples:

- Volume 1 is mapped to Host 1 with SCSI LUN ID 1.
- Volume 2 is mapped to Host 1 with SCSI LUN ID 2.
- Volume 3 is mapped to Host 1 with SCSI LUN ID 4.

When the device driver scans the HBA, it might stop after discovering volumes 1 and 2 because no SCSI LUN is mapped with ID 3.

> **Important:** Ensure that the SCSI LUN ID allocation is contiguous.

It is not possible to map a volume to a host more than one time at separate LUNs (Example 7-25).

*Example 7-25 The mkvdiskhostmap command*

```
IBM_Storwize:ITSO:superuser>mkvdiskhostmap -host Siam volume_A
Virtual Disk to Host map, id [0], successfully created
```

This command maps the volume that is called `volume_A` to the host that is called `Siam`.

All tasks that are required to assign a volume to an attached host are complete.

## 7.10.14 Showing volumes to host mapping

Use the **`lshostvdiskmap`** command to show the volumes that are assigned to a specific host, as shown in Example 7-26.

*Example 7-26 The lshostvdiskmap command*

```
IBM_Storwize:ITSO:superuser>lshostvdiskmap -delim , Siam
id,name,SCSI_id,vdisk_id,vdisk_name,wwpn,vdisk_UID
3,Siam,0,0,volume_A,210000E08B18FF8A,60050768018301BF280000000000000C
```

From this command, you can see that the host `Siam` has only one assigned volume that is called `volume_A`. The SCSI LUN ID is also shown, which is the ID by which the volume is presented to the host. If no host is specified, all defined host-to-volume mappings are returned.

> **Specifying the flag before the host name:** Although the **`-delim`** flag normally comes at the end of the command string, in this case, you must specify this flag before the host name. Otherwise, it returns the following message:
>
> ```
> CMMVC6070E An invalid or duplicated parameter, unaccompanied argument, or
> incorrect argument sequence has been detected. Ensure that the input is as per
> the help.
> ```

## 7.10.15 Deleting a volume to host mapping

When you are deleting a volume mapping, you are not deleting the volume, only the connection from the host to the volume. If you mapped a volume to a host by mistake or you want to reassign the volume to another host, use the **rmvdiskhostmap** command to unmap a volume from a host, as shown in Example 7-27.

*Example 7-27 The rmvdiskhostmap command*

```
IBM_Storwize:ITSO:superuser>rmvdiskhostmap -host Tiger volume_D
```

This command unmaps the volume that is called `volume_D` from the host that is called `Tiger`.

## 7.10.16 Migrating a volume

You might want to migrate volumes from one set of MDisks to another set of MDisks to decommission an old disk subsystem to have better balanced performance across your virtualized environment, or to migrate data into the IBM Spectrum Virtualize environment transparently by using image mode. For more information about migration, see Chapter 9, “Storage migration” on page 347.

**Important:** After migration is started, it continues until completion unless it is stopped or suspended by an error condition or the volume that is being migrated is deleted.

As you can see from the parameters that are shown in Example 7-28 on page 275, before you can migrate your volume, you must know the name of the volume that you want to migrate and the name of the storage pool to which you want to migrate it. To discover the names, run the **lsvdisk** and **lsmdiskgrp** commands.

After you know these details, you can run the **migratevdisk** command, as shown in Example 7-28.

*Example 7-28 migratevdisk command*

```
IBM_Storwize:ITSO:superuser>migratevdisk -mdiskgrp STGPool_DS5000-1 -vdisk volume_C
```

This command moves `volume_C` to the storage pool named `STGPool_DS5000-1`.

**Tips:** If insufficient extents are available within your target storage pool, you receive an error message. Ensure that the source MDisk group and target MDisk group have the same extent size.

By using the optional threads parameter, you can assign a priority to the migration process. The default is `4`, which is the highest priority setting. However, if you want the process to take a lower priority over other types of I/O, you can specify `3`, `2`, or `1`.

You can run the **lsmigrate** command at any time to see the status of the migration process, as shown in Example 7-29.

*Example 7-29 lsmigrate command*

```
IBM_Storwize:ITSO:superuser>lsmigrate
migrate_type MDisk_Group_Migration
progress 0
migrate_source_vdisk_index 27
```

```
migrate_target_mdisk_grp 2
max_thread_count 4
migrate_source_vdisk_copy_id 0

IBM_Storwize:ITSO:superuser>lsmigrate
migrate_type MDisk_Group_Migration
progress 76
migrate_source_vdisk_index 27
migrate_target_mdisk_grp 2
max_thread_count 4
migrate_source_vdisk_copy_id 0
```

**Progress:** The progress is shown as a percentage complete. If you receive no more replies, it means that the process finished.

## 7.10.17 Migrating a fully managed volume to an image mode volume

Migrating a fully managed volume to an image mode volume enables the IBM Spectrum Virtualize to be removed from the data path, which might be useful where the IBM Spectrum Virtualize is used as a data mover appliance. You can use the **migratetoimage** command.

To migrate a fully managed volume to an image mode volume, the following rules apply:

- The destination MDisk must be greater than or equal to the size of the volume.
- The MDisk that is specified as the target must be in an unmanaged state.
- Regardless of the mode in which the volume starts, it is reported as a managed mode during the migration.
- Both of the MDisks that are involved are reported as being image mode volumes during the migration.
- If the migration is interrupted by a system recovery or cache problem, the migration resumes after the recovery completes.

Example 7-30 shows an example of the command.

*Example 7-30 migratetoimage command*

```
IBM_Storwize:ITSO:superuser>migratetoimage -vdisk volume_A -mdisk mdisk10 -mdiskgrp
STGPool_IMAGE
```

In this example, you migrate the data from `volume_A` onto `mdisk10`, and the MDisk must be put into the `STGPool_IMAGE` storage pool.

## 7.10.18 Shrinking a volume

The **shrinkvdisksize** command reduces the capacity that is allocated to the particular volume by the amount that you specify. You cannot shrink the real size of a thin-provisioned volume to less than its used size. All capacities (including changes) must be in multiples of 512 bytes. An entire extent is reserved even if it is only partially used. The default capacity units are MBs.

You can use this command to shrink the physical capacity that is allocated to a particular volume by the specified amount. You also can use this command to shrink the virtual capacity of a thin-provisioned volume without altering the physical capacity that is assigned to the volume. Use the following parameters:

- For a non-thin-provisioned volume, use the **-size** parameter.
- For a thin-provisioned volume's real capacity, use the **-rsize** parameter.
- For the thin-provisioned volume's virtual capacity, use the **-size** parameter.

When the virtual capacity of a thin-provisioned volume is changed, the warning threshold is automatically scaled to match. The new threshold is stored as a percentage.

The system arbitrarily reduces the capacity of the volume by removing a partial extent, one extent, or multiple extents from those extents that are allocated to the volume. You cannot control which extents are removed. Therefore, you cannot assume that it is unused space that is removed.

Image mode volumes cannot be reduced in size. Instead, they must first be migrated to fully managed mode. To run the **shrinkvdisksize** command on a mirrored volume, all copies of the volume must be synchronized.

> **Important:** Consider the following guidelines when you are shrinking a disk:
>
> - If the volume contains data, do not shrink the disk.
> - Certain operating systems or file systems use the outer edge of the disk for performance reasons. This command can shrink a FlashCopy target volume to the same capacity as the source.
> - Before you shrink a volume, validate that the volume is not mapped to any host objects. If the volume is mapped, data is displayed. You can determine the exact capacity of the source or master volume by issuing the **svcinfo lsvdisk -bytes vdiskname** command. Shrink the volume by the required amount by issuing the **shrinkvdisksize -size disk_size -unit b | kb | mb | gb | tb | pb vdisk_name | vdisk_id** command.

Assuming that your operating system supports it, you can use the **shrinkvdisksize** command to decrease the capacity of a volume, as shown in Example 7-31.

*Example 7-31 shrinkvdisksize command*

```
IBM_Storwize:ITSO:superuser>shrinkvdisksize -size 44 -unit gb volume_D
```

This command shrinks a volume that is called `volume_D` from a total size of 80 GB by 44 GB, to a new total size of 36 GB.

## 7.10.19 Showing a volume on an MDisk

Use the **lsmdiskmember** command to display information about the volume that is using space on a specific MDisk, as shown in Example 7-32.

*Example 7-32 lsmdiskmember command*

```
IBM_Storwize:ITSO:superuser>lsmdiskmember mdisk8
id copy_id
24 0
27 0
```

This command displays a list of all of the volume IDs that correspond to the volume copies that use `mdisk8`.

To correlate the IDs that are displayed in this output to volume names, we can run the **lsvdisk** command.

## 7.10.20 Showing which volumes are using a storage pool

Use the **lsvdisk -filtervalue** command to see which volumes are part of a specific storage pool, as shown in Example 7-33. This command shows all of the volumes that are part of the storage pool that is named STGPool_DS3500-2.

*Example 7-33 lsvdisk -filtervalue command: VDisks in the managed disk group (MDG)*

```
IBM_Storwize:ITSO:superuser>lsvdisk -filtervalue mdisk_grp_name=STGPool_DS3500-2 -delim ,
id,name,IO_group_id,IO_group_name,status,mdisk_grp_id,mdisk_grp_name,capacity,type,FC_id,FC
_name,RC_id,RC_name,vdisk_UID,fc_map_count,copy_count,fast_write_state,se_copy_count,RC_cha
nge
7,W2K3_SRV2_VOL01,0,io_grp0,online,1,STGPool_DS3500-2,10.00GB,striped,,,,,6005076801AF813F1
000000000000008,0,1,empty,0,0,no
8,W2K3_SRV2_VOL02,0,io_grp0,online,1,STGPool_DS3500-2,10.00GB,striped,,,,,6005076801AF813F1
000000000000009,0,1,empty,0,0,no
9,W2K3_SRV2_VOL03,0,io_grp0,online,1,STGPool_DS3500-2,10.00GB,striped,,,,,6005076801AF813F1
00000000000000A,0,1,empty,0,0,no
10,W2K3_SRV2_VOL04,0,io_grp0,online,1,STGPool_DS3500-2,10.00GB,striped,,,,,6005076801AF813F
100000000000000B,0,1,empty,0,0,no
11,W2K3_SRV2_VOL05,0,io_grp0,online,1,STGPool_DS3500-2,10.00GB,striped,,,,,6005076801AF813F
100000000000000C,0,1,empty,0,0,no
12,W2K3_SRV2_VOL06,0,io_grp0,online,1,STGPool_DS3500-2,10.00GB,striped,,,,,6005076801AF813F
100000000000000D,0,1,empty,0,0,no
16,AIX_SRV2_VOL01,0,io_grp0,online,1,STGPool_DS3500-2,20.00GB,striped,,,,,6005076801AF813F1
000000000000011,0,1,empty,0,0,no
```

## 7.10.21 Showing which MDisks are used by a specific volume

Use the **lsvdiskmember** command to show from which MDisks a specific volume's extents came, as shown in Example 7-34.

*Example 7-34 The lsvdiskmember command*

```
IBM_Storwize:ITSO:superuser>lsvdiskmember 0
id
4
5
6
7
```

If you want to know more about these MDisks, you can run the **lsmdisk** command (by using the ID that is displayed in Example 7-34 rather than the name).

## 7.10.22 Showing from which storage pool a volume has its extents

Use the **lsvdisk** command to show to which storage pool a specific volume belongs, as shown in Example 7-35.

*Example 7-35 lsvdisk command: Storage pool name*

```
IBM_Storwize:ITSO:superuser>lsvdisk Volume_D
id 25
name Volume_D
IO_group_id 0
IO_group_name io_grp0
status online
mdisk_grp_id 0
mdisk_grp_name STGPool_DS3500-1
capacity 10.00GB
type striped
formatted no
mdisk_id
mdisk_name
FC_id
FC_name
RC_id
RC_name
vdisk_UID 6005076801AF813F100000000000001E
throttling 0
preferred_node_id 1
fast_write_state empty
cache readwrite
udid
fc_map_count 0
sync_rate 50
copy_count 1
se_copy_count 1
filesystem
mirror_write_priority latency

copy_id 0
status online
sync yes
primary yes
mdisk_grp_id 0
mdisk_grp_name STGPool_DS3500-1
type striped
mdisk_id
mdisk_name
fast_write_state empty
used_capacity 0.41MB
real_capacity 2.02GB
free_capacity 2.02GB
overallocation 496
autoexpand on
warning 80
grainsize 32
se_copy yes
easy_tier on
easy_tier_status inactive
tier generic_ssd
tier_capacity 0.00MB
tier generic_hdd
tier_capacity 2.02GB
```

To learn more about these storage pools, you can run the **lsmdiskgrp** command, as described in Chapter 6, “Storage pools” on page 165.

### 7.10.23 Showing the host to which the volume is mapped

To show the hosts to which a specific volume was assigned, run the **lsvdiskhostmap** command, as shown in Example 7-36.

*Example 7-36 lsvdiskhostmap command*

```
IBM_Storwize:ITSO:superuser>lsvdiskhostmap -delim , volume_B
id,name,SCSI_id,host_id,host_name,vdisk_UID
26,volume_B,0,2,Almaden,6005076801AF813F1000000000000020
```

This command shows the host or hosts to which the `volume_B` volume was mapped. Duplicate entries are normal because more paths exist between the clustered system and the host. To be sure that the operating system on the host sees the disk only one time, you must install and configure a multipath software application, such as IBM Subsystem Device Driver (SDD).

> **Specifying the -delim flag:** Although the optional **-delim** flag normally comes at the end of the command string, you must specify this flag before the volume name in this case. Otherwise, the command does not return any data.

### 7.10.24 Showing the volume to which the host is mapped

To show the volume to which a specific host was assigned, run the **lshostvdiskmap** command, as shown in Example 7-37.

*Example 7-37 lshostvdiskmap command example*

```
IBM_Storwize:ITSO:superuser>lshostvdiskmap -delim , Almaden
id,name,SCSI_id,vdisk_id,vdisk_name,vdisk_UID
2,Almaden,0,26,volume_B,60050768018301BF2800000000000005
2,Almaden,1,27,volume_A,60050768018301BF2800000000000004
```

This command shows which volumes are mapped to the host called `Almaden`.

> **Specifying the -delim flag:** Although the optional **-delim** flag normally comes at the end of the command string, you must specify this flag before the volume name in this case. Otherwise, the command does not return any data.

### 7.10.25 Tracing a volume from a host back to its physical disk

In many cases, you must verify exactly which physical disk is presented to the host, for example, from which storage pool a specific volume comes. However, from the host side, it is not possible for the server administrator who is using the GUI to see on which physical disks the volumes are running. Instead, you must enter the command that is shown in Example 7-38 from your multipath command prompt.

Complete the following steps:

1. On your host, run the **datapath query device** command. You see a long disk serial number for each vpath device, as shown in Example 7-38.

*Example 7-38 datapath query device command*

```
DEV#:   0  DEVICE NAME: Disk1 Part0  TYPE: 2145       POLICY: OPTIMIZED
SERIAL: 60050768018301BF2800000000000005
```

```
============================================================================
Path#            Adapter/Hard Disk       State  Mode       Select     Errors
    0     Scsi Port2 Bus0/Disk1 Part0     OPEN   NORMAL         20          0
    1     Scsi Port3 Bus0/Disk1 Part0     OPEN   NORMAL       2343          0

DEV#:   1  DEVICE NAME: Disk2 Part0  TYPE: 2145       POLICY: OPTIMIZED
SERIAL: 60050768018301BF2800000000000004
============================================================================
Path#            Adapter/Hard Disk       State  Mode       Select     Errors
    0     Scsi Port2 Bus0/Disk2 Part0     OPEN   NORMAL       2335          0
    1     Scsi Port3 Bus0/Disk2 Part0     OPEN   NORMAL          0          0

DEV#:   2  DEVICE NAME: Disk3 Part0  TYPE: 2145       POLICY: OPTIMIZED
SERIAL: 60050768018301BF2800000000000006
============================================================================
Path#            Adapter/Hard Disk       State  Mode       Select     Errors
    0     Scsi Port2 Bus0/Disk3 Part0     OPEN   NORMAL       2331          0
    1     Scsi Port3 Bus0/Disk3 Part0     OPEN   NORMAL          0          0
```

> **State:** In Example 7-38, the state of each path is OPEN. Sometimes, the state is CLOSED. This state does not necessarily indicate a problem because it might be a result of the path's processing stage.

2. Run the **`lshostvdiskmap`** command to return a list of all assigned volumes, as shown in Example 7-39.

   *Example 7-39 lshostvdiskmap command*

   ```
   IBM_Storwize:ITSO:superuser>lshostvdiskmap -delim , Almaden
   id,name,SCSI_id,vdisk_id,vdisk_name,vdisk_UID
   2,Almaden,0,26,volume_B,60050768018301BF2800000000000005
   2,Almaden,1,27,volume_A,60050768018301BF2800000000000004
   2,Almaden,2,28,volume_C,60050768018301BF2800000000000006
   ```

   Look for the disk serial number that matches your **`datapath query device`** output. This host was defined in our IBM Spectrum Virtualize system as `Almaden`.

3. Run the **`lsvdiskmember`** ***`vdiskname`*** command for the MDisk or a list of the MDisks that make up the specified volume, as shown in Example 7-40.

   *Example 7-40 lsvdiskmember command*

   ```
   IBM_Storwize:ITSO:superuser>lsvdiskmember volume_E
   id
   0
   1
   2
   3
   4
   10
   11
   13
   15
   16
   17
   ```

4. Query the MDisks with the **`lsmdisk mdiskID`** command to discover their controller and LUN information, as shown in Example 7-41. The output displays the controller name and the controller LUN ID to help you to track back to a LUN within the disk subsystem (if you gave your controller a unique name, such as a serial number). See Example 7-41.

*Example 7-41 lsmdisk command*

```
IBM_Storwize:ITSO:superuser>lsmdisk 0
id 0
name mdisk0
status online
mode managed
mdisk_grp_id 0
mdisk_grp_name STGPool_DS3500-1
capacity 128.0GB
quorum_index 1
block_size 512
controller_name ITSO-DS3500
ctrl_type 4
ctrl_WWNN 20080080E51B09E8
controller_id 2
path_count 4
max_path_count 4
ctrl_LUN_# 0000000000000000
UID 60080e50001b0b62000007b04e731e4d00000000000000000000000000000000
preferred_WWPN 20580080E51B09E8
active_WWPN 20580080E51B09E8
fast_write_state empty
raid_status
raid_level
redundancy
strip_size
spare_goal
spare_protection_min
balanced
tier generic_hdd
```

8

# Hosts

This chapter describes the host configuration procedures that are required to attach supported hosts to the IBM Spectrum Virtualize system. The chapter also introduces new concepts about Host Clusters, and N-Port Virtualization ID (NPIV) support from a hosts perspective.

This chapter describes the following topics:

- 8.1, “Host attachment overview” on page 284
- 8.2, “Host Clusters” on page 285
- 8.3, “N-Port Virtualization ID (NPIV) Support” on page 285
- 8.4, “Hosts operations using the GUI” on page 293
- 8.5, “Hosts operations using CLI” on page 336

## 8.1 Host attachment overview

IBM Spectrum Virtualize system supports a wide range of host types (both IBM and non-IBM), which makes it possible to consolidate storage in an open systems environment into a common pool of storage. Then, you can use and manage the storage pool more efficiently as a single entity from a central point on the storage area network (SAN).

The ability to consolidate storage for attached open systems hosts provides the following benefits:

- Unified, easier storage management
- Increased utilization rate of the installed storage capacity
- Advanced Copy Services functions offered across storage systems from separate vendors
- Consider only one kind of multipath driver for attached hosts

Hosts can be connected to IBM Spectrum Virtualize system using any of the following protocols:

- Fibre Channel (FC)
- Fibre Channel over Ethernet (FCoE)
- Internet Small Computer System Interface (iSCSI)

For hosts connecting to the IBM Spectrum Virtualize system using the fabric switches using FC or FCoE protocol, they need to be zoned appropriately as indicated in 3.3.2, “SAN zoning and SAN connections” on page 50.

For hosts connecting to the IBM Spectrum Virtualize system iSCSI protocol, they need to be configured appropriately as indicated in 3.3.3, “iSCSI IP addressing plan” on page 52.

> **Note:** Certain host operating systems can be directly connected to the IBM Spectrum Virtualize system without the need for FC fabric switches. Check the IBM System Storage Interoperation Center at
> https://www.ibm.com/systems/support/storage/ssic/interoperability.wss

For better load balancing and redundancy on the host side, the use of a host multipathing driver is strongly recommended. A host multipathing I/O driver is required in the following situations:

- To protect a host from fabric link failures including port failures on the IBM Spectrum Virtualize system nodes
- To protect a host from an HBA failure (if two HBAs are in use)
- To protect a host from fabric failures if the host is connected through two HBAs to two separate fabrics.
- To provide load balancing on the server’s HBA.

To learn about various host operating systems and versions supported with Storwize V7000 refer to the IBM System Storage Interoperability Center (SSIC) at:

https://www.ibm.com/systems/support/storage/ssic/interoperability.wss

To learn about how to attach various supported host operating systems to IBM Storwize V7000 refer to the IBM Knowledge Center for Storwize V7000 at:

https://ibm.biz/BdsKhv

If the desired host operating system is not in SSIC, then an IBM representative can be requested to submit a special request for support via Storage Customer Opportunity REquest (SCORE) tool for evaluation at:

`https://ibm.biz/Bdsr6P`

## 8.2 Host Clusters

IBM Spectrum Virtualize software supports host clusters starting with V7.7.1 onwards. The host cluster allows a user to create a group of hosts to form a cluster, which is treated as one single entity, thus allowing multiple hosts to have access to the same volumes.

Volumes mapped to that host cluster will be assigned to all members of the host cluster with the same SCSI ID.

A typical use-case is to define a host cluster containing all the WWPNs belonging to the hosts participating in a host operating system based cluster, such as IBM PowerHA or Microsoft Cluster Server (MSCS) and such.

The following new commands have been added to deal with host clusters:

- **`lshostcluster`**
- **`lshostclustermember`**
- **`lshostclustervolumemap`**
- **`mkhost`** (modified to put host in a host cluster on creation)
- **`rmhostclustermember`**
- **`rmhostcluster`**
- **`rmvolumehostclustermap`**

**Note:** Host clusters allows for the creation of individual hosts and adding them to a host cluster. Care must be taken to make sure that no loss of access occurs when transitioning to host clusters.

**CLI only at the time of writing:** As of this writing, host cluster operations have not yet been incorporated into the Spectrum Virtualize GUI, however, they can be carried out using the CLI as described in 8.5.5, “Host Cluster Operations” on page 343.

## 8.3 N-Port Virtualization ID (NPIV) Support

The usage model for all Spectrum Virtualize products is based around two-way active/active node models. That is a pair of distinct control modules that share active/active access for a given volume. These nodes each have their own Fibre Channel WWNN, and thus all ports presented from each node have a set of WWPNs that are presented to the fabric.

Traditionally, should one node fail or be removed for some reason, the paths presented for volumes from that node would go offline, and it is up to the native OS multipathing software to failover from using both sets of WWPN to just those that remain online. While this is exactly what multipathing software is designed to do, occasionally it can be problematic, particularly if paths are not seen as coming back online for some reason.

Starting with Spectrum Virtualize V7.7.0, IBM Spectrum Virtualize system can be enabled in NPIV mode. When NPIV mode is enabled on the IBM Spectrum Virtualize system, ports do not come up until they are ready to service I/O, which improves host behavior around node unpends. In addition, path failures due to an offline node are masked from host multipathing.

When NPIV is enabled on IBM Spectrum Virtualize system nodes, each physical WWPN will report up to three virtual WWPNs as shown in Table 8-1.

*Table 8-1 Spectrum Virtualize NPIV Ports's*

| NPIV port | Port description |
|---|---|
| Primary NPIV Port | This is the WWPN that will communicate with backend storage, and may be used for node to node traffic. (Local or remote) |
| Primary Host Attach Port | This is the WWPN that will communicate with hosts. It is a target port only, and this is the primary port, therefore represents this local nodes WWNN. |
| Failover Host Attach Port | This is a standby WWPN that will communicate with hosts and will only be brought online on this node, if the partner node in this I/O Group goes offline. This will be the same as the Primary Host Attach WWPN on the partner node. |

Figure 8-1 on page 286 depicts the three WWPNs associated with a Storwize V7000 port when NPIV is enabled.

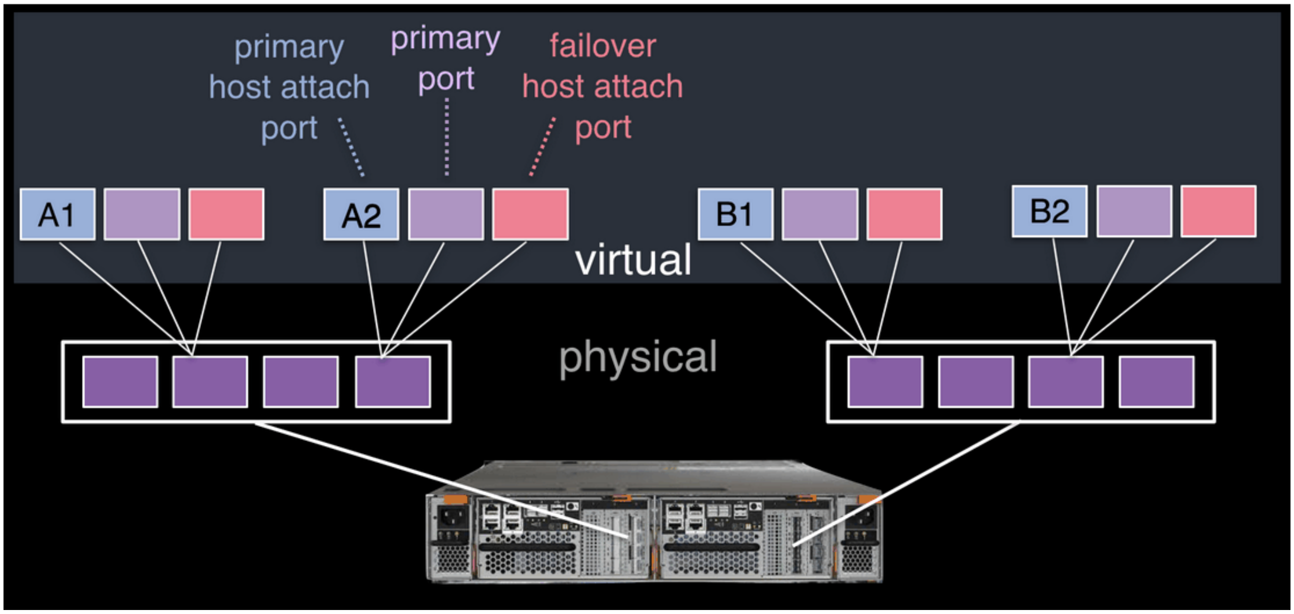


*Figure 8-1 Allocation of NPIV virtual WWPN ports per physical port*

The failover host attach port (in pink) is not active at this time. Figure 8-2 shows what happens when the second node fails. Subsequent to the node failure, the failover host attach ports on the remaining node are active and have taken on the WWPN of the failed node's primary host attach port.

> **Note:** The picture shows only two ports per node in detail, but the same applies for all physical ports.

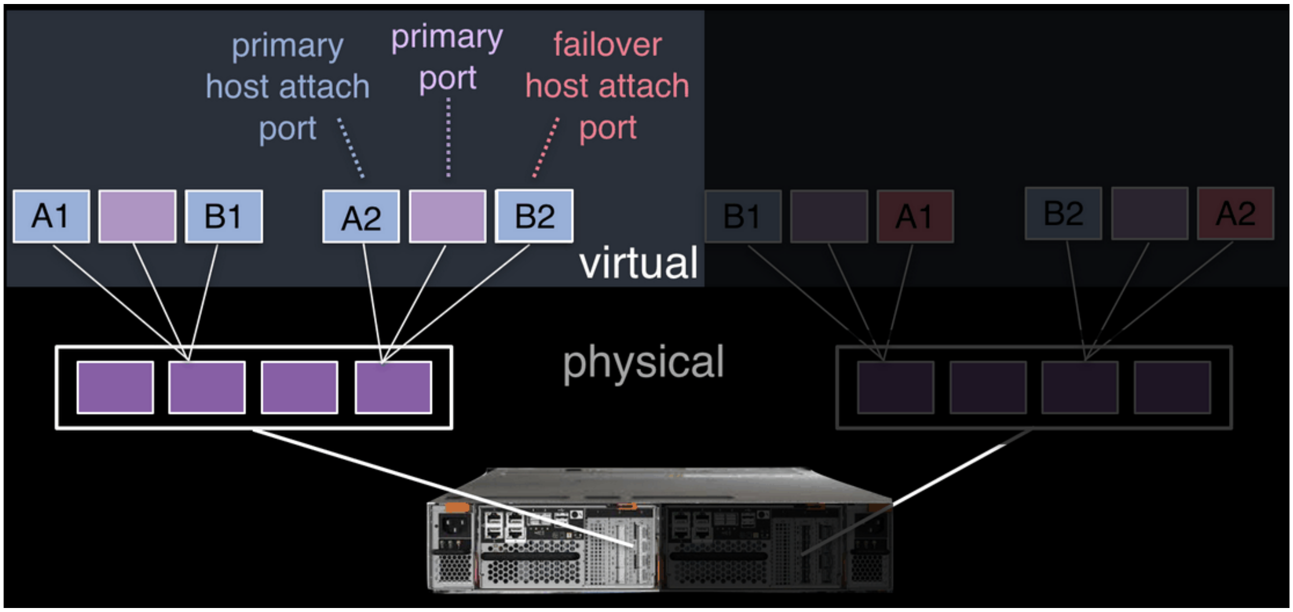


*Figure 8-2 Allocation of NPIV virtual WWPN ports per physical port after a node failure*

With V7.7.0 onwards, this all happens automatically when NPIV is enabled at a system level in Storwize V7000. At this time, the failover only happens automatically between the two nodes in an I/O Group.

There is a transitional mode for backward compatibility during the transition period as well.

The processes for enabling NPIV on a new Spectrum Virtualize system is slightly different than on an existing system. Refer to the Knowledge Center for more details at:

`https://ibm.biz/BdsKgx`

> **Note:** NPIV is only supported for Fibre Channel protocol. It is not supported for FCoE protocol.

## 8.3.1 NPIV Pre-requisites

The following key points should be considered for NPIV enablement:

- For NPIV enablement, the IBM Spectrum Virtualize system has to be at V7.7.0 or later
- A V7.7.0 or later system with NPIV enabled as backend storage for a system that is earlier than V7.7.0 is not supported.
- Both nodes in an IO group should have identical hardware to allow failover to work as expected.
- Fibre Channel switches must permit each physically connected IBM Spectrum Virtualize system port the ability to create two additional NPIV ports.

## 8.3.2 Enabling NPIV on a new system

For V7.7.0 onwards, a new Spectrum Virtualize system should have NPIV enabled by default. For any case, where it is not enabled by default and NPIV is desired then, NPIV can be enabled on a new Spectrum Virtualize system by executing the following steps:

1. Execute the **lsiogrp** command to list the I/O groups present in a system as shown in Example 8-1.

*Example 8-1 Listing the I/O groups in a system*

```
IBM_Storwize:ITSO:superuser>lsiogrp
id name            node_count vdisk_count host_count site_id site_name
0  io_grp0         2          2           2
1  io_grp1         0          0           2
2  io_grp2         0          0           2
3  io_grp3         0          0           2
4  recovery_io_grp 0          0           0
```

2. Execute the **lsiogrp** command to view the status of N_Port ID Virtualization (NPIV) as shown in Example 8-2.

*Example 8-2 Checking NPIV mode via fctargetportmode field*

```
IBM_Storwize:ITSO:superuser>lsiogrp 0
id 0
name io_grp0
node_count 2
vdisk_count 2
host_count 2
flash_copy_total_memory 20.0MB
flash_copy_free_memory 20.0MB
remote_copy_total_memory 20.0MB
remote_copy_free_memory 20.0MB
mirroring_total_memory 20.0MB
mirroring_free_memory 20.0MB
raid_total_memory 40.0MB
raid_free_memory 38.8MB
maintenance no
compression_active yes
accessible_vdisk_count 2
compression_supported yes
max_enclosures 10
encryption_supported no
flash_copy_maximum_memory 552.0MB
site_id
site_name
fctargetportmode enabled
compression_total_memory 2047.9MB
```

3. If the resulting output is *fctargetportnode*:**enabled** as shown in Example 8-2, then NPIV is enabled.
4. The virtual WWPNs can be listed using the **lstargetportfc** command as shown in Example 8-3.

*Example 8-3 Listing the virtual WWPNs*

```
IBM_Storwize:ITSO:superuser>lstargetportfc
id WWPN             WWNN             port_id owning_node_id current_node_id nportid
host_io_permitted virtualized
1  50050768021000EF 50050768020000EF 1       1              1               010200  no
no
```

```
2  5005076802990OEF 50050768020000EF 1     1          1          010201  yes
yes
3  50050768022000EF 50050768020000EF 2     1          1          020200  no
no
4  5005076802A900EF 50050768020000EF 2     1          1          020201  yes
yes
5  50050768023000EF 50050768020000EF 3     1          1          0A83C0  no
no
6  5005076802B900EF 50050768020000EF 3     1          1          0A83C1  yes
yes
7  50050768024000EF 50050768020000EF 4     1          1          011400  no
no
8  5005076802C900EF 50050768020000EF 4     1          1          011401  yes
yes
33 50050768021000F0 50050768020000F0 1     2          2          010300  no
no
34 50050768029900F0 50050768020000F0 1     2          2          010301  yes
yes
35 50050768022000F0 50050768020000F0 2     2          2          020300  no
no
36 5005076802A900F0 50050768020000F0 2     2          2          020301  yes
yes
37 50050768023000F0 50050768020000F0 3     2          2          011500  no
no
38 5005076802B900F0 50050768020000F0 3     2          2          011501  yes
yes
39 50050768024000F0 50050768020000F0 4     2          2          0A82C0  no
no
40 5005076802C900F0 50050768020000F0 4     2          2          0A82C1  yes
yes
```

5. At this point you can configure your zones for hosts using the primary host attach ports (virtual WWPNs) of the Spectrum Virtualize ports as shown in **bold** in the output of Example 8-3.
6. If the status of “*fctargetportmode*” is **disabled**, run the **chiogrp** command to get into **transitional** mode for NPIV as shown in Example 8-4.

*Example 8-4 Change the NPIV mode to transitional*

```
IBM_Storwize:ITSO:superuser>chiogrp -fctargetportmode transitional 0
```

7. The ***transitional*** mode can be verified using the **lsiogrp** command as shown in Example 8-5.

*Example 8-5 NPIV transitional mode*

```
IBM_Storwize:ITSO:superuser>lsiogrp 0
id 0
name io_grp0
node_count 2
vdisk_count 2
host_count 2
flash_copy_total_memory 20.0MB
flash_copy_free_memory 20.0MB
remote_copy_total_memory 20.0MB
remote_copy_free_memory 20.0MB
mirroring_total_memory 20.0MB
mirroring_free_memory 20.0MB
raid_total_memory 40.0MB
raid_free_memory 38.8MB
maintenance no
```

```
compression_active yes
accessible_vdisk_count 2
compression_supported yes
max_enclosures 10
encryption_supported no
flash_copy_maximum_memory 552.0MB
site_id
site_name
fctargetportmode transitional
compression_total_memory 2047.9MB
```

8. In ***transitional*** mode, host I/O is permitted on primary ports and primary host attach ports (virtual WWPN) as shown in Example 8-6 under "**host_io_permitted**" column.

*Example 8-6 WWPNs in transitional mode*

```
IBM_Storwize:ITSO:superuser>lstargetportfc
id WWPN             WWNN             port_id owning_node_id current_node_id nportid
host_io_permitted virtualized
1  50050768021000EF 50050768020000EF 1     1           1           010200  yes
no
2  50050768029900EF 50050768020000EF 1     1           1           010201  yes
yes
3  50050768022000EF 50050768020000EF 2     1           1           020200  yes
no
4  5005076802A900EF 50050768020000EF 2     1           1           020201  yes
yes
5  50050768023000EF 50050768020000EF 3     1           1           0A83C0  yes
no
6  5005076802B900EF 50050768020000EF 3     1           1           0A83C1  yes
yes
7  50050768024000EF 50050768020000EF 4     1           1           011400  yes
no
8  5005076802C900EF 50050768020000EF 4     1           1           011401  yes
yes
33 50050768021000F0 50050768020000F0 1     2           2           010300  yes
no
34 50050768029900F0 50050768020000F0 1     2           2           010301  yes
yes
35 50050768022000F0 50050768020000F0 2     2           2           020300  yes
no
36 5005076802A900F0 50050768020000F0 2     2           2           020301  yes
yes
37 50050768023000F0 50050768020000F0 3     2           2           011500  yes
no
38 5005076802B900F0 50050768020000F0 3     2           2           011501  yes
yes
39 50050768024000F0 50050768020000F0 4     2           2           0A82C0  yes
no
40 5005076802C900F0 50050768020000F0 4     2           2           0A82C1  yes
yes
```

9. Enable NPIV by changing the mode from ***transitional*** to ***enabled*** as shown in Example 8-7.

*Example 8-7 Enabling NPIV*

```
IBM_Storwize:ITSO:superuser>chiogrp -fctargetportmode enabled 0
```

10. NPIV enablement can be verified by checking the **fctargetportmode** field as shown in Example 8-8 on page 291.

*Example 8-8 NPIV enablement verification*

```
IBM_Storwize:ITSO:superuser>lsiogrp 0
id 0
name io_grp0
node_count 2
vdisk_count 2
host_count 2
flash_copy_total_memory 20.0MB
flash_copy_free_memory 20.0MB
remote_copy_total_memory 20.0MB
remote_copy_free_memory 20.0MB
mirroring_total_memory 20.0MB
mirroring_free_memory 20.0MB
raid_total_memory 40.0MB
raid_free_memory 38.8MB
maintenance no
compression_active yes
accessible_vdisk_count 2
compression_supported yes
max_enclosures 10
encryption_supported no
flash_copy_maximum_memory 552.0MB
site_id
site_name
fctargetportmode enabled
compression_total_memory 2047.9MB
```

At this point you can configure your zones for hosts using the primary host attach ports (virtual WWPNs) of the Spectrum Virtualize ports as shown in **bold** in the output of Example 8-3.

### 8.3.3 Enabling NPIV on an existing system

When IBM Spectrum Virtualize systems running code prior to V7.7.1, are upgraded to V7.7.1 or higher, then by default, the NPIV feature is not turned on as it may require changes to host side zoning.

Enabling N_Port ID Virtualization (NPIV) on an existing system requires that you complete the following steps after meeting the pre-requisites.

1. Audit your SAN fabric layout and zoning rules as NPIV has stricter requirements. Ensure that equivalent ports are on the same fabric and in the same zone. For more information, see the topic about zoning considerations for N_Port ID Virtualization in IBM Knowledge Center at:

   `https://ibm.biz/BdsKhA`

2. Check the path count between your hosts and the IBM Spectrum Virtualize system to ensure that the number of paths is half of the usual supported maximum. For more information, see the topic about zoning considerations for N_Port ID Virtualization in IBM Knowledge Center at:

   `https://ibm.biz/BdsKhA`

3. Execute **`lstargetportfc`** command to note down the primary host attach WWPNs (virtual WWPNs) as shown in **bold** in the Example 8-9.

*Example 8-9* **`lstargetportfc`** *command to get primary host WWPNs (virtual WWPNs)*

```
IBM_Storwize:ITSO:superuser>lstargetportfc
```

```
id WWPN             WWNN             port_id owning_node_id current_node_id nportid
host_io_permitted virtualized
1  50050768021000EF 50050768020000EF 1      1           1           010200  yes
no
2  50050768029900EF 50050768020000EF 1      1                       000000  no
yes
3  50050768022000EF 50050768020000EF 2      1           1           020200  yes
no
4  5005076802A900EF 50050768020000EF 2      1                       000000  no
yes
5  50050768023000EF 50050768020000EF 3      1           1           0A83C0  yes
no
6  5005076802B900EF 50050768020000EF 3      1                       000000  no
yes
7  50050768024000EF 50050768020000EF 4      1           1           011400  yes
no
8  5005076802C900EF 50050768020000EF 4      1                       000000  no
yes
33 50050768021000F0 50050768020000F0 1      2           2           010300  yes
no
34 50050768029900F0 50050768020000F0 1      2                       000000  no
yes
35 50050768022000F0 50050768020000F0 2      2           2           020300  yes
no
36 5005076802A900F0 50050768020000F0 2      2                       000000  no
yes
37 50050768023000F0 50050768020000F0 3      2           2           011500  yes
no
38 5005076802B900F0 50050768020000F0 3      2                       000000  no
yes
39 50050768024000F0 50050768020000F0 4      2           2           0A82C0  yes
no
40 5005076802C900F0 50050768020000F0 4      2                       000000  no
yes
```

4. Include the primary host attach ports (virtual WWPNs) to your host zones.
5. Enable transitional mode for NPIV on IBM Spectrum Virtualize system as shown in Example 8-10.

*Example 8-10 NPIV in transitional mode*

```
IBM_Storwize:ITSO:superuser>chiogrp -fctargetportmode transitional 0
```

Ensure that the primary host attach WWPNs (virtual WWPNs) now allows host traffic as shown in **bold** in Example 8-11.

*Example 8-11 Host attach WWPNs (virtual WWPNs) permitting host traffic*

```
IBM_Storwize:ITSO:superuser>lstargetportfc
id WWPN             WWNN             port_id owning_node_id current_node_id nportid
host_io_permitted virtualized
1  50050768021000EF 50050768020000EF 1      1           1           010200  yes
no
2  50050768029900EF 50050768020000EF 1      1           1           010201  yes
yes
3  50050768022000EF 50050768020000EF 2      1           1           020200  yes
no
4  5005076802A900EF 50050768020000EF 2      1           1           020201  yes
yes
```

```
5  50050768023000EF 50050768020000EF 3      1           1            0A83C0  yes
no
6  5005076802B900EF 50050768020000EF 3      1           1            0A83C1  yes
yes
7  50050768024000EF 50050768020000EF 4      1           1            011400  yes
no
8  5005076802C900EF 50050768020000EF 4      1           1            011401  yes
yes
33 50050768021000F0 50050768020000F0 1      2           2            010300  yes
no
34 50050768029900F0 50050768020000F0 1      2           2            010301  yes
yes
35 50050768022000F0 50050768020000F0 2      2           2            020300  yes
no
36 5005076802A900F0 50050768020000F0 2      2           2            020301  yes
yes
37 50050768023000F0 50050768020000F0 3      2           2            011500  yes
no
38 5005076802B900F0 50050768020000F0 3      2           2            011501  yes
yes
39 50050768024000F0 50050768020000F0 4      2           2            0A82C0  yes
no
40 5005076802C900F0 50050768020000F0 4      2           2            0A82C1  yes
yes
```

6. Ensure that the hosts are using the NPIV ports for host I/O.

> **Note:**
>
> - You can verify that you are logged in to them by entering the **`lsfabric -host host_id_or_name`** command. If I/O activity is occurring, each host has at least one line in the command output that corresponds to a host port and shows active in the activity field.
>   - Hosts where no I/O was issued in the past 5 minutes do not show active for any login.
>   - Hosts that do not adhere to preferred paths might still be processing I/O to primary ports.
> - Depending on the host operating system, rescanning of the SAN may be required on some hosts to recognize additional paths now provided via primary host attach ports (virtual WWPNs).

7. After a minimum of 15 minutes has passed since entering **transitional** mode, change the system to enabled mode by entering the command as shown in Example 8-12.

*Example 8-12 Enabling the NPIV*

```
IBM_Storwize:ITSO:superuser>chiogrp -fctargetportmode enabled 0
```

Now NPIV has been enabled on the IBM Spectrum Virtualize system and hosts should also be using the virtualized WWPNs for I/O. At this point, the host zones can be amended appropriately to use primary host attach port WWPNs (virtual WWPNs) only.

## 8.4 Hosts operations using the GUI

This section describes the following host operations using the Spectrum Virtualize GUI.

- Creating hosts
- Advanced host administration
- Adding and deleting host ports
- Host mappings overview

## 8.4.1 Creating hosts

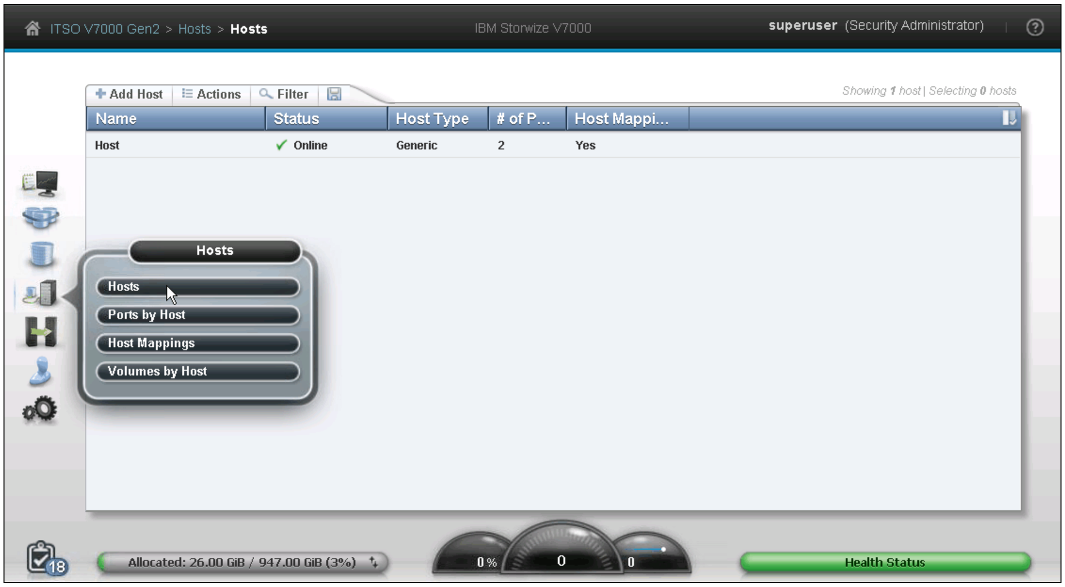
This section describes how to create Fibre Channel and iSCSI hosts using the IBM Spectrum Virtualize GUI. It is assumed that hosts are prepared for attachment, as described in the host attachment section of Spectrum Virtualize V7.8 knowledge center, and that the host WWPNs and their iSCSI initiator names are known.

The host attachment section of Spectrum Virtualize v7.8 can be found here:

`https://ibm.biz/BdsKgU`

To create a host, complete the described steps:

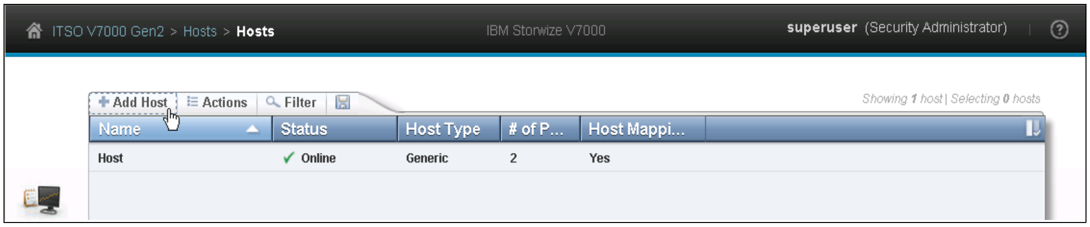
1. Open the host configuration window by clicking **Hosts** (Figure 8-3).


ITSO V7000 Gen2 > Hosts > Hosts
IBM Storwize V7000
superuser (Security Administrator)
Add Host
Actions
Filter
Showing 1 host | Selecting 0 hosts
Hosts
Hosts
Ports by Host
Host Mappings
Volumes by Host
18
Allocated: 26.00 GiB / 947.00 GiB (3%)
0 %
0
0
Health Status

| Name | Status | Host Type | # of P... | Host Mappi... |
|---|---|---|---|---|
| Host | ✓ Online | Generic | 2 | Yes |

*Figure 8-3 Open the host window*

2. To create a host, click **Add Host** to start the wizard (Figure 8-4).

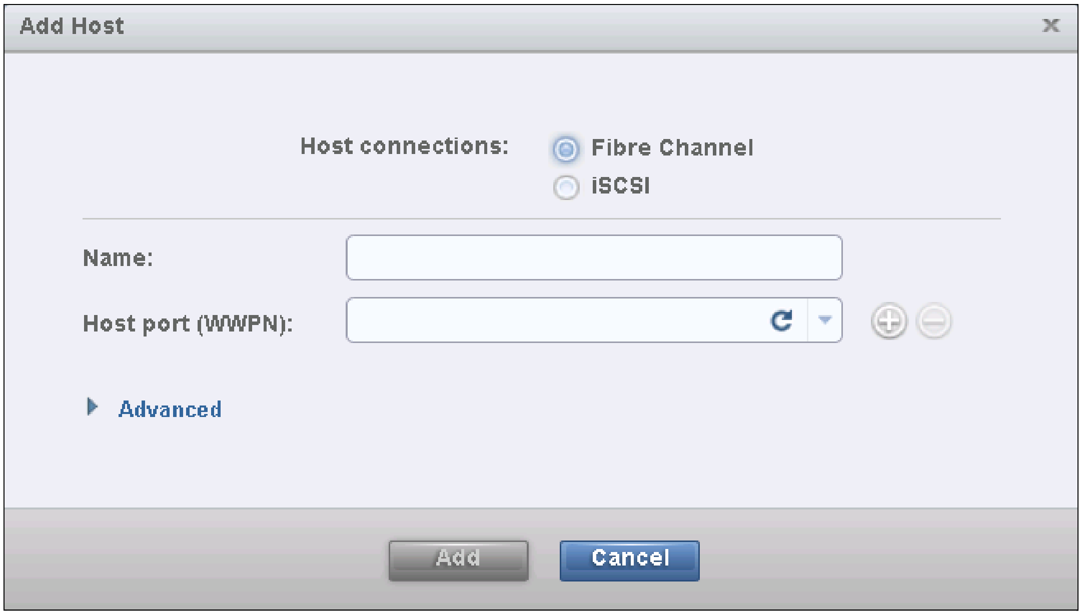

ITSO V7000 Gen2 > Hosts > Hosts
IBM Storwize V7000
superuser (Security Administrator)
Add Host
Actions
Filter
Showing 1 host | Selecting 0 hosts

| Name | Status | Host Type | # of P... | Host Mappi... |
|---|---|---|---|---|
| Host | ✓ Online | Generic | 2 | Yes |

*Figure 8-4 Add Host*

If you want to create a Fibre Channel host, continue with , “Creating Fibre Channel hosts” on page 295. To create iSCSI hosts, go to , “Creating iSCSI hosts” on page 299.

3. After pressing **Add Host**, the host selection menu opens, as shown in Figure 8-5.


Add Host

Host connections: ● Fibre Channel ○ iSCSI

Name:

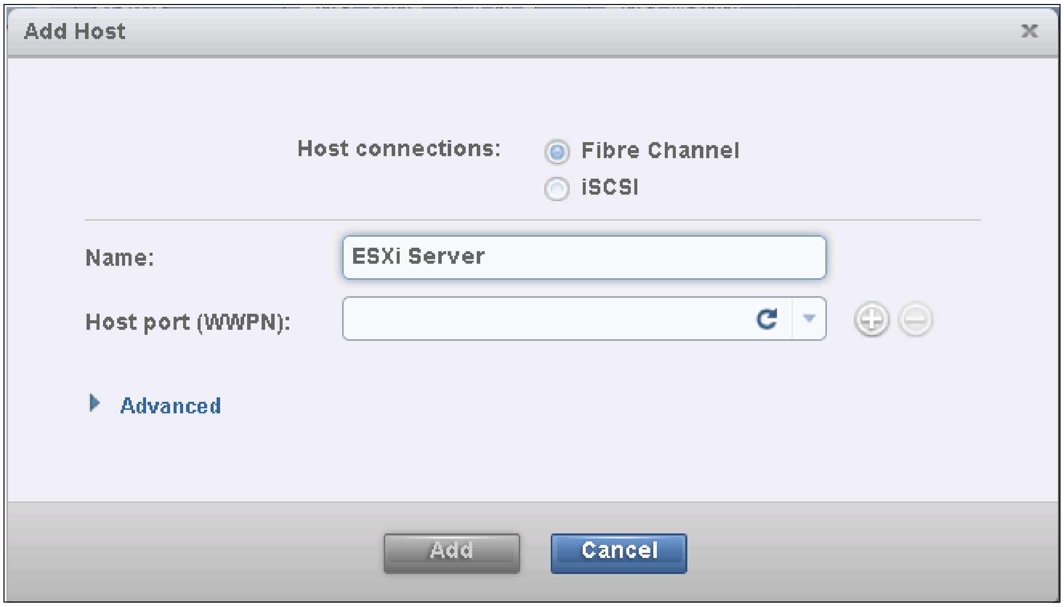
Host port (WWPN):

*Figure 8-5 Add host menu*

## Creating Fibre Channel hosts

To create Fibre Channel hosts, complete the following steps:

1. Click **Fibre Channel** (Figure 8-5 on page 295). The Fibre Channel configuration wizard opens (Figure 8-6).


Add Host

Host connections: ● Fibre Channel ○ iSCSI

Name: ESXi Server

Host port (WWPN):

*Figure 8-6 Create a Fibre Channel host*

2. Enter a host name and click the **Host Port** menu to get a list of all known WWPNs (Figure 8-7).

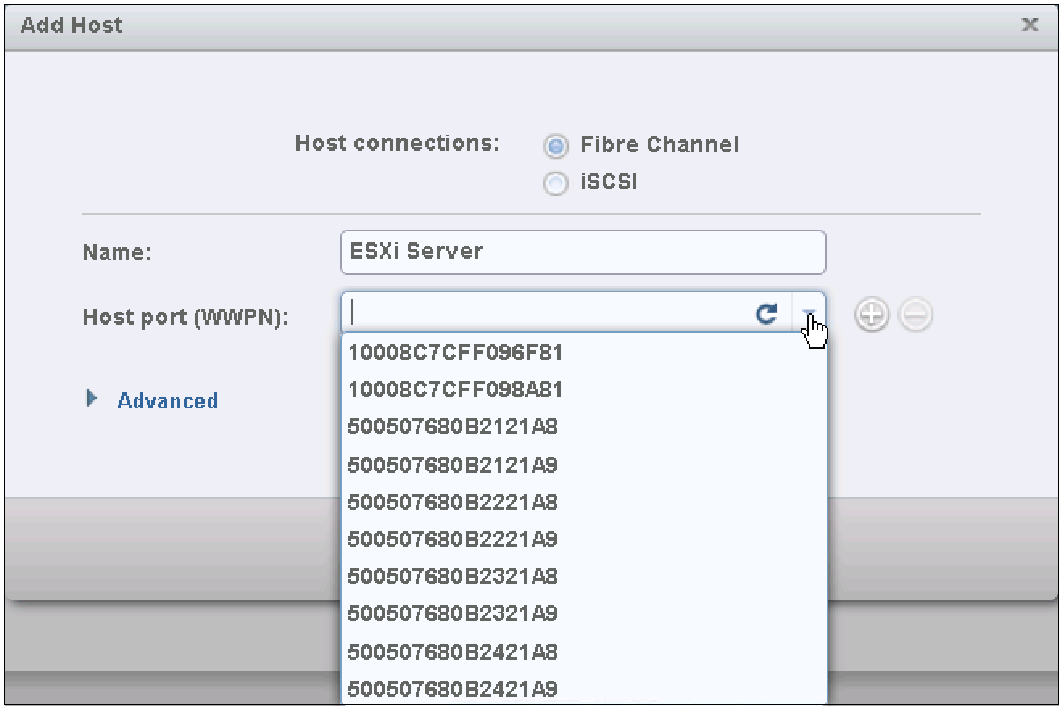


*Figure 8-7 Available WWPNs*

The IBM Spectrum Virtualize should have the host port WWPNs available if the host is prepared described in the 7.8 knowledge center for host attachment. If they do not appear in the list, scan for new disks as per he procedure based on the respective operating system and click **Rescan** in the configuration wizard. If they still do not appear, check the SAN zoning and repeat the scanning.

3. Select the WWPN for your host. (Figure 8-8).

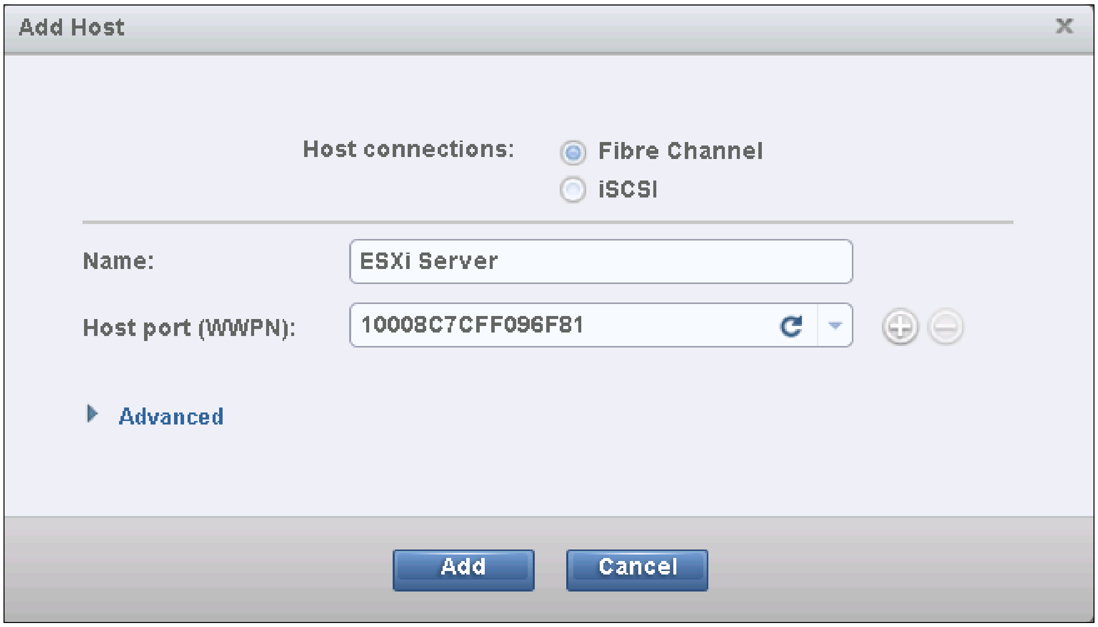
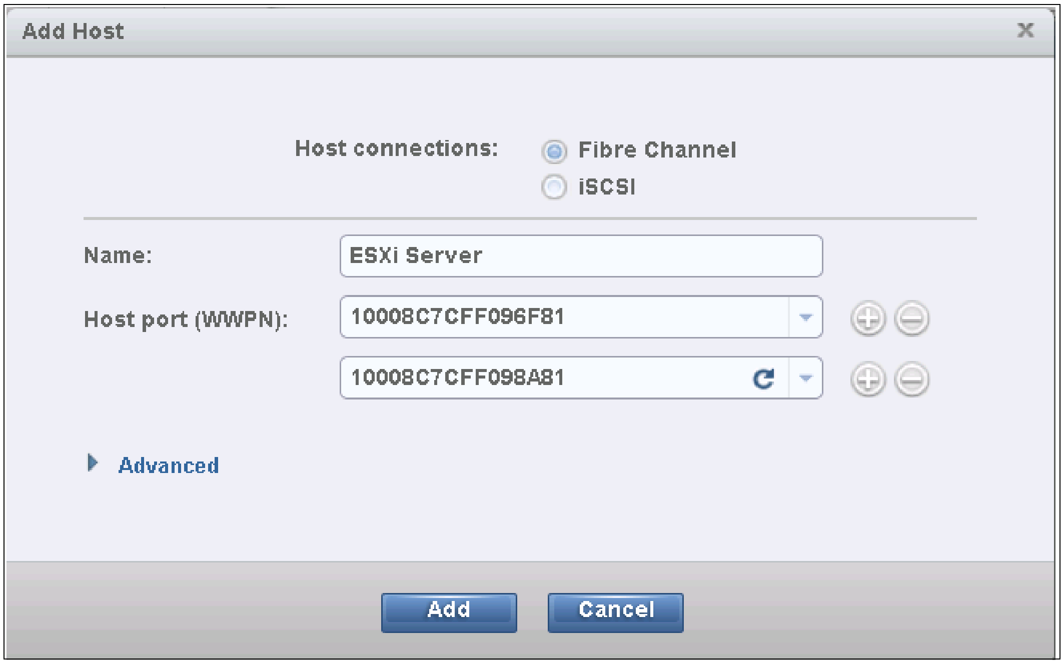


*Figure 8-8 Add a port to a list*

4. If you want to add additional ports to your Host, click the Plus sign (+).

5. Add all ports that belong to the host (Figure 8-9).



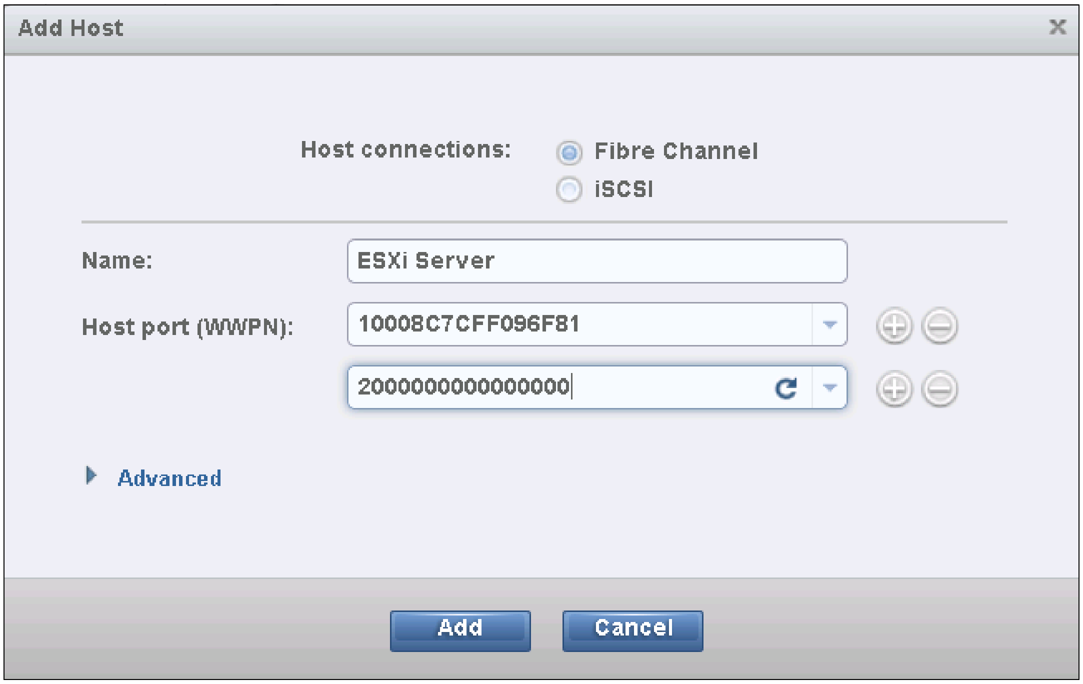
Name: ESXi Server

Host port (WWPN): 10008C7CFF096F81

10008C7CFF098A81

Figure 8-9 Add all WWPNs

**Creating offline hosts:** If you want to create hosts that are offline, or not connected at the moment, it is also possible to enter the WWPNs manually. Type them into the Host Ports Box and also add them to the list. See Figure 8-10.



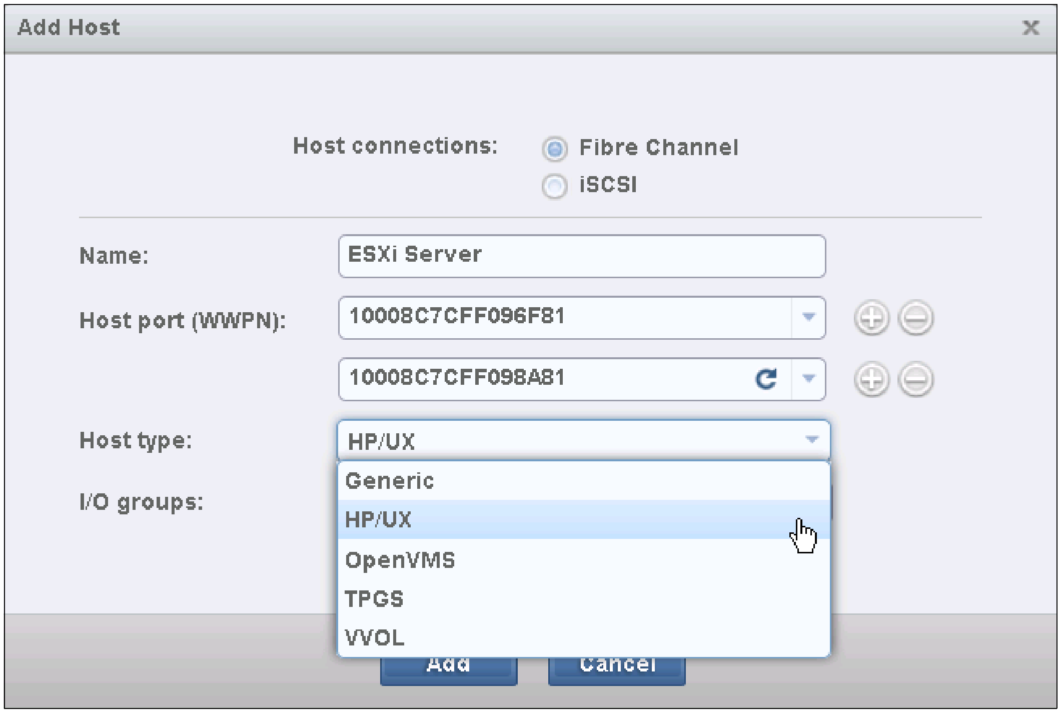
Name: ESXi Server

Host port (WWPN): 10008C7CFF096F81

2000000000000000

Figure 8-10 Manually added WWPN

6. If you are creating a Hewlett-Packard UNIX (HP-UX) or Target Port Group Support (TPGS) host, select the **Advanced** check box, and more options appear (Figure 8-11). Select your host type.



Add Host

Host connections: ● Fibre Channel ○ iSCSI

Name: ESXi Server

Host port (WWPN): 10008C7CFF096F81

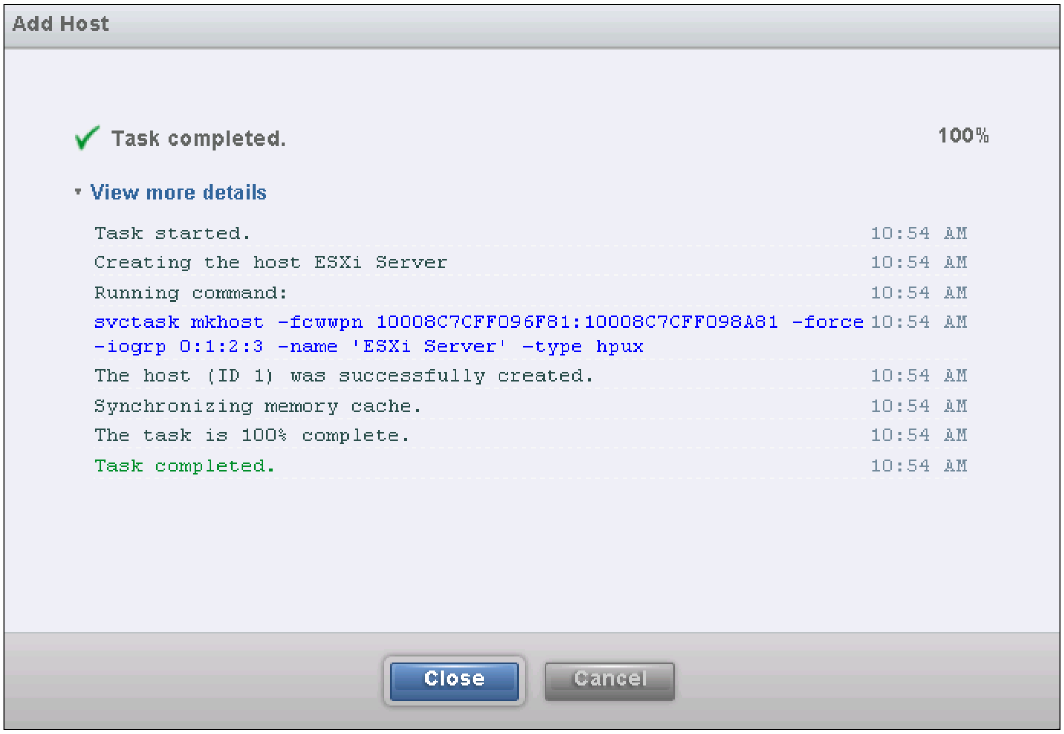
10008C7CFF098A81

Host type: HP/UX

I/O groups:

*Figure 8-11 Create Host: Advanced Settings*

7. Click **Add Host** and the wizard creates the host (Figure 8-12).



Add Host

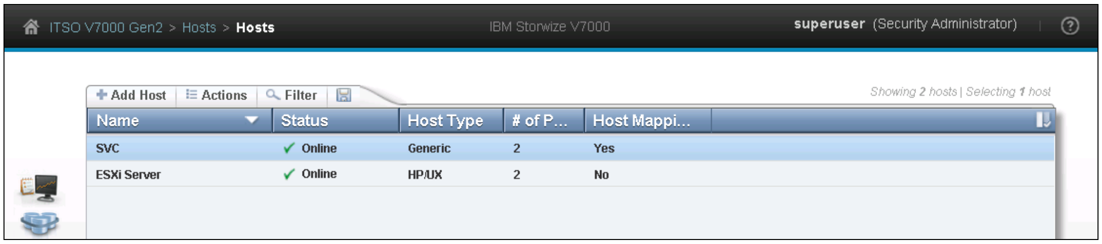
✓ Task completed. 100%

View more details

```
Task started.                                                   10:54 AM
Creating the host ESXi Server                                   10:54 AM
Running command:                                                10:54 AM
svctask mkhost -fcwwpn 10008C7CFF096F81:10008C7CFF098A81 -force 10:54 AM
-iogrp 0:1:2:3 -name 'ESXi Server' -type hpux
The host (ID 1) was successfully created.                       10:54 AM
Synchronizing memory cache.                                     10:54 AM
The task is 100% complete.                                      10:54 AM
Task completed.                                                 10:54 AM
```

*Figure 8-12 Add Host completes*

8. Click **Close** to return to the host window. Repeat these steps for all of your Fibre Channel hosts. Figure 8-13 shows the **All Hosts** window after creating a second host.


ITSO V7000 Gen2 > Hosts > Hosts
IBM Storwize V7000
superuser (Security Administrator)
Add Host
Actions
Filter
Showing 2 hosts | Selecting 1 host

| Name | Status | Host Type | # of P... | Host Mappi... |
|---|---|---|---|---|
| SVC | ✓ Online | Generic | 2 | Yes |
| ESXi Server | ✓ Online | HP/UX | 2 | No |

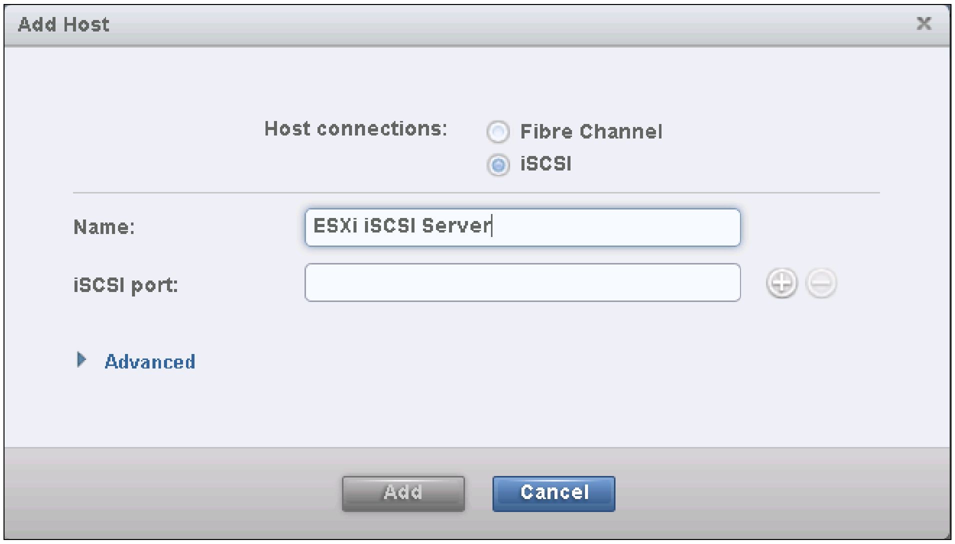
*Figure 8-13 All Hosts: After creating a second host*

After you complete the adding of Fibre Channel hosts, go to Chapter 7, “Volumes” on page 213 to create volumes and map them to the created hosts.

## Creating iSCSI hosts

To create iSCSI hosts, complete the following steps:

1. Click **iSCSI** and the iSCSI configuration wizard opens (Figure 8-14).



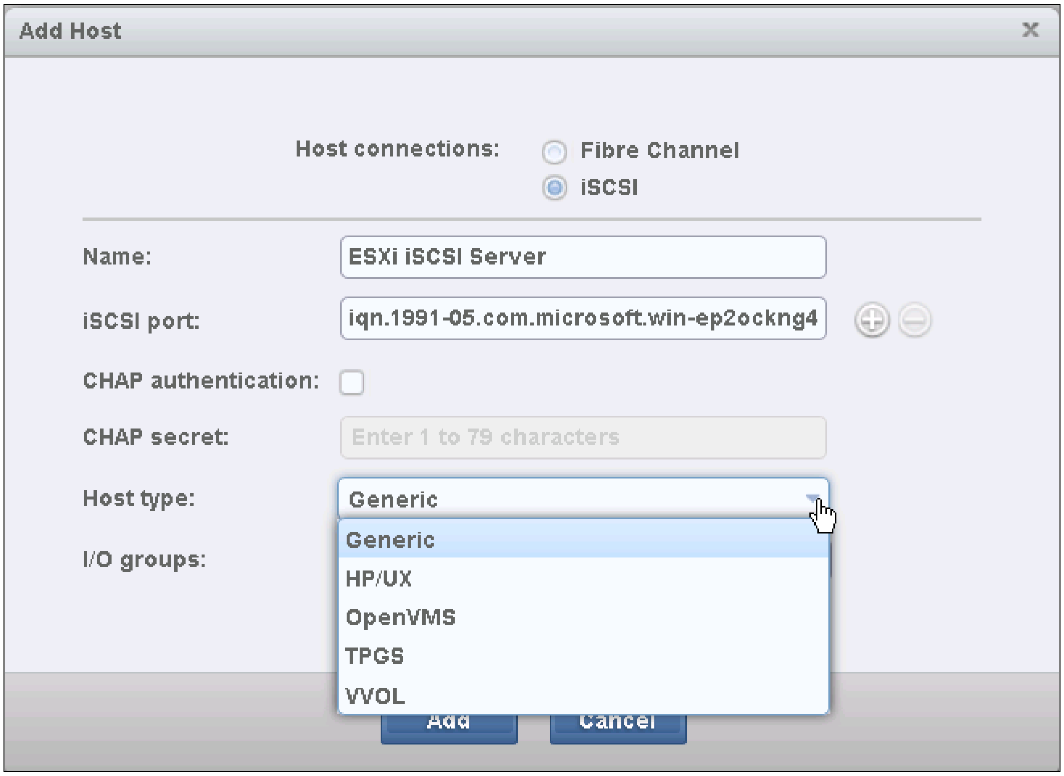
Host connections: ○ Fibre Channel ● iSCSI

Name: ESXi iSCSI Server

iSCSI port:

*Figure 8-14 Add an iSCSI host*

2. Enter a host name, type the iSCSI initiator name into the iSCSI Ports box, and insert the iSCSI port information into the **iSCSI port** field. If you want to add several initiator names to one host, repeat this step by clicking the plus sign (+).
3. If you are connecting an HP-UX or TPGS host, select the **Advanced** check box (Figure 8-15) and select the correct host type.



Add Host

Host connections: ○ Fibre Channel ● iSCSI

Name: ESXi iSCSI Server

iSCSI port: iqn.1991-05.com.microsoft.win-ep2ockng4

CHAP authentication: □

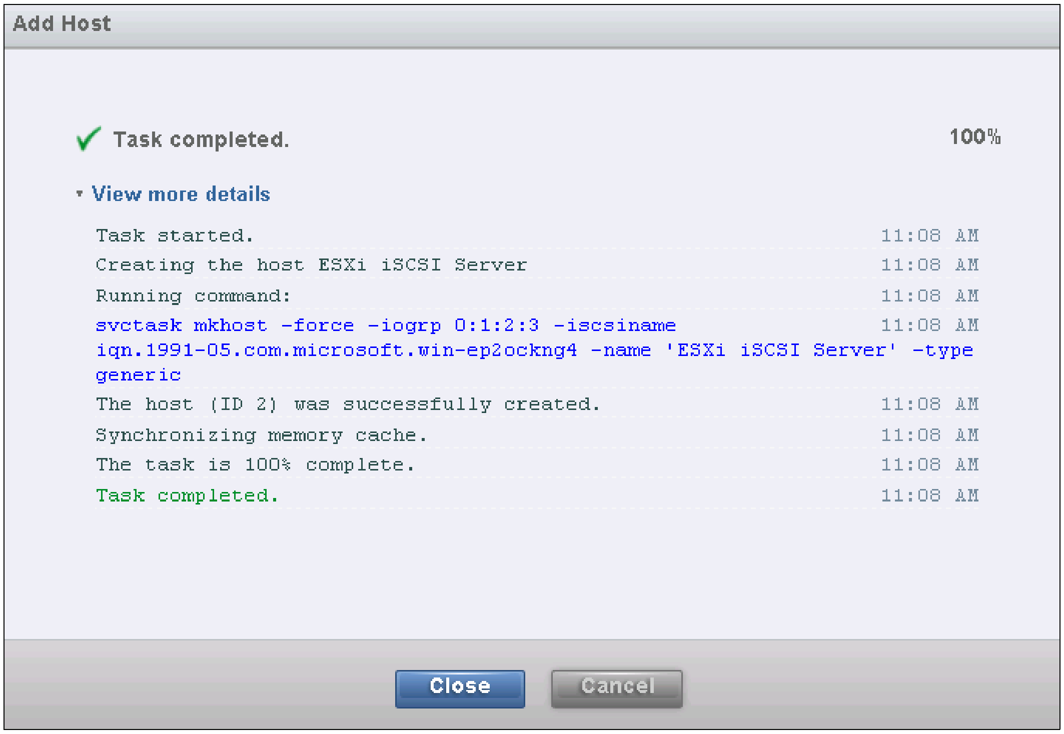
CHAP secret: Enter 1 to 79 characters

Host type: Generic

I/O groups:

*Figure 8-15 Create an iSCSI host: Advanced Settings*

4. Click **Add** and the wizard completes (Figure 8-16). Click **Close**.



Add Host

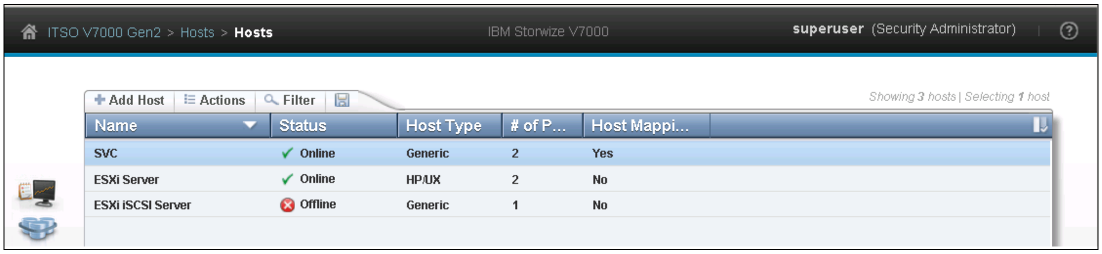
✓ Task completed. 100%

▾ View more details

```
Task started.                                                11:08 AM
Creating the host ESXi iSCSI Server                          11:08 AM
Running command:                                             11:08 AM
svctask mkhost -force -iogrp 0:1:2:3 -iscsiname              11:08 AM
iqn.1991-05.com.microsoft.win-ep2ockng4 -name 'ESXi iSCSI Server' -type
generic
The host (ID 2) was successfully created.                    11:08 AM
Synchronizing memory cache.                                  11:08 AM
The task is 100% complete.                                   11:08 AM
Task completed.                                              11:08 AM
```

*Figure 8-16 Add an iSCSI host: Complete*

5. Repeat these steps for every iSCSI host that you want to create. Figure 8-17 shows the All Hosts window after creating two Fibre Channel and one iSCSI host.


ITSO V7000 Gen2 > Hosts > Hosts
IBM Storwize V7000
superuser (Security Administrator)
Add Host
Actions
Filter
Showing 3 hosts | Selecting 1 host

| Name | Status | Host Type | # of P... | Host Mappi... |
|---|---|---|---|---|
| SVC | ✓ Online | Generic | 2 | Yes |
| ESXi Server | ✓ Online | HP/UX | 2 | No |
| ESXi iSCSI Server | ☒ Offline | Generic | 1 | No |

*Figure 8-17 All Hosts*

The iSCSI host is now configured. To provide connectivity, the iSCSI Ethernet ports must also be configured.

Complete the following steps to enable iSCSI connectivity:

1. Switch to the Configuration window and select **Network** (Figure 8-18).

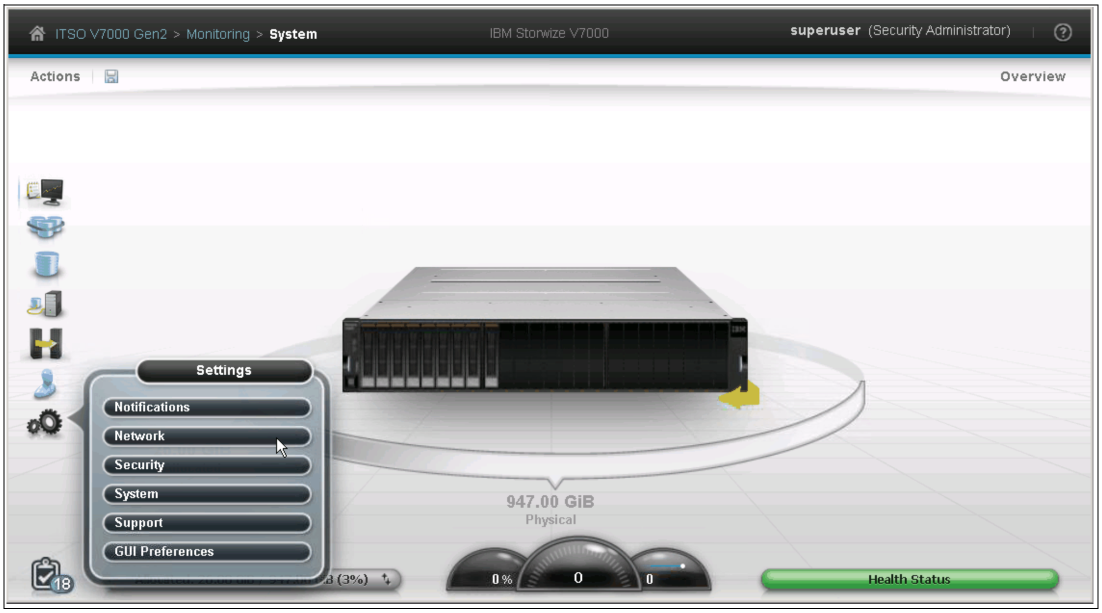
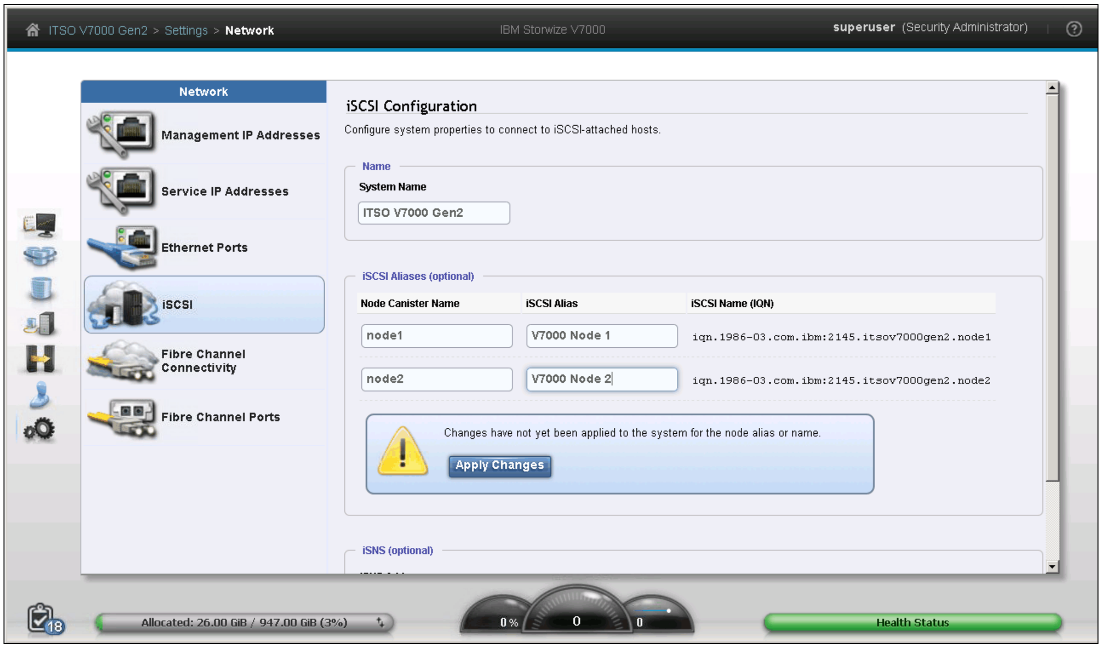


*Figure 8-18 Configuration: Network*

2. Select **iSCSI** and the iSCSI Configuration window opens (Figure 8-19).



iSCSI Configuration

Configure system properties to connect to iSCSI-attached hosts.

Name

System Name: ITSO V7000 Gen2

iSCSI Aliases (optional)

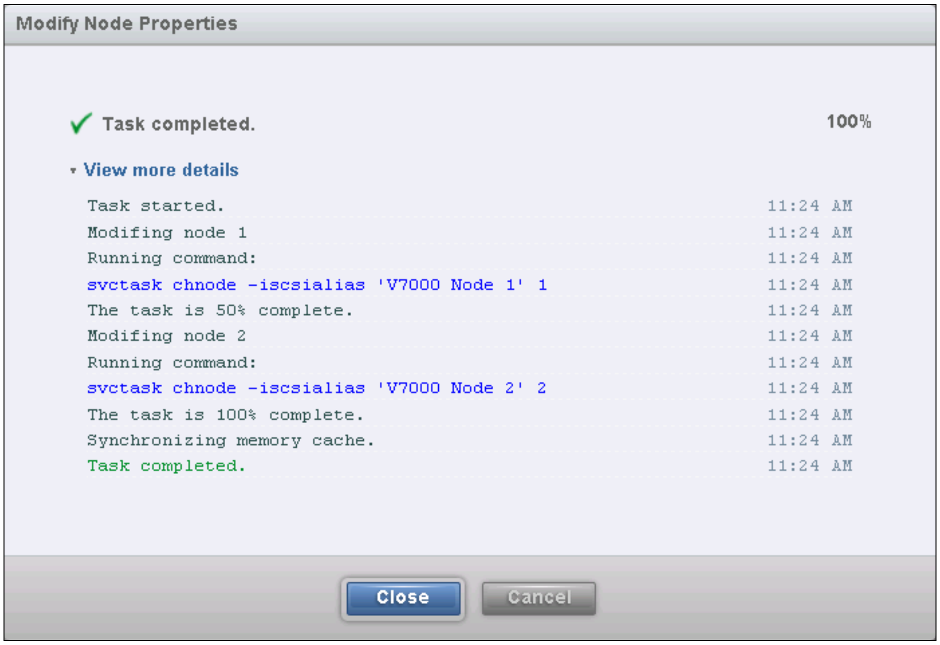
| Node Canister Name | iSCSI Alias | iSCSI Name (IQN) |
|---|---|---|
| node1 | V7000 Node 1 | iqn.1986-03.com.ibm:2145.itsov7000gen2.node1 |
| node2 | V7000 Node 2 | iqn.1986-03.com.ibm:2145.itsov7000gen2.node2 |

Changes have not yet been applied to the system for the node alias or name.

iSNS (optional)

*Figure 8-19 iSCSI Configuration window*

3. The system waits until you apply the changes that you made. Click **Apply Changes**. All changes are applied, as shown in Figure 8-20.



Task completed. 100%

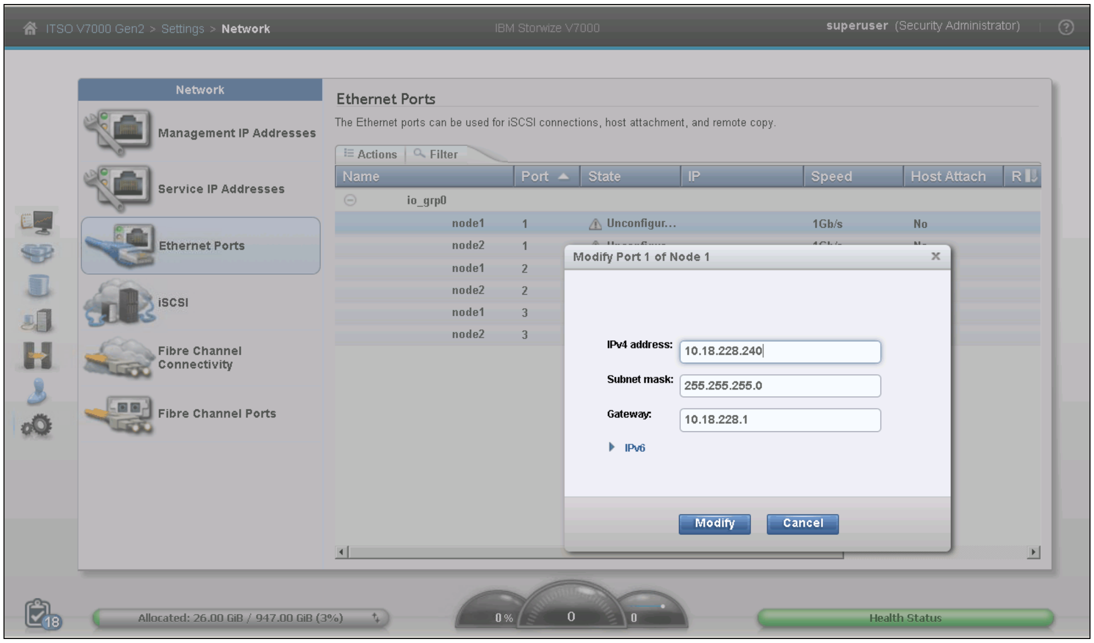
View more details

```
Task started.                                      11:24 AM
Modifing node 1                                    11:24 AM
Running command:                                   11:24 AM
svctask chnode -iscsialias 'V7000 Node 1' 1        11:24 AM
The task is 50% complete.                          11:24 AM
Modifing node 2                                    11:24 AM
Running command:                                   11:24 AM
svctask chnode -iscsialias 'V7000 Node 2' 2        11:24 AM
The task is 100% complete.                         11:24 AM
Synchronizing memory cache.                        11:24 AM
Task completed.                                    11:24 AM
```

*Figure 8-20 Applying all iSCSI changes*

In the configuration window, you have an overview of all of the iSCSI settings for the Storwize V7000. You can configure iSCSI Alias, Internet Storage Name Service (iSNS) Addresses, and Challenge Handshake Authentication Protocol (CHAP) on this window, and the iSCSI IP address, which we also edit in the basic setup.

4. Click **Ethernet Ports** to enter the iSCSI IP address (Figure 8-21). Repeat this step for each port that you want to use for iSCSI traffic.



Ethernet Ports

The Ethernet ports can be used for iSCSI connections, host attachment, and remote copy.

| Name | Port | State | IP | Speed | Host Attach |
|---|---|---|---|---|---|
| io_grp0 | | | | | |
| node1 | 1 | Unconfigur... | | 1Gb/s | No |
| node2 | 1 | | | | |
| node1 | 2 | | | | |
| node2 | 2 | | | | |
| node1 | 3 | | | | |
| node2 | 3 | | | | |

Modify Port 1 of Node 1

IPv4 address: 10.18.228.240

Subnet mask: 255.255.255.0

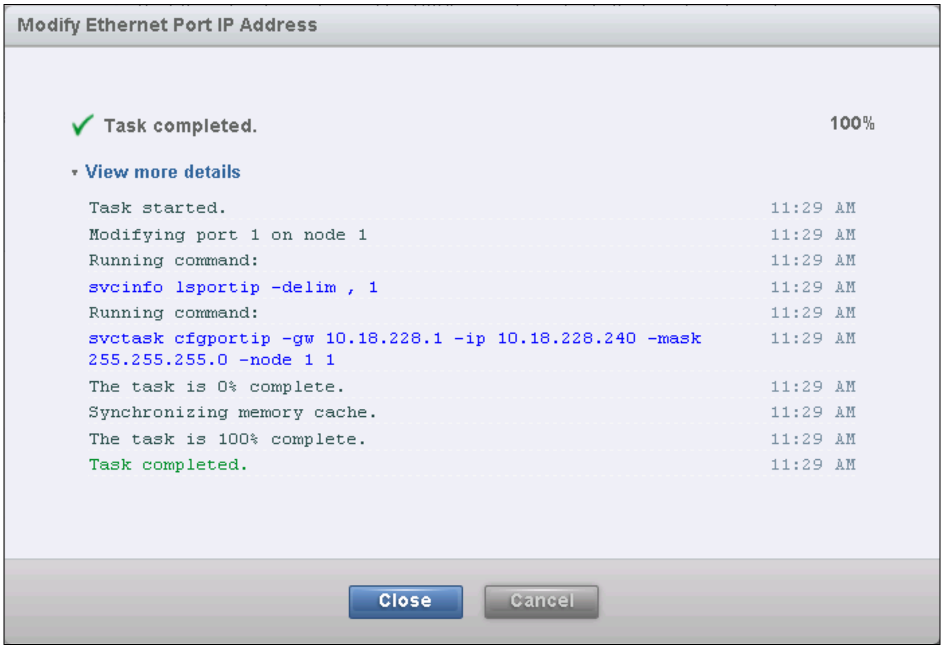
Gateway: 10.18.228.1

IPv6

Modify Cancel

*Figure 8-21 Enter an iSCSI IP address*

5. After you enter the IP address for each port, click **Modify** to enable the configuration.
6. After the changes are successfully applied, click **Close** (Figure 8-22).



Modify Ethernet Port IP Address

✓ Task completed. 100%

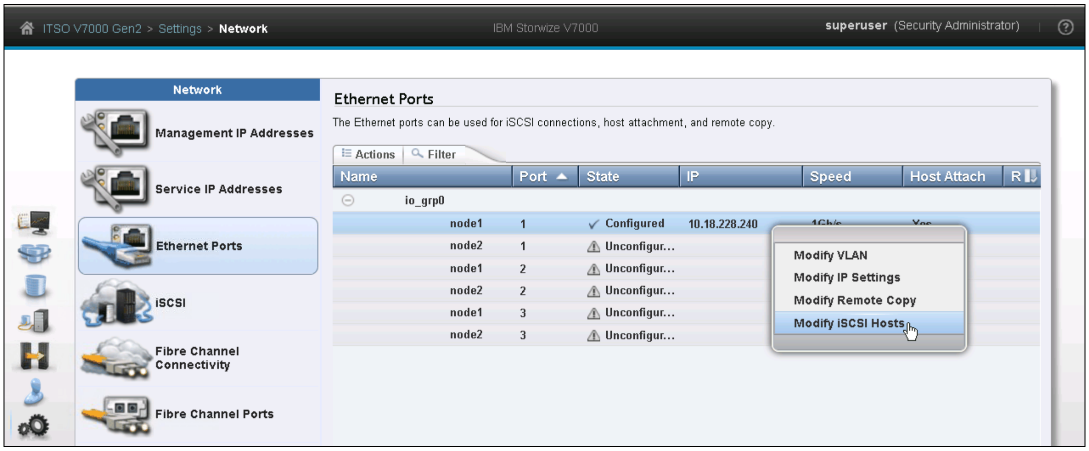
View more details

```
Task started.                                              11:29 AM
Modifying port 1 on node 1                                 11:29 AM
Running command:                                           11:29 AM
svcinfo lsportip -delim , 1                                11:29 AM
Running command:                                           11:29 AM
svctask cfgportip -gw 10.18.228.1 -ip 10.18.228.240 -mask  11:29 AM
255.255.255.0 -node 1 1
The task is 0% complete.                                   11:29 AM
Synchronizing memory cache.                                11:29 AM
The task is 100% complete.                                 11:29 AM
Task completed.                                            11:29 AM
```

Close Cancel

*Figure 8-22 iSCSI IP successfully modified*

7. Under **Actions,** you can check if all Hosts you want are iSCSI enabled. See Figure 8-23.



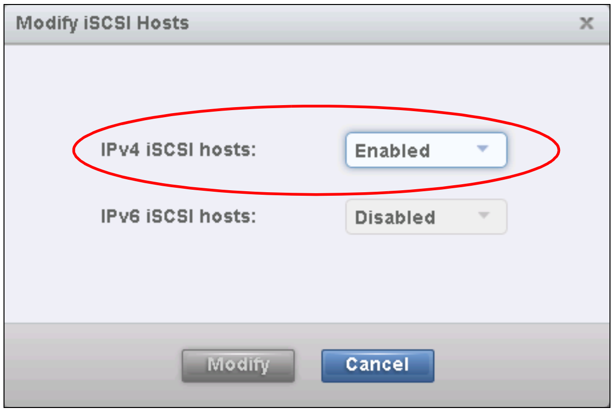
Ethernet Ports

The Ethernet ports can be used for iSCSI connections, host attachment, and remote copy.

| Name | Port | State | IP | Speed | Host Attach | R |
|---|---|---|---|---|---|---|
| io_grp0 | | | | | | |
| node1 | 1 | ✓ Configured | 10.18.228.240 | 1Gb/s | Yes | |
| node2 | 1 | ⚠ Unconfigur... | | | | |
| node1 | 2 | ⚠ Unconfigur... | | | | |
| node2 | 2 | ⚠ Unconfigur... | | | | |
| node1 | 3 | ⚠ Unconfigur... | | | | |
| node2 | 3 | ⚠ Unconfigur... | | | | |

*Figure 8-23 Action menu to modify iSCSI hosts*

8. By default all iSCSI hosts are enabled. (Figure 8-24).



IPv4 iSCSI hosts: Enabled

IPv6 iSCSI hosts: Disabled

*Figure 8-24 iSCSI host enabled*

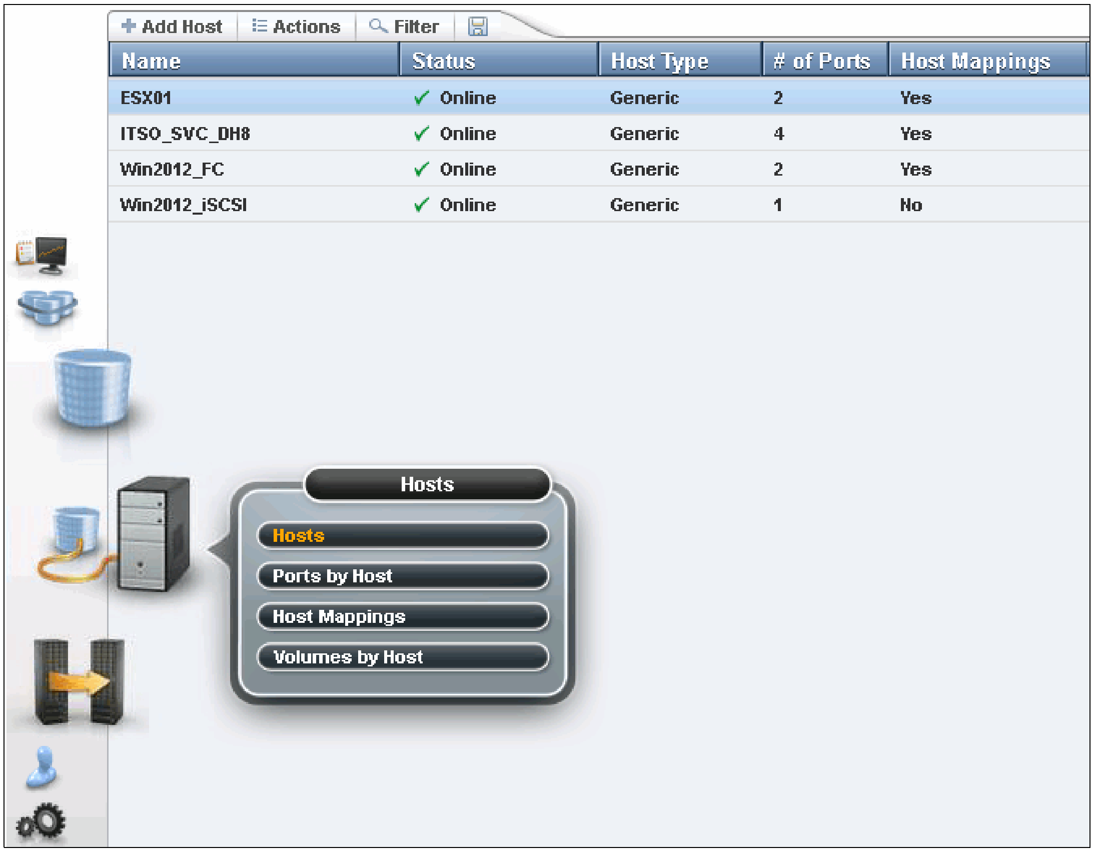
The IBM Storwize V7000 is now configured and ready for iSCSI use. Note the initiator names of your Storwize V7000 node canisters (Figure 8-19 on page 302) because you need them later. To create volumes and map them to a host, go to Chapter 7, “Volumes” on page 213.

## 8.4.2 Advanced host administration

This section covers host administration, including topics such as host modification, host mappings, and deleting hosts. Basic host creation using Fibre Channel (FC) and Internet Small Computer System Interface (iSCSI) connectivity has been described in 8.4.1, “Creating hosts” on page 294.

It is assumed that hosts have been created in your IBM Spectrum Virtualize GUI, and that some volumes are already mapped to them. This section describes three functions that are covered in the Hosts section of the IBM Spectrum Virtualize GUI (Figure 8-25):

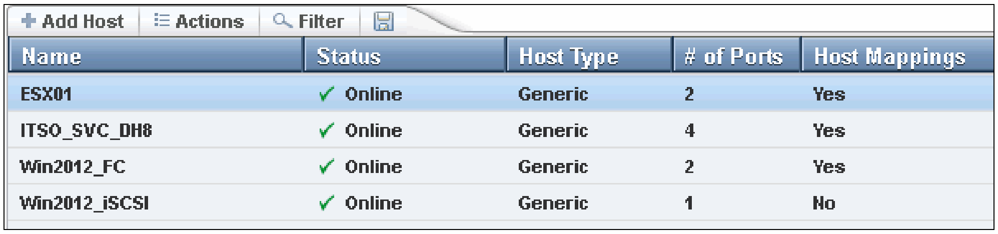
- Hosts (, “Modifying Mappings menu” on page 306)
- Ports by Host (8.4.3, “Adding and deleting host ports” on page 325)
- Host Mappings (8.4.4, “Host mappings overview” on page 334)



| Name | Status | Host Type | # of Ports | Host Mappings |
|---|---|---|---|---|
| ESX01 | ✓ Online | Generic | 2 | Yes |
| ITSO_SVC_DH8 | ✓ Online | Generic | 4 | Yes |
| Win2012_FC | ✓ Online | Generic | 2 | Yes |
| Win2012_iSCSI | ✓ Online | Generic | 1 | No |

*Figure 8-25 IBM Spectrum Virtualize Hosts menu*

If you hover the mouse cursor over the **Hosts** selection in the IBM Spectrum Virtualize GUI dynamic menu, the Hosts pane opens (Figure 8-26).

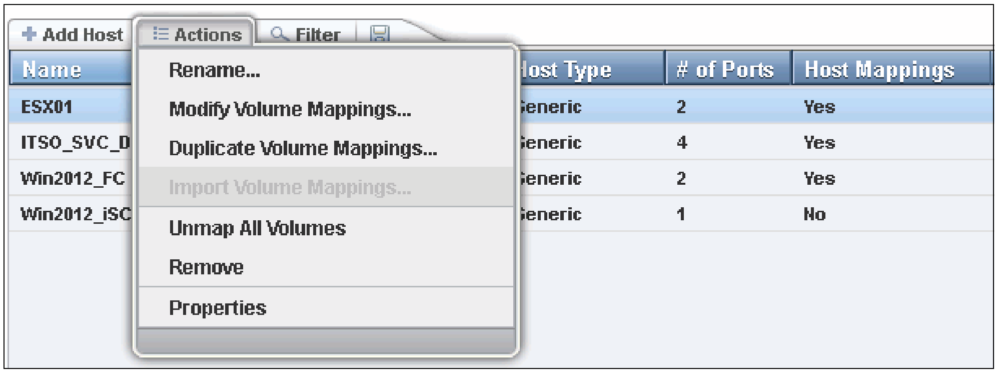


| Name | Status | Host Type | # of Ports | Host Mappings |
|---|---|---|---|---|
| ESX01 | ✓ Online | Generic | 2 | Yes |
| ITSO_SVC_DH8 | ✓ Online | Generic | 4 | Yes |
| Win2012_FC | ✓ Online | Generic | 2 | Yes |
| Win2012_iSCSI | ✓ Online | Generic | 1 | No |

*Figure 8-26 Hosts pane*

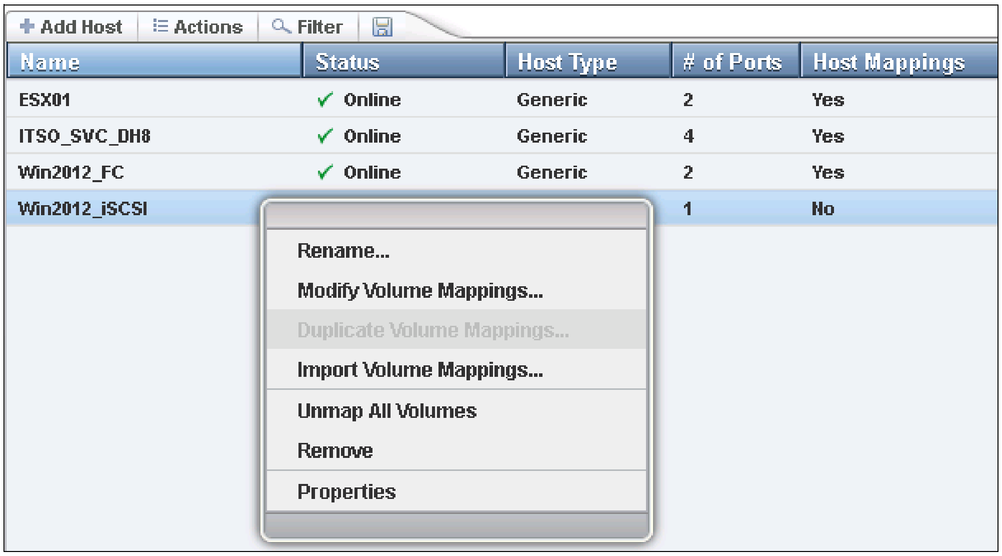
As our example shows, four hosts have been created, and volumes are already mapped to some of them. We use these hosts to show the modification possibilities.

9. If you highlight a host, you can either click **Actions** (Figure 8-27) or right-click the host (Figure 8-28) to see all the available tasks.


Add Host
Actions
Filter
Rename...
Modify Volume Mappings...
Duplicate Volume Mappings...
Import Volume Mappings...
Unmap All Volumes
Remove
Properties

| Name | ost Type | # of Ports | Host Mappings |
|---|---|---|---|
| ESX01 | eneric | 2 | Yes |
| ITSO_SVC_D | eneric | 4 | Yes |
| Win2012_FC | eneric | 2 | Yes |
| Win2012_iSC | eneric | 1 | No |

*Figure 8-27 Host Actions*


Add Host
Actions
Filter
Rename...
Modify Volume Mappings...
Duplicate Volume Mappings...
Import Volume Mappings...
Unmap All Volumes
Remove
Properties

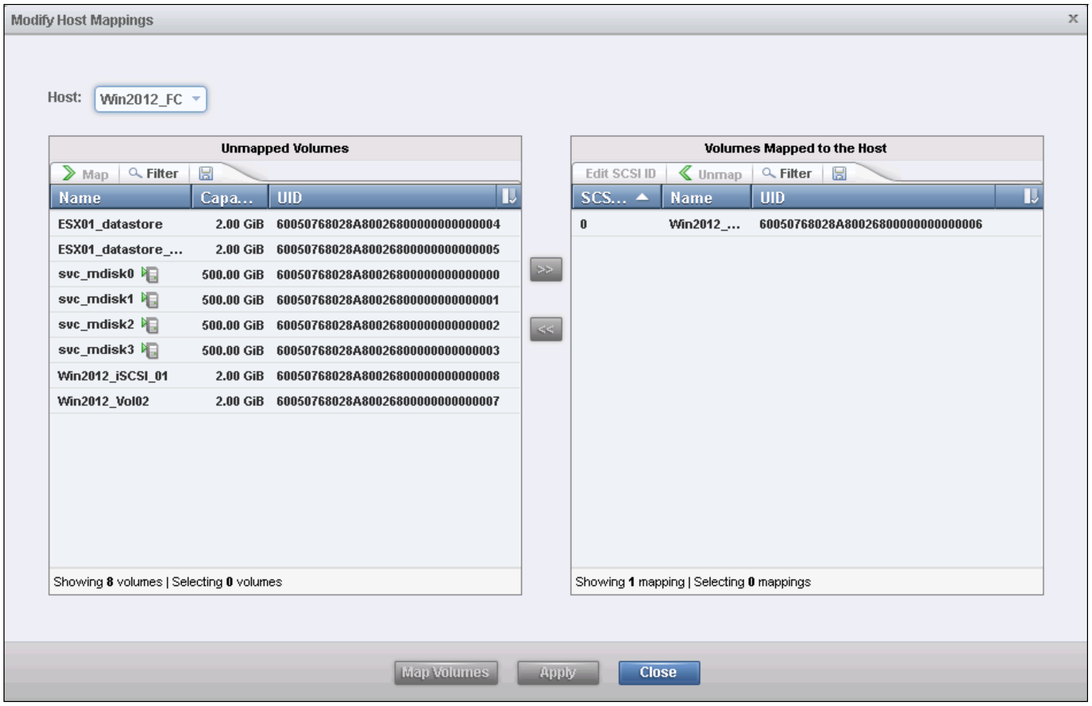
| Name | Status | Host Type | # of Ports | Host Mappings |
|---|---|---|---|---|
| ESX01 | ✓ Online | Generic | 2 | Yes |
| ITSO_SVC_DH8 | ✓ Online | Generic | 4 | Yes |
| Win2012_FC | ✓ Online | Generic | 2 | Yes |
| Win2012_iSCSI | | | 1 | No |

*Figure 8-28 Host menu*

## Modifying Mappings menu

To modify mappings, complete the following steps:

1. On the Hosts pane, select a host and click **Actions** → **Modify Mappings** (Figure 8-27 on page 306). The window shown in Figure 8-29 opens. At the upper left, you see that the highlighted host is selected. The two list boxes show all available unmapped and mapped volumes. The left list includes the volumes that are ready for mapping to this host. The right list includes the volumes already mapped. In our example, one volume with SCSI ID `0` is mapped to the host `Win2012_FC`, and eight more volumes are available (unmapped).



Unmapped Volumes

| Name | Capa... | UID |
|---|---|---|
| ESX01_datastore | 2.00 GiB | 60050768028A80026800000000000004 |
| ESX01_datastore_... | 2.00 GiB | 60050768028A80026800000000000005 |
| svc_mdisk0 | 500.00 GiB | 60050768028A80026800000000000000 |
| svc_mdisk1 | 500.00 GiB | 60050768028A80026800000000000001 |
| svc_mdisk2 | 500.00 GiB | 60050768028A80026800000000000002 |
| svc_mdisk3 | 500.00 GiB | 60050768028A80026800000000000003 |
| Win2012_iSCSI_01 | 2.00 GiB | 60050768028A80026800000000000008 |
| Win2012_Vol02 | 2.00 GiB | 60050768028A80026800000000000007 |

Showing 8 volumes | Selecting 0 volumes

Volumes Mapped to the Host

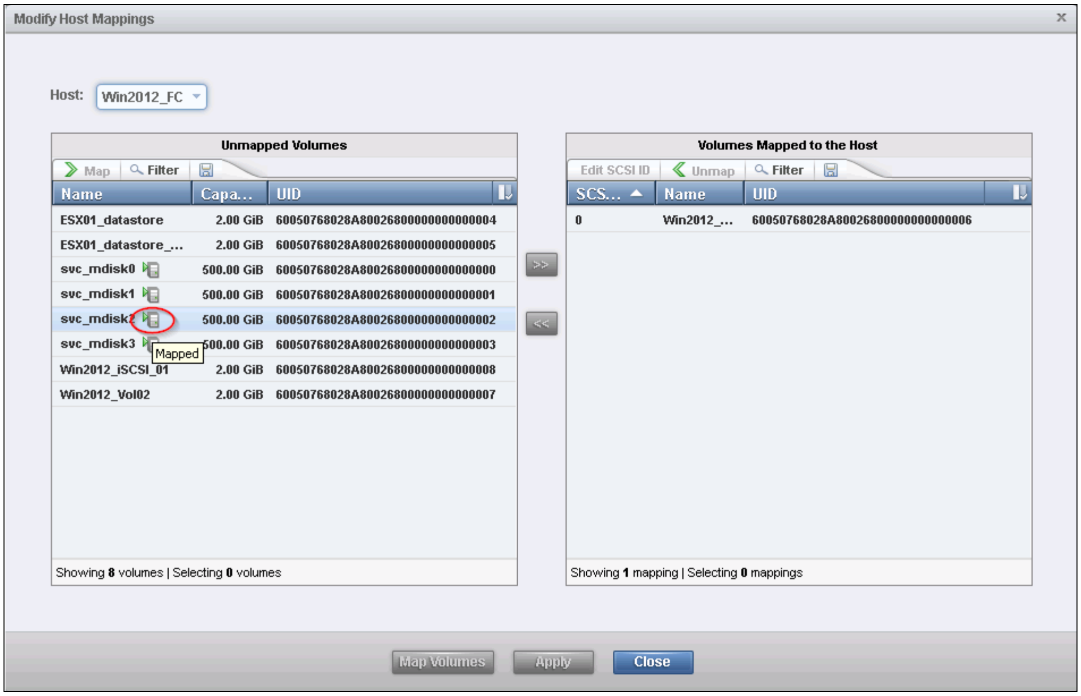
| SCS... | Name | UID |
|---|---|---|
| 0 | Win2012_... | 60050768028A80026800000000000006 |

Showing 1 mapping | Selecting 0 mappings

Figure 8-29 Modify Host Mappings window

You can easily identify if the volume that you want to map is already mapped to another host. The already mapped volumes on the left pane have a green arrow and host icon by them.

2. When you hover a mouse cursor over this icon, the little Box icon with the **Mapped** sign shows, as seen in Figure 8-30.



Unmapped Volumes

| Name | Capa... | UID |
|---|---|---|
| ESX01_datastore | 2.00 GiB | 60050768028A80026800000000000004 |
| ESX01_datastore_... | 2.00 GiB | 60050768028A80026800000000000005 |
| svc_mdisk0 | 500.00 GiB | 60050768028A80026800000000000000 |
| svc_mdisk1 | 500.00 GiB | 60050768028A80026800000000000001 |
| svc_mdisk2 | 500.00 GiB | 60050768028A80026800000000000002 |
| svc_mdisk3 | 500.00 GiB | 60050768028A80026800000000000003 |
| Win2012_iSCSI_01 | 2.00 GiB | 60050768028A80026800000000000008 |
| Win2012_Vol02 | 2.00 GiB | 60050768028A80026800000000000007 |

Showing 8 volumes | Selecting 0 volumes

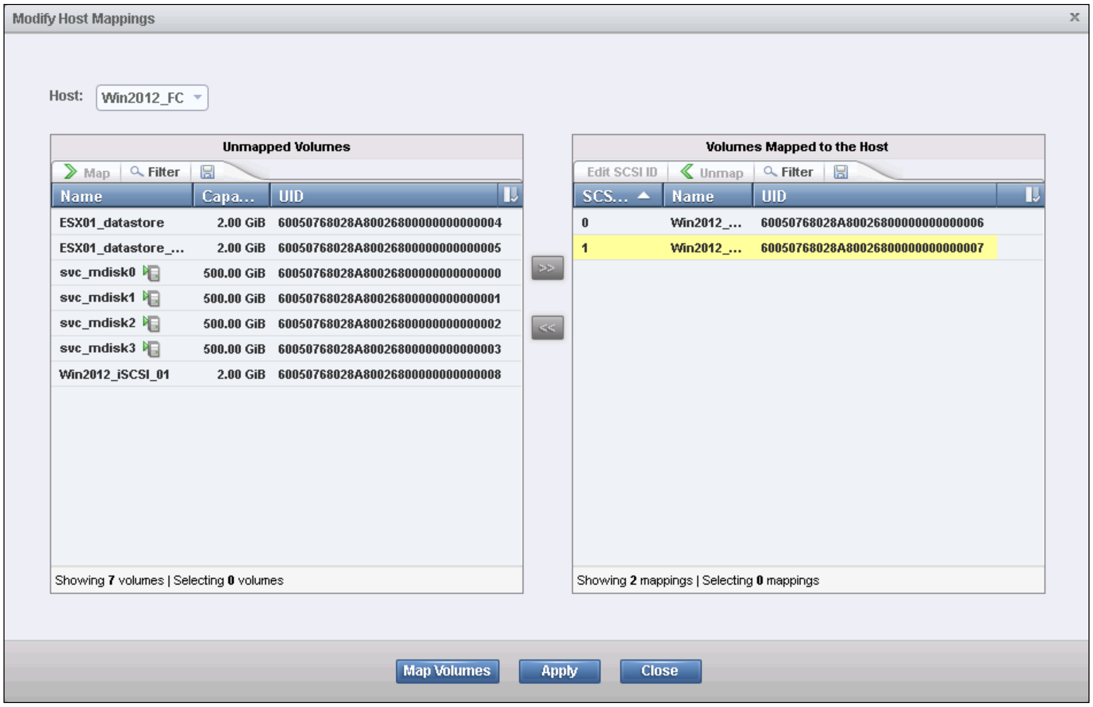
Volumes Mapped to the Host

| SCS... | Name | UID |
|---|---|---|
| 0 | Win2012_... | 60050768028A80026800000000000006 |

Showing 1 mapping | Selecting 0 mappings

Figure 8-30 Volumes already mapped to other hosts

3. To map a volume, select it in the left pane, and then click the right arrow (>>) to move the volume to the list on the right, where mapped volumes are placed (Figure 8-31).



Host: Win2012_FC

Unmapped Volumes

| Name | Capa... | UID |
|---|---|---|
| ESX01_datastore | 2.00 GiB | 60050768028A80026800000000000004 |
| ESX01_datastore_... | 2.00 GiB | 60050768028A80026800000000000005 |
| svc_mdisk0 | 500.00 GiB | 60050768028A80026800000000000000 |
| svc_mdisk1 | 500.00 GiB | 60050768028A80026800000000000001 |
| svc_mdisk2 | 500.00 GiB | 60050768028A80026800000000000002 |
| svc_mdisk3 | 500.00 GiB | 60050768028A80026800000000000003 |
| Win2012_iSCSI_01 | 2.00 GiB | 60050768028A80026800000000000008 |

Showing **7** volumes | Selecting **0** volumes

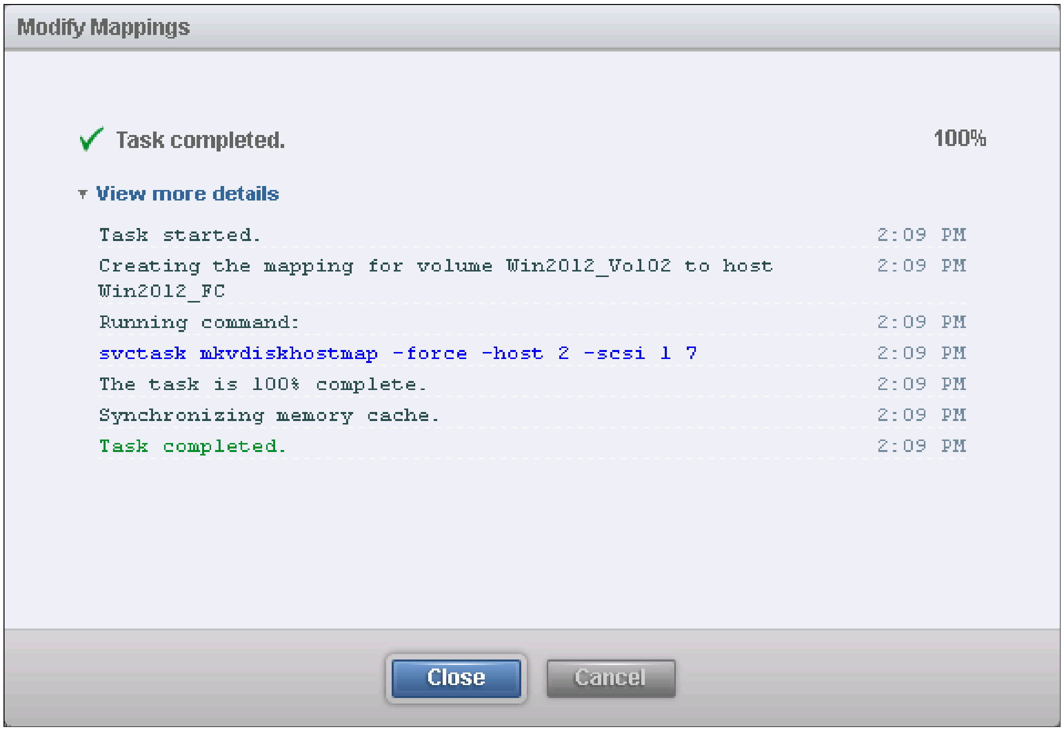
Volumes Mapped to the Host

| SCS... | Name | UID |
|---|---|---|
| 0 | Win2012_... | 60050768028A80026800000000000006 |
| 1 | Win2012_... | 60050768028A80026800000000000007 |

Showing **2** mappings | Selecting **0** mappings

*Figure 8-31 Modify Host Mappings: Map volume*

4. The changes are marked in yellow and now the **Map Volumes** and **Apply** buttons are enabled. If you click **Map Volumes**, the changes are applied (Figure 8-32) and the window closes. If you click **Apply**, the changes are submitted to the system (Figure 8-32), but the window remains open for further changes (Figure 8-33).



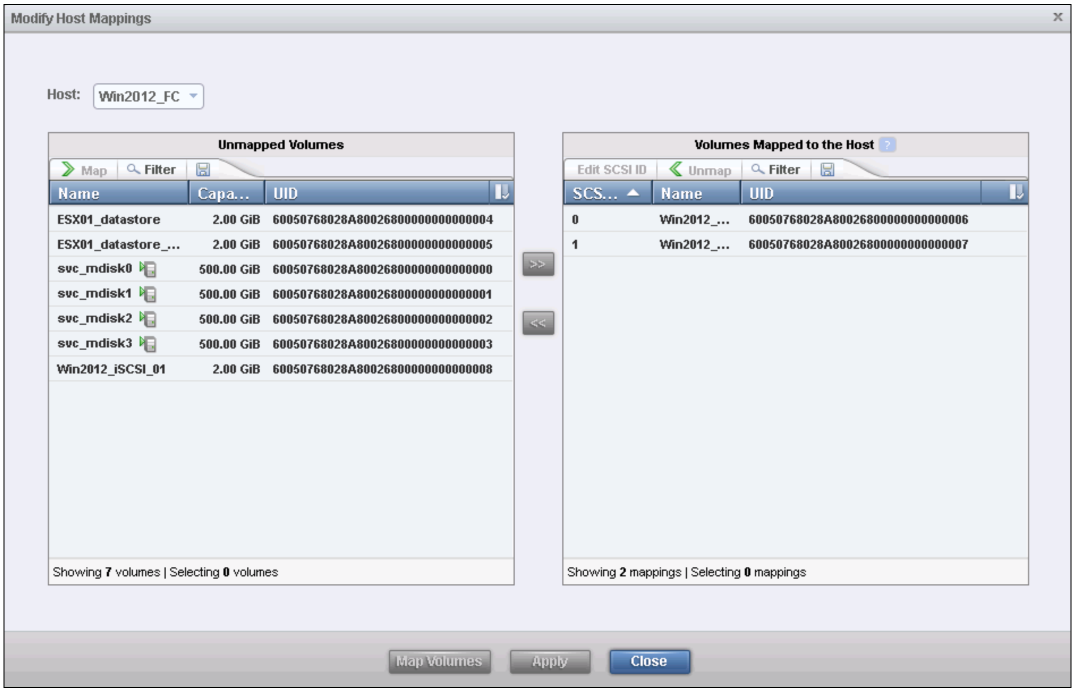
Modify Mappings

✓ Task completed. 100%

▾ View more details

```
Task started.                                                 2:09 PM
Creating the mapping for volume Win2012_Vol02 to host         2:09 PM
Win2012_FC
Running command:                                              2:09 PM
svctask mkvdiskhostmap -force -host 2 -scsi 1 7               2:09 PM
The task is 100% complete.                                    2:09 PM
Synchronizing memory cache.                                   2:09 PM
Task completed.                                               2:09 PM
```

*Figure 8-32 Modify Mappings task completion window*



Modify Host Mappings

Host: Win2012_FC

Unmapped Volumes

| Name | Capa... | UID |
|---|---|---|
| ESX01_datastore | 2.00 GiB | 60050768028A8002680000000000000004 |
| ESX01_datastore_... | 2.00 GiB | 60050768028A8002680000000000000005 |
| svc_mdisk0 | 500.00 GiB | 60050768028A8002680000000000000000 |
| svc_mdisk1 | 500.00 GiB | 60050768028A8002680000000000000001 |
| svc_mdisk2 | 500.00 GiB | 60050768028A8002680000000000000002 |
| svc_mdisk3 | 500.00 GiB | 60050768028A8002680000000000000003 |
| Win2012_iSCSI_01 | 2.00 GiB | 60050768028A8002680000000000000008 |

Showing 7 volumes | Selecting 0 volumes

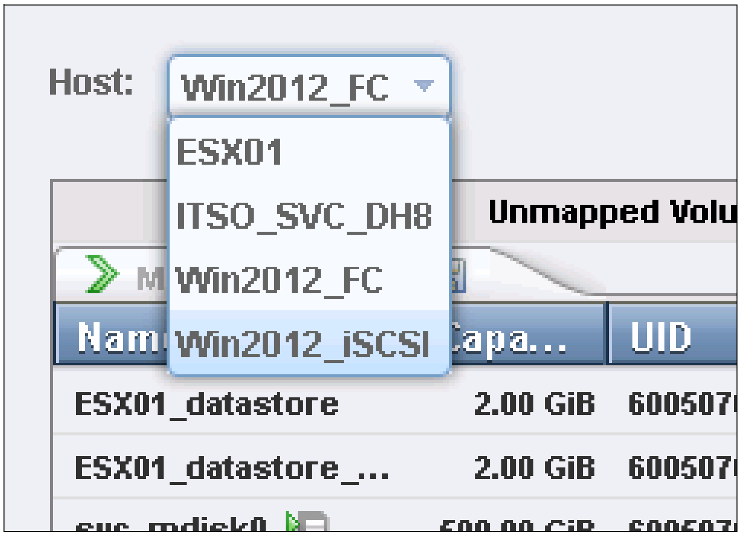
Volumes Mapped to the Host

| SCS... | Name | UID |
|---|---|---|
| 0 | Win2012_... | 60050768028A8002680000000000000006 |
| 1 | Win2012_... | 60050768028A8002680000000000000007 |

Showing 2 mappings | Selecting 0 mappings

*Figure 8-33 Modify Mappings: Applied changes*

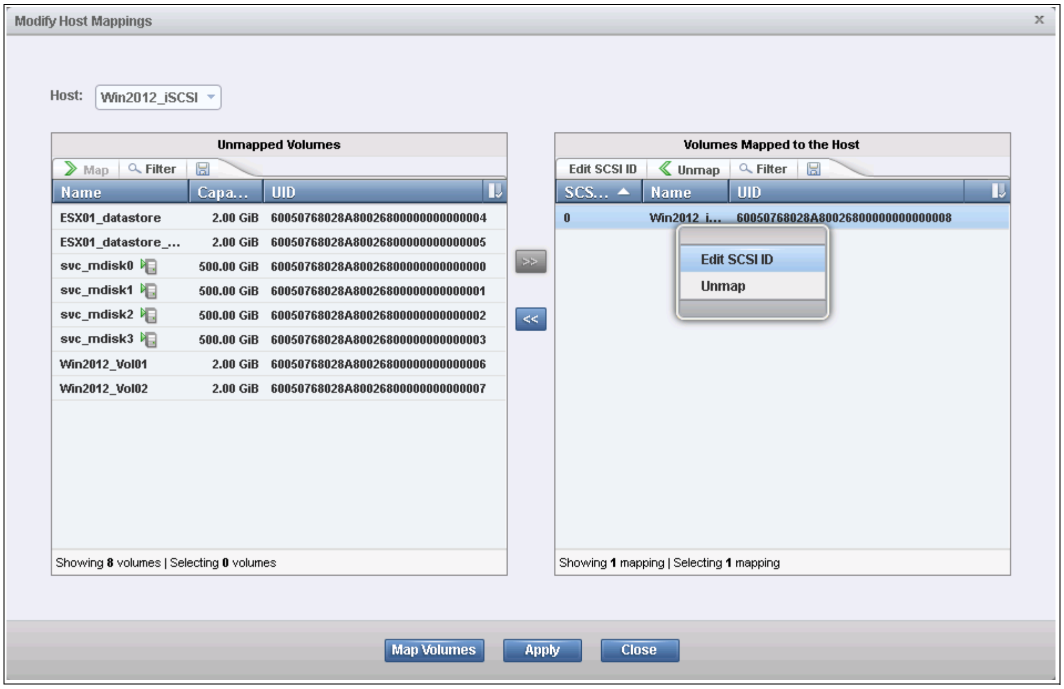
You can now select another host in the Host menu (Figure 8-34) to modify the host settings for it, or continue working with the one that is selected (Figure 8-33 on page 309).



| Nam | Capa... | UID |
|---|---|---|
| ESX01_datastore | 2.00 GiB | 600507 |
| ESX01_datastore_... | 2.00 GiB | 600507 |

Figure 8-34 Modify another hosts mappings

5. Select the volume to be modified, again, and click the right arrow to move the volume to the list in the right pane. The changes are shown in yellow there. If you right-click the highlighted volume, you are able to change the SCSI ID, which is used for the host mapping (Figure 8-35). Select **Edit SCSI ID**.



Unmapped Volumes

| Name | Capa... | UID |
|---|---|---|
| ESX01_datastore | 2.00 GiB | 60050768028A8002680000000000000004 |
| ESX01_datastore_... | 2.00 GiB | 60050768028A8002680000000000000005 |
| svc_mdisk0 | 500.00 GiB | 60050768028A8002680000000000000000 |
| svc_mdisk1 | 500.00 GiB | 60050768028A8002680000000000000001 |
| svc_mdisk2 | 500.00 GiB | 60050768028A8002680000000000000002 |
| svc_mdisk3 | 500.00 GiB | 60050768028A8002680000000000000003 |
| Win2012_Vol01 | 2.00 GiB | 60050768028A8002680000000000000006 |
| Win2012_Vol02 | 2.00 GiB | 60050768028A8002680000000000000007 |

Showing 8 volumes | Selecting 0 volumes

Volumes Mapped to the Host

| SCS... | Name | UID |
|---|---|---|
| 0 | Win2012_i... | 60050768028A8002680000000000000008 |

Showing 1 mapping | Selecting 1 mapping

Figure 8-35 Edit the SCSI ID

6. Enter a SCSI ID and click **OK** to change the ID (Figure 8-36).

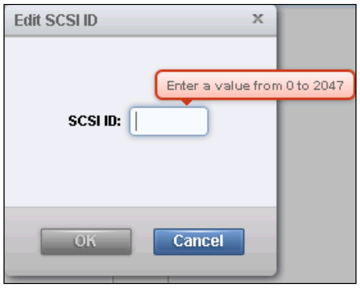
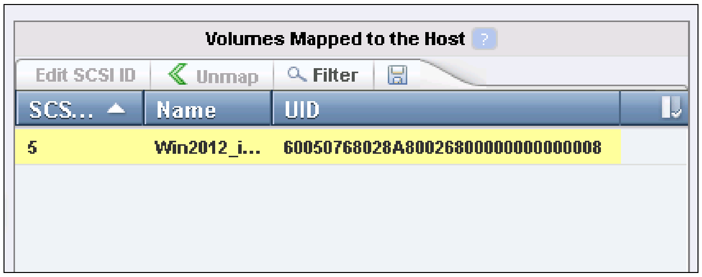


*Figure 8-36 Enter a new SCSI ID*

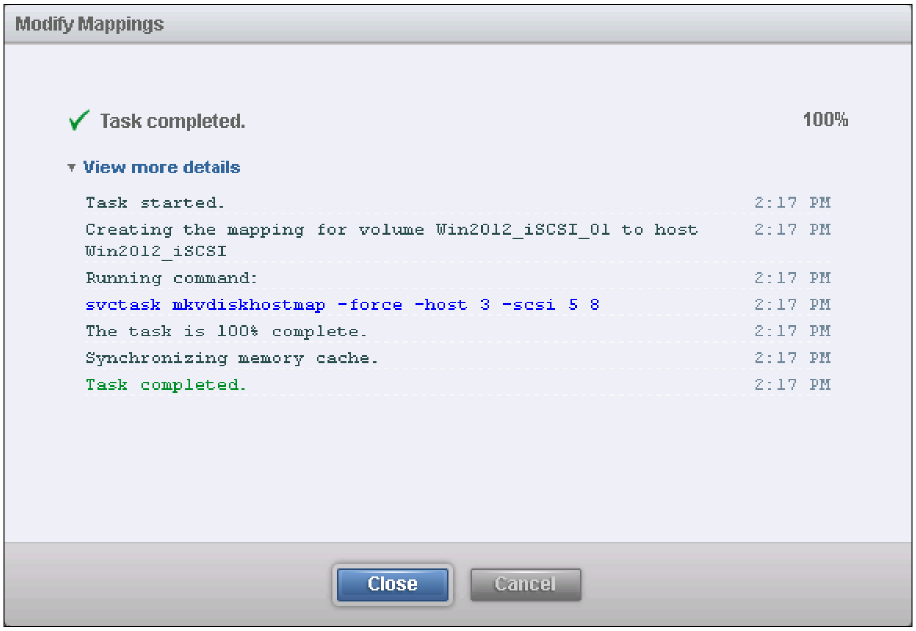
The changes are shown in the Modify Host Mappings window (Figure 8-37).



| SCS... | Name | UID |
|---|---|---|
| 5 | Win2012_i... | 60050768028A80026800000000000008 |

*Figure 8-37 Modify Host Mappings*

7. Click **Apply** to submit the changes. The resulting output is shown in Figure 8-38.

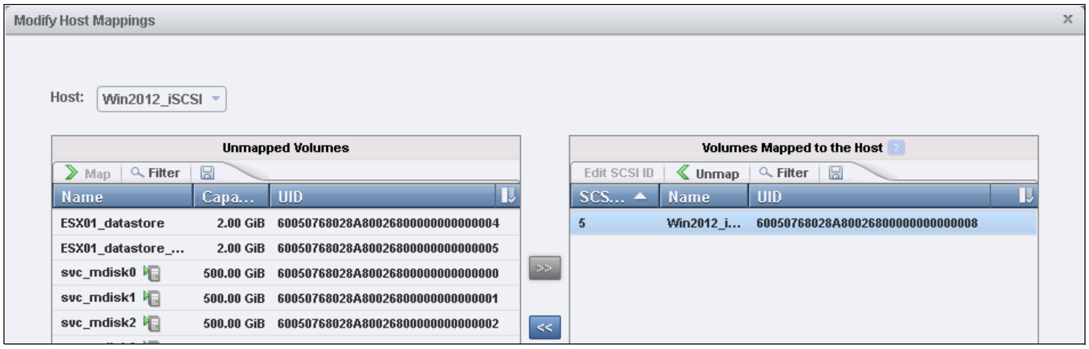


```
Task started.                                                          2:17 PM
Creating the mapping for volume Win2012_iSCSI_01 to host               2:17 PM
Win2012_iSCSI
Running command:                                                       2:17 PM
svctask mkvdiskhostmap -force -host 3 -scsi 5 8                        2:17 PM
The task is 100% complete.                                             2:17 PM
Synchronizing memory cache.                                            2:17 PM
Task completed.                                                        2:17 PM
```

*Figure 8-38 Modify Mappings task completion window*

> **Note:** The SCSI ID of the volume can be changed only before it is mapped to a host. Changing it after is not possible unless the volume is unmapped it again.

8. To remove a host mapping, the steps are similar, except that you select a volume in the right pane and click the left arrow (<<) to remove the mapping (Figure 8-39).



Host: Win2012_iSCSI

Unmapped Volumes

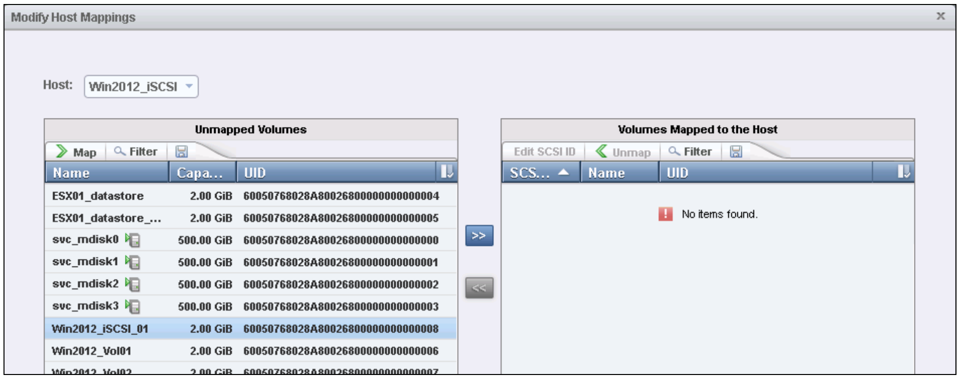
| Name | Capa... | UID |
|---|---|---|
| ESX01_datastore | 2.00 GiB | 60050768028A80026800000000000004 |
| ESX01_datastore_... | 2.00 GiB | 60050768028A80026800000000000005 |
| svc_mdisk0 | 500.00 GiB | 60050768028A80026800000000000000 |
| svc_mdisk1 | 500.00 GiB | 60050768028A80026800000000000001 |
| svc_mdisk2 | 500.00 GiB | 60050768028A80026800000000000002 |

Volumes Mapped to the Host

| SCS... | Name | UID |
|---|---|---|
| 5 | Win2012_i... | 60050768028A80026800000000000008 |

*Figure 8-39 Modify Host Mappings: Remove mapping*

The selected volume has been moved to the left pane for unmapping (Figure 8-40).



Host: Win2012_iSCSI

Unmapped Volumes

| Name | Capa... | UID |
|---|---|---|
| ESX01_datastore | 2.00 GiB | 60050768028A80026800000000000004 |
| ESX01_datastore_... | 2.00 GiB | 60050768028A80026800000000000005 |
| svc_mdisk0 | 500.00 GiB | 60050768028A80026800000000000000 |
| svc_mdisk1 | 500.00 GiB | 60050768028A80026800000000000001 |
| svc_mdisk2 | 500.00 GiB | 60050768028A80026800000000000002 |
| svc_mdisk3 | 500.00 GiB | 60050768028A80026800000000000003 |
| Win2012_iSCSI_01 | 2.00 GiB | 60050768028A80026800000000000008 |
| Win2012_Vol01 | 2.00 GiB | 60050768028A80026800000000000006 |
| Win2012_Vol02 | 2.00 GiB | 60050768028A80026800000000000007 |

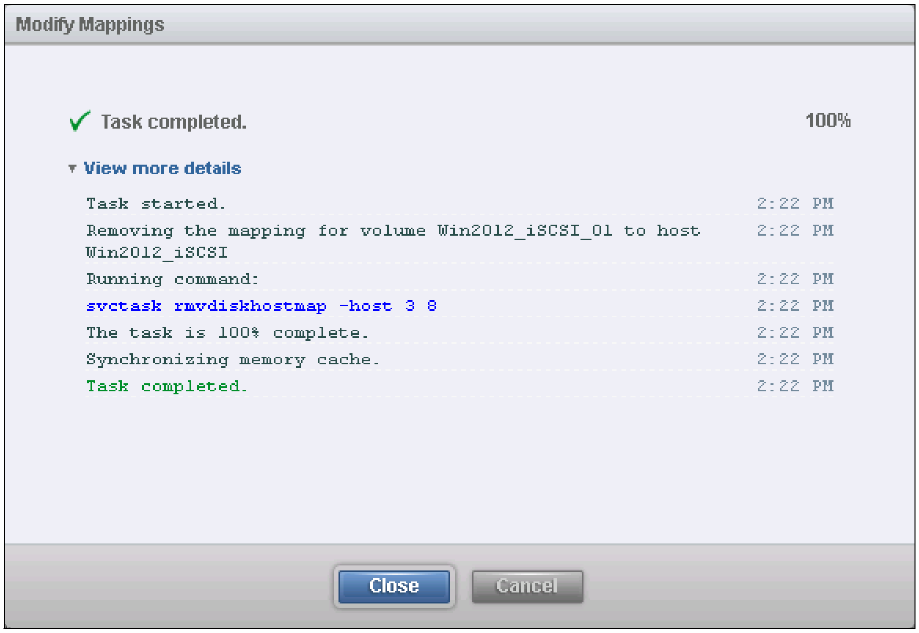
Volumes Mapped to the Host

| SCS... | Name | UID |
|---|---|---|

No items found.

*Figure 8-40 Modify Host Mappings: Mapping removed*

9. Click **Apply** to submit the changes to the system (Figure 8-41).



✓ Task completed. 100%

▾ View more details

```
Task started.                                                  2:22 PM
Removing the mapping for volume Win2012_iSCSI_01 to host       2:22 PM
Win2012_iSCSI
Running command:                                               2:22 PM
svctask rmvdiskhostmap -host 3 8                               2:22 PM
The task is 100% complete.                                     2:22 PM
Synchronizing memory cache.                                    2:22 PM
Task completed.                                                2:22 PM
```

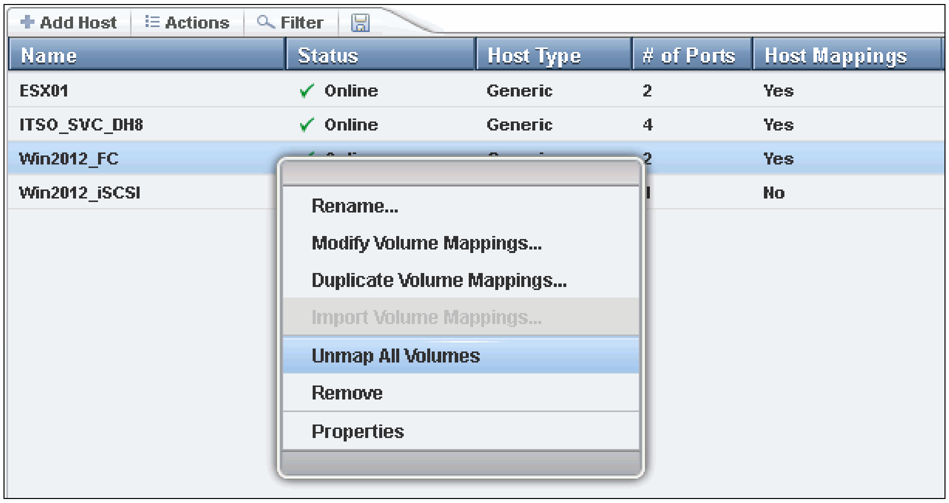
*Figure 8-41 Modify Host Mappings: Removal complete*

10. After you are done with all host mapping modifications, click **Close** to return to the Modify Mappings window (Figure 8-29 on page 307).

## Unmapping all volumes from a host

A host is able to access only those volumes on your IStorwize V7000 system that are mapped to it. If you want to remove access to all volumes for one host, regardless of how many volumes are mapped to it, you can complete this task in several simple steps:

1. From the Hosts pane, select the host, click **Actions** → **Unmap All Volumes** to remove all access that this host has to its volumes (Figure 8-42).

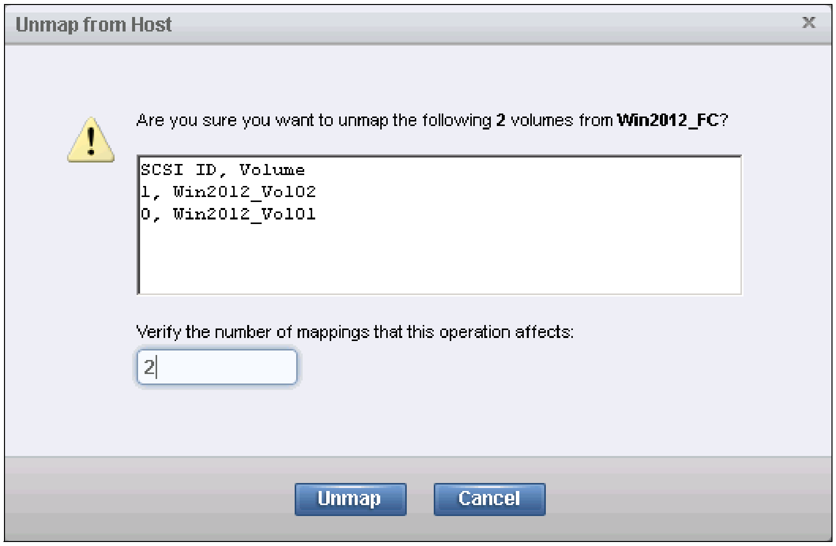


Add Host
Actions
Filter
Rename...
Modify Volume Mappings...
Duplicate Volume Mappings...
Import Volume Mappings...
Unmap All Volumes
Remove
Properties

| Name | Status | Host Type | # of Ports | Host Mappings |
|---|---|---|---|---|
| ESX01 | ✓ Online | Generic | 2 | Yes |
| ITSO_SVC_DH8 | ✓ Online | Generic | 4 | Yes |
| Win2012_FC | | | 2 | Yes |
| Win2012_iSCSI | | | 1 | No |

*Figure 8-42 Unmap All Volumes action*

2. You are prompted about the number of mappings you want to remove. Enter the number and click **Unmap** (Figure 8-43). In our example, we remove two mappings.



Are you sure you want to unmap the following **2** volumes from **Win2012_FC**?

```
SCSI ID, Volume
1, Win2012_Vol02
0, Win2012_Vol01
```

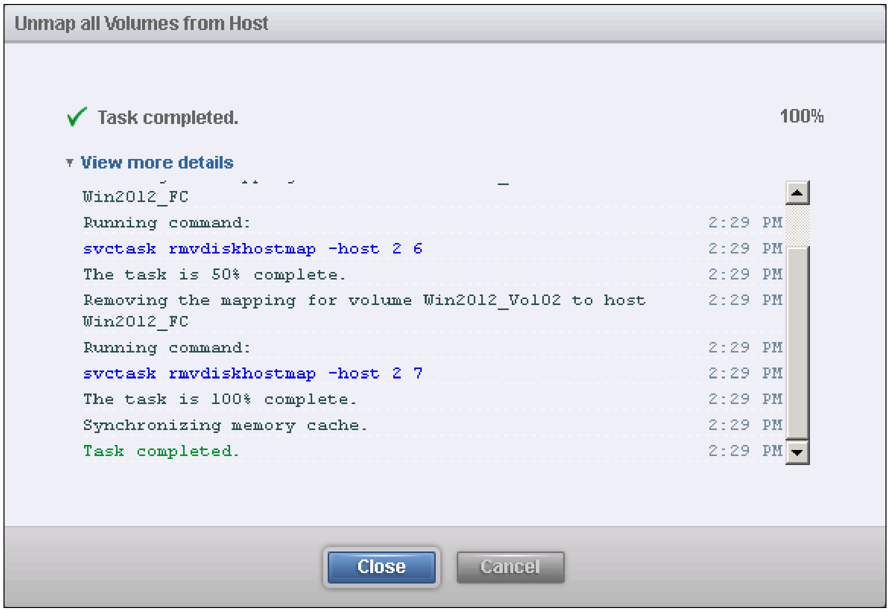
Verify the number of mappings that this operation affects:

2

*Figure 8-43 Enter the number of mappings to be removed*

> **Unmapping:** If you click **Unmap**, all access for this host to volumes that are controlled by Storwize V7000 system, is removed. Ensure that you run the required procedures on your host operating system (OS), such as unmounting file system, taking disk offline, or disabling volume group, before removing the volume mappings from your host object on IBM Spectrum Virtualize GUI.

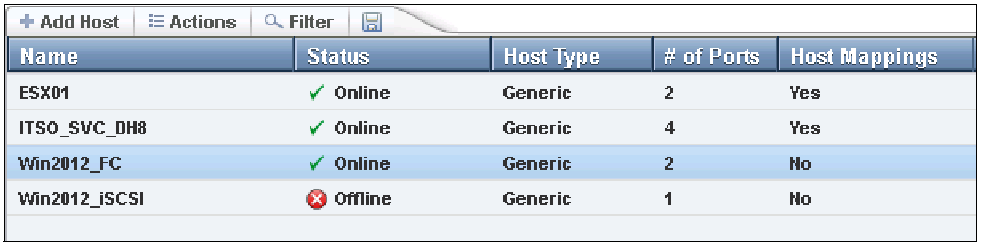
3. The changes are applied to the system (Figure 8-44). Click **Close** after you review the output.



```
Win2012_FC
Running command:                                              2:29 PM
svctask rmvdiskhostmap -host 2 6                              2:29 PM
The task is 50% complete.                                     2:29 PM
Removing the mapping for volume Win2012_Vol02 to host         2:29 PM
Win2012_FC
Running command:                                              2:29 PM
svctask rmvdiskhostmap -host 2 7                              2:29 PM
The task is 100% complete.                                    2:29 PM
Synchronizing memory cache.                                   2:29 PM
Task completed.                                               2:29 PM
```

Figure 8-44 Unmap all Volumes from Host task completion window

Figure 8-45 shows that the selected host does not have host mappings anymore.



| Name | Status | Host Type | # of Ports | Host Mappings |
|---|---|---|---|---|
| ESX01 | ✓ Online | Generic | 2 | Yes |
| ITSO_SVC_DH8 | ✓ Online | Generic | 4 | Yes |
| Win2012_FC | ✓ Online | Generic | 2 | No |
| Win2012_iSCSI | ☒ Offline | Generic | 1 | No |

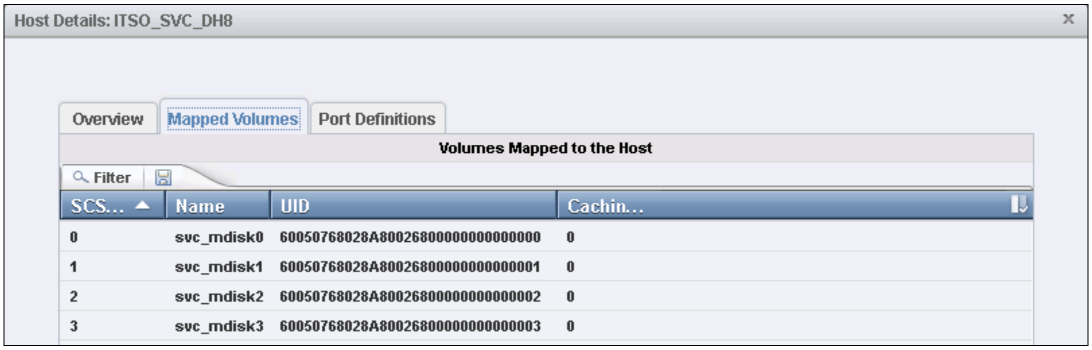
Figure 8-45 All mappings for host Win2012_FC have been removed

## Duplicating and importing mappings

Volumes that are assigned to a host can be mapped to another host object. You can do this, for example, when you add a node to the host cluster and want to ensure that the new host node has access to same set of volumes as the source host.

Verify the mappings on the existing source host object:

1. From the Hosts Actions menu (Figure 8-27 on page 306), right-click the host, select **Properties**, and click the **Mapped Volumes** tab (Figure 8-46).

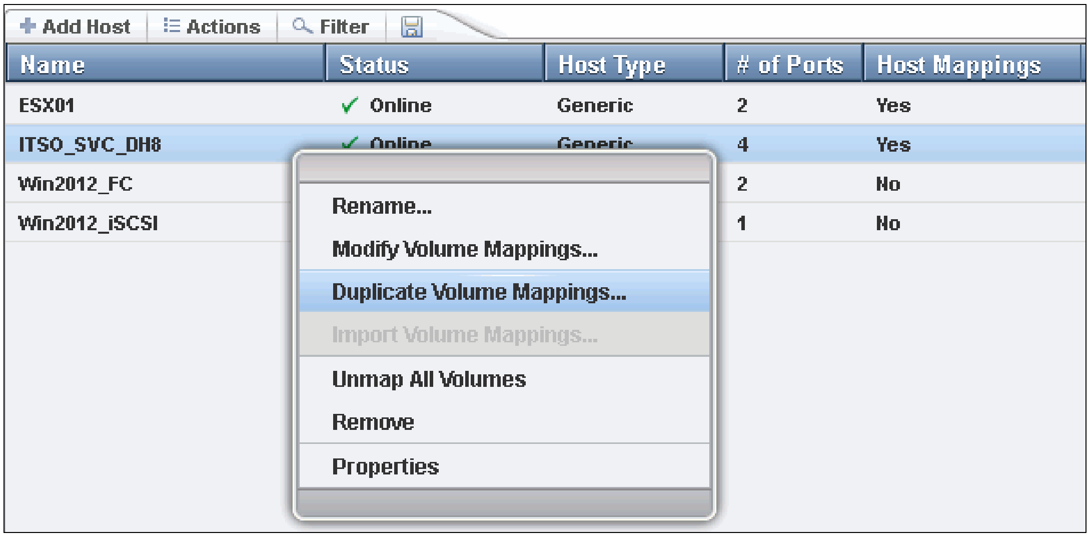


Volumes Mapped to the Host

| SCS... | Name | UID | Cachin... |
|---|---|---|---|
| 0 | svc_mdisk0 | 60050768028A80026800000000000000 | 0 |
| 1 | svc_mdisk1 | 60050768028A80026800000000000001 | 0 |
| 2 | svc_mdisk2 | 60050768028A80026800000000000002 | 0 |
| 3 | svc_mdisk3 | 60050768028A80026800000000000003 | 0 |

*Figure 8-46 Host Details: Mapped Volumes*

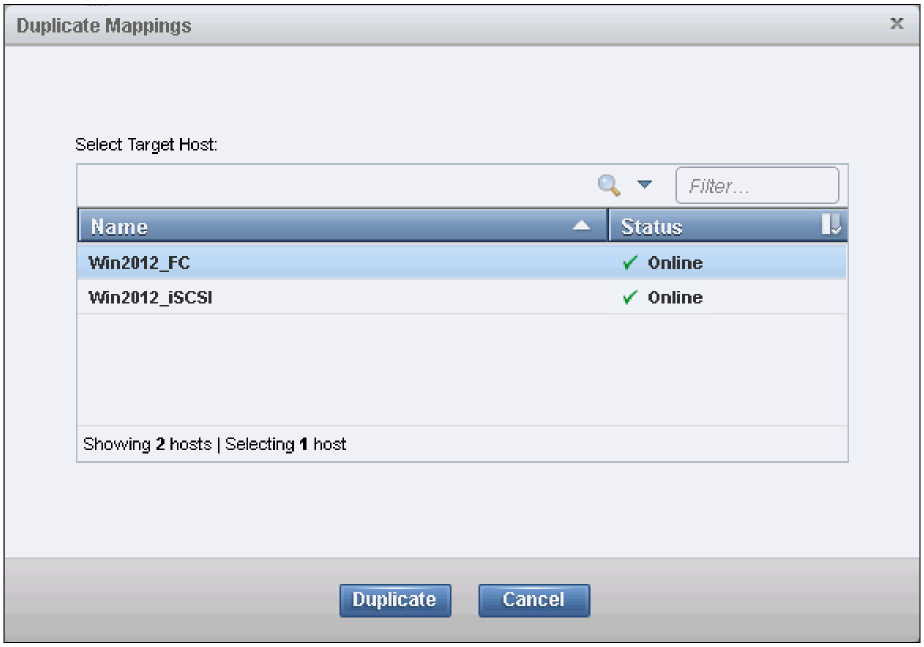
2. Select a host whose mappings you want to duplicate, and then click **Actions** → **Duplicate Volume Mappings** (Figure 8-47).


Add Host
Actions
Filter
Rename...
Modify Volume Mappings...
Duplicate Volume Mappings...
Import Volume Mappings...
Unmap All Volumes
Remove
Properties

| Name | Status | Host Type | # of Ports | Host Mappings |
|---|---|---|---|---|
| ESX01 | ✓ Online | Generic | 2 | Yes |
| ITSO_SVC_DH8 | ✓ Online | Generic | 4 | Yes |
| Win2012_FC | | | 2 | No |
| Win2012_iSCSI | | | 1 | No |

*Figure 8-47 Duplicate Volume Mappings action on source host object*

3. The Duplicate Mappings window opens. Select a target host object to which you want to add all the existing source host mappings and click **Duplicate** (Figure 8-48).



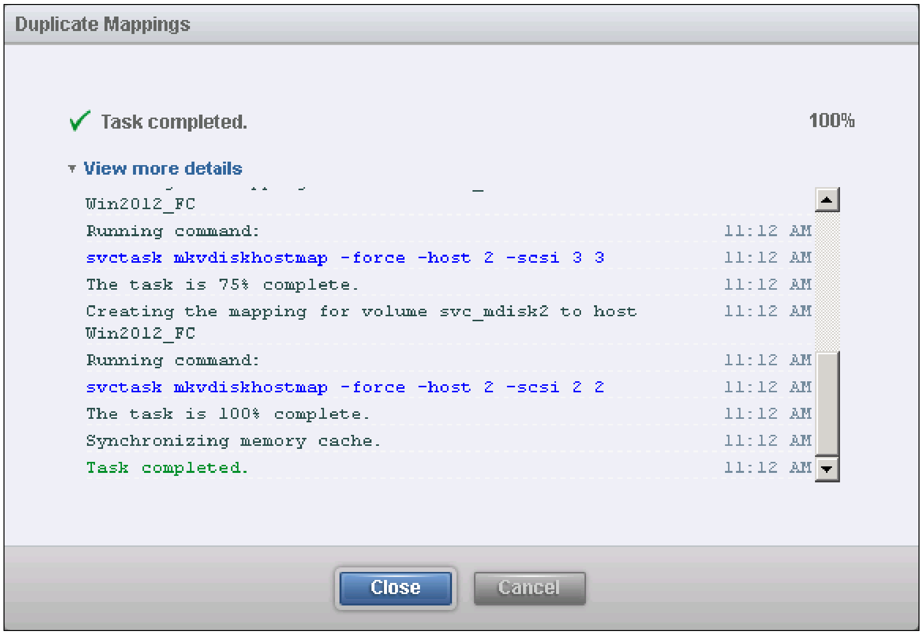
Select Target Host:

| Name | Status |
|---|---|
| Win2012_FC | ✓ Online |
| Win2012_iSCSI | ✓ Online |

Showing **2** hosts | Selecting **1** host

*Figure 8-48 Duplicate mappings window*

4. After the task completion is displayed (Figure 8-49), verify the new mappings on the target host object. From the Hosts Actions menu (Figure 8-27 on page 306), right-click the target host and select **Properties**.

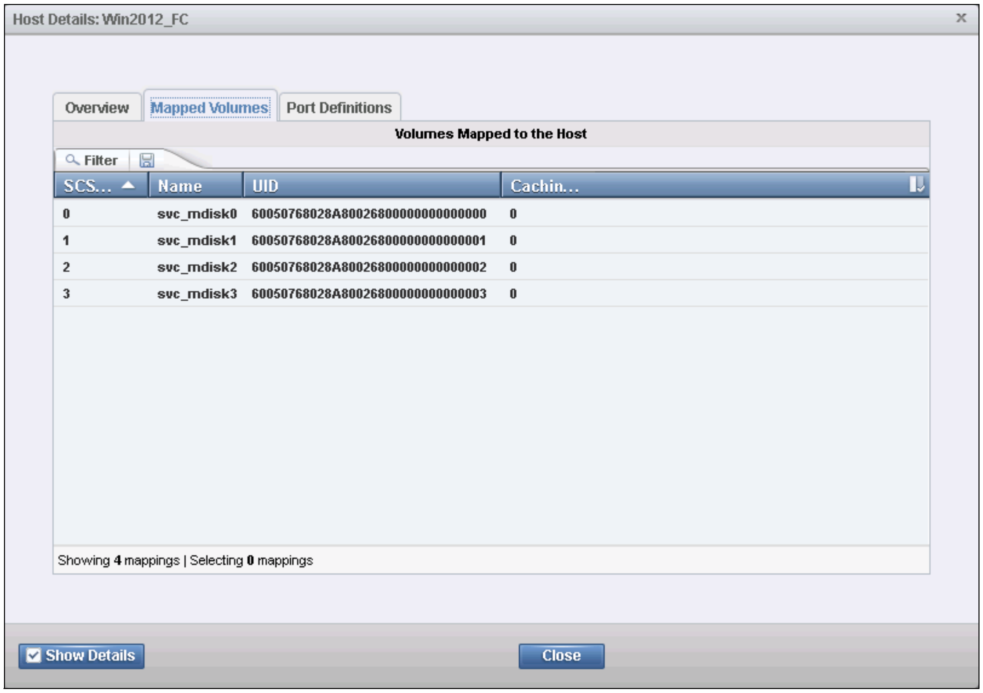


✓ Task completed. 100%

```
Win2012_FC
Running command:                                          11:12 AM
svctask mkvdiskhostmap -force -host 2 -scsi 3 3          11:12 AM
The task is 75% complete.                                 11:12 AM
Creating the mapping for volume svc_mdisk2 to host        11:12 AM
Win2012_FC
Running command:                                          11:12 AM
svctask mkvdiskhostmap -force -host 2 -scsi 2 2          11:12 AM
The task is 100% complete.                                11:12 AM
Synchronizing memory cache.                               11:12 AM
Task completed.                                           11:12 AM
```

*Figure 8-49 Duplicate Mappings task completion window*

5. Click the **Mapped Volumes** tab (Figure 8-50).



Volumes Mapped to the Host

| SCS... ▲ | Name | UID | Cachin... |
|---|---|---|---|
| 0 | svc_mdisk0 | 60050768028A80026800000000000000 | 0 |
| 1 | svc_mdisk1 | 60050768028A80026800000000000001 | 0 |
| 2 | svc_mdisk2 | 60050768028A80026800000000000002 | 0 |
| 3 | svc_mdisk3 | 60050768028A80026800000000000003 | 0 |

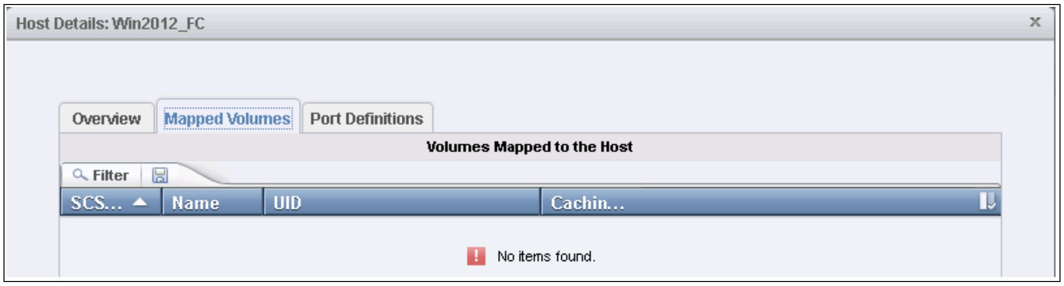
Showing **4** mappings | Selecting **0** mappings

*Figure 8-50 Host Details, new mappings on target host*

> **Note:** You can duplicate mappings only to a host that has no volumes mapped.

You can perform the same action from the Actions menu of the target host object. You can also import existing source host mappings provided the target host has no existing mappings defined.

6. Verify that no mappings are on the target host object. From the Hosts Actions menu (Figure 8-27 on page 306), right-click the host, select **Properties**, and then click the **Mapped Volumes** tab (Figure 8-51).



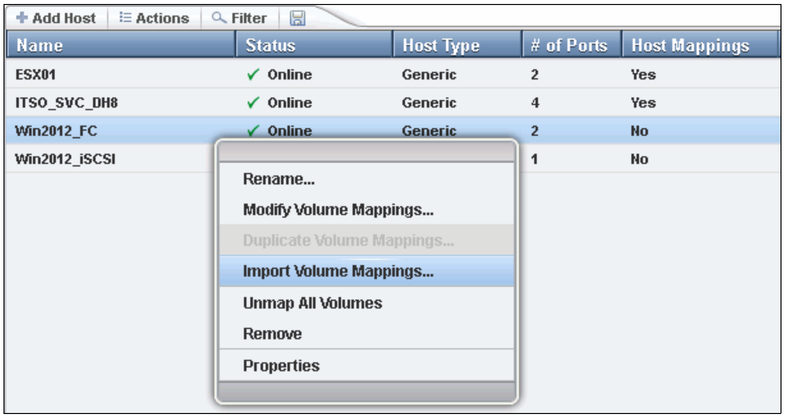
Volumes Mapped to the Host

| SCS... ▲ | Name | UID | Cachin... |
|---|---|---|---|

! No items found.

*Figure 8-51 Import Mappings, verify that there are no mappings on target host object*

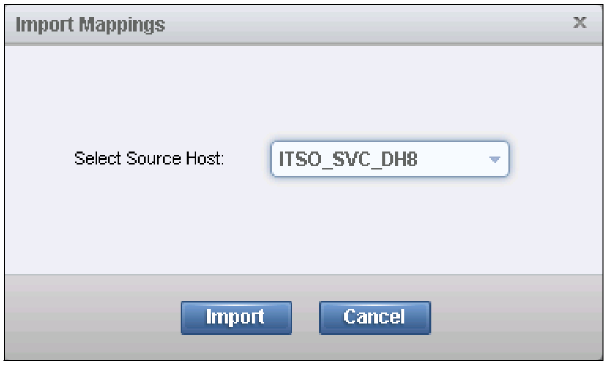
7. From the Hosts Actions menu, select a host to which you want to import existing mappings, and then click **Actions** → **Import Mappings** (Figure 8-52).


Add Host
Actions
Filter
Rename...
Modify Volume Mappings...
Duplicate Volume Mappings...
Import Volume Mappings...
Unmap All Volumes
Remove
Properties

| Name | Status | Host Type | # of Ports | Host Mappings |
|---|---|---|---|---|
| ESX01 | ✓ Online | Generic | 2 | Yes |
| ITSO_SVC_DH8 | ✓ Online | Generic | 4 | Yes |
| Win2012_FC | ✓ Online | Generic | 2 | No |
| Win2012_iSCSI | | | 1 | No |

*Figure 8-52 Import Mappings action on target host object*

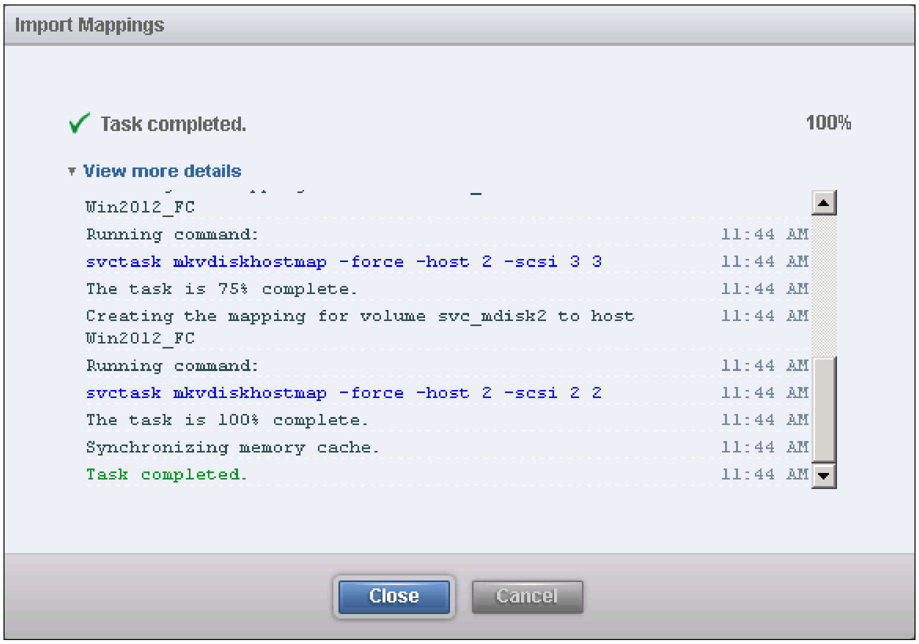
8. The Import Mappings window opens. Select the source host object from the list and click **Import** (Figure 8-53).



Select Source Host: ITSO_SVC_DH8

*Figure 8-53 Import Mappings window*

9. After the task completion window is displayed (Figure 8-54), verify the new mappings on the target host object. From the Hosts Actions menu (Figure 8-27 on page 306), right-click the target host, select **Properties**, and click the **Mapped Volumes** tab (Figure 8-55).

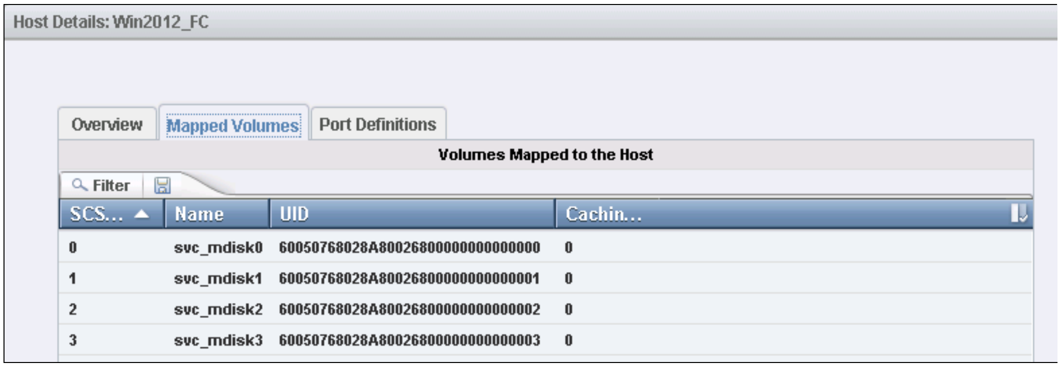


Task completed. 100%

View more details

```
Win2012_FC
Running command:                                          11:44 AM
svctask mkvdiskhostmap -force -host 2 -scsi 3 3           11:44 AM
The task is 75% complete.                                 11:44 AM
Creating the mapping for volume svc_mdisk2 to host        11:44 AM
Win2012_FC
Running command:                                          11:44 AM
svctask mkvdiskhostmap -force -host 2 -scsi 2 2           11:44 AM
The task is 100% complete.                                11:44 AM
Synchronizing memory cache.                               11:44 AM
Task completed.                                           11:44 AM
```

*Figure 8-54 Import Mappings task completion window*



Volumes Mapped to the Host

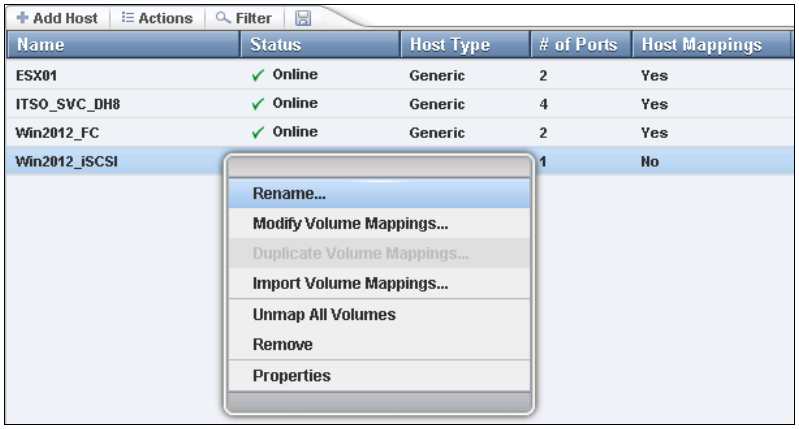
| SCS... | Name | UID | Cachin... |
|---|---|---|---|
| 0 | svc_mdisk0 | 60050768028A80026800000000000000 | 0 |
| 1 | svc_mdisk1 | 60050768028A80026800000000000001 | 0 |
| 2 | svc_mdisk2 | 60050768028A80026800000000000002 | 0 |
| 3 | svc_mdisk3 | 60050768028A80026800000000000003 | 0 |

*Figure 8-55 Host Details, new mappings on target host*

## Renaming a host

To rename a host, complete the following steps:

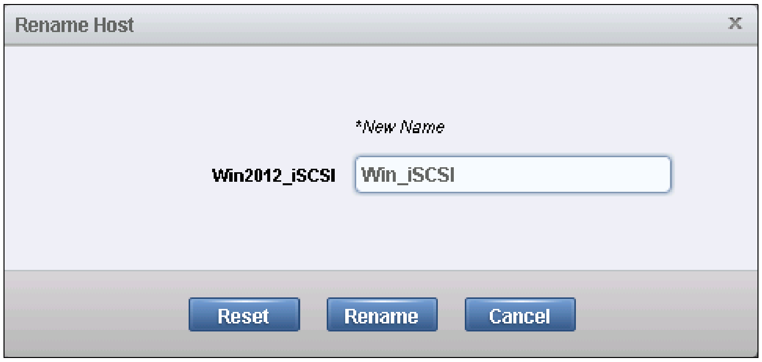
1. Select the host, and then right-click and select **Rename** (Figure 8-56).


Add Host
Actions
Filter
Rename...
Modify Volume Mappings...
Duplicate Volume Mappings...
Import Volume Mappings...
Unmap All Volumes
Remove
Properties

| Name | Status | Host Type | # of Ports | Host Mappings |
|---|---|---|---|---|
| ESX01 | ✓ Online | Generic | 2 | Yes |
| ITSO_SVC_DH8 | ✓ Online | Generic | 4 | Yes |
| Win2012_FC | ✓ Online | Generic | 2 | Yes |
| Win2012_iSCSI | | | 1 | No |

*Figure 8-56 Rename a host*

2. Enter a new name and click **Rename** (Figure 8-57). If you click **Reset**, the changes are reset to the original host name.

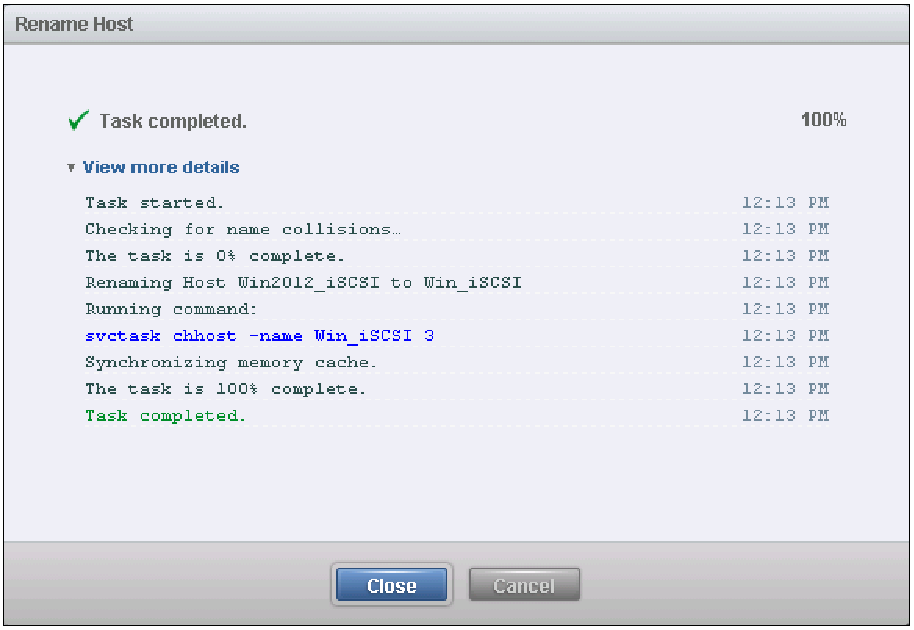


*New Name

Win2012_iSCSI Win_iSCSI

*Figure 8-57 Rename Host window*

3. After the changes are applied to the system, click **Close** (Figure 8-58).



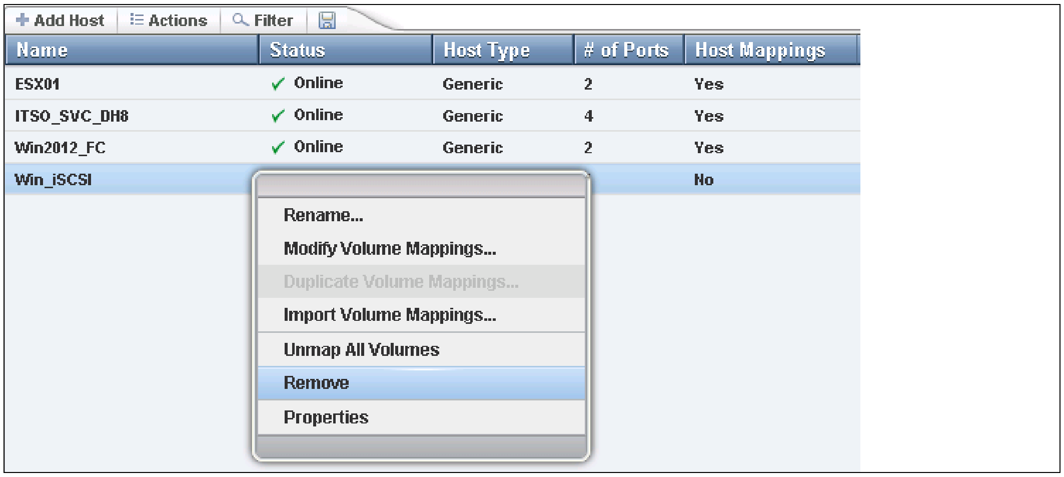
```
Task started.                                            12:13 PM
Checking for name collisions…                            12:13 PM
The task is 0% complete.                                 12:13 PM
Renaming Host Win2012_iSCSI to Win_iSCSI                 12:13 PM
Running command:                                         12:13 PM
svctask chhost -name Win_iSCSI 3                         12:13 PM
Synchronizing memory cache.                              12:13 PM
The task is 100% complete.                               12:13 PM
Task completed.                                          12:13 PM
```

*Figure 8-58 Rename Host task completion window*

## Removing a host

To remove a host, complete the following steps:

1. From the Hosts pane, select the host and right-click it or click **Actions** → **Remove** (Figure 8-59).

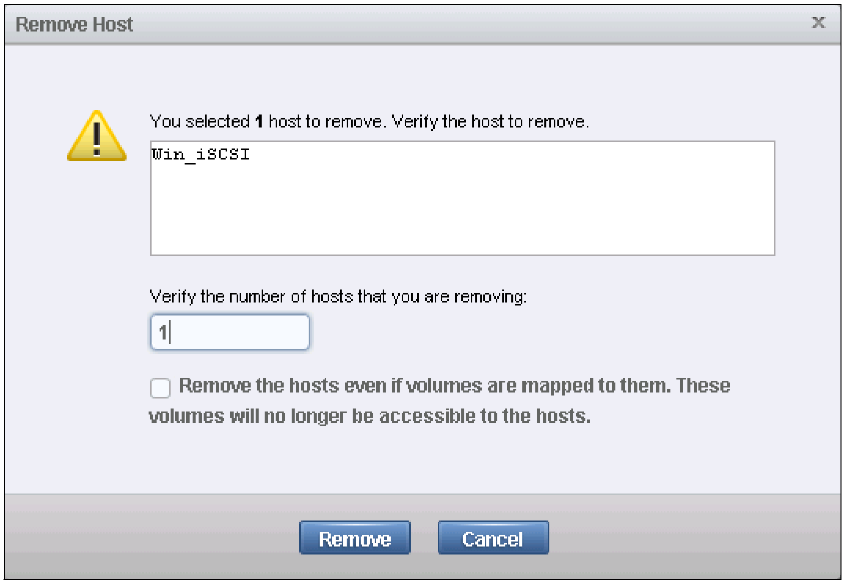


Add Host
Actions
Filter
Rename...
Modify Volume Mappings...
Duplicate Volume Mappings...
Import Volume Mappings...
Unmap All Volumes
Remove
Properties

| Name | Status | Host Type | # of Ports | Host Mappings |
|---|---|---|---|---|
| ESX01 | ✓ Online | Generic | 2 | Yes |
| ITSO_SVC_DH8 | ✓ Online | Generic | 4 | Yes |
| Win2012_FC | ✓ Online | Generic | 2 | Yes |
| Win_iSCSI | | | | No |

*Figure 8-59 Remove a host*

2. Confirm the number of hosts that you want to remove and click **Delete** (Figure 8-60).



You selected 1 host to remove. Verify the host to remove.

Win_iSCSI

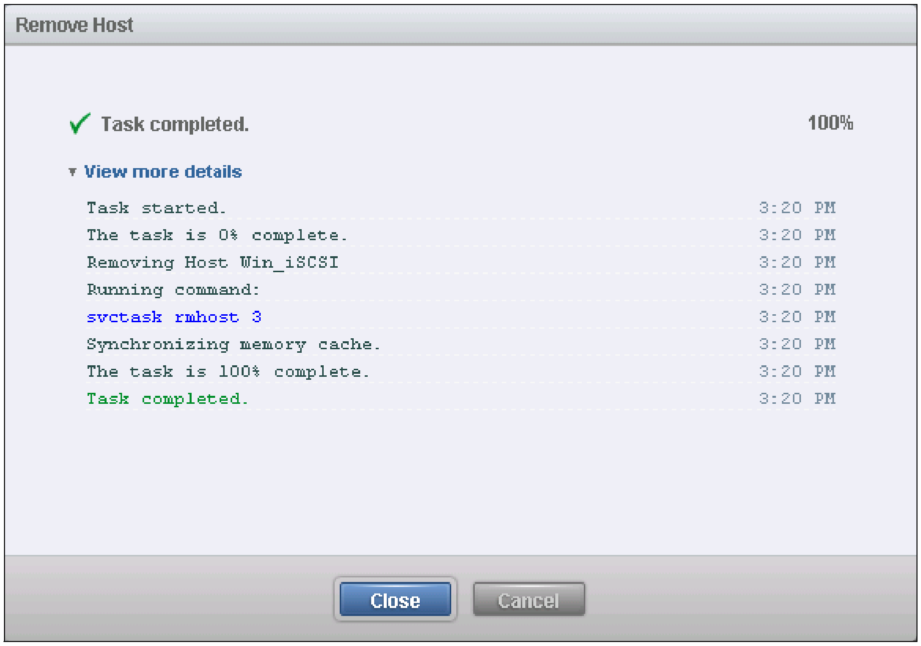
Verify the number of hosts that you are removing:

1

□ Remove the hosts even if volumes are mapped to them. These volumes will no longer be accessible to the hosts.

*Figure 8-60 Confirm the removal of the host*

3. If you want to remove a host that has volumes mapped, you must force the removal by selecting the check box in the lower part of the window. If you select this check box, the host is removed and it no longer has access to this system.
4. After the task is completed, click **Close** (Figure 8-61) to return to the mappings window.



✓ Task completed. 100%

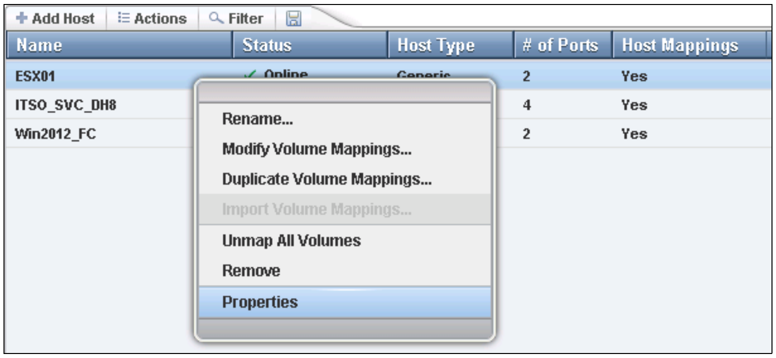
```
Task started.                                   3:20 PM
The task is 0% complete.                        3:20 PM
Removing Host Win_iSCSI                         3:20 PM
Running command:                                3:20 PM
svctask rmhost 3                                3:20 PM
Synchronizing memory cache.                     3:20 PM
The task is 100% complete.                      3:20 PM
Task completed.                                 3:20 PM
```

*Figure 8-61 Remove host task completion window*

## Host properties

To access host properties, complete the following steps:

1. From the IBM Spectrum Virtualize GUI Hosts pane, select the host, and right-click it or click **Actions** → **Properties** (Figure 8-62).

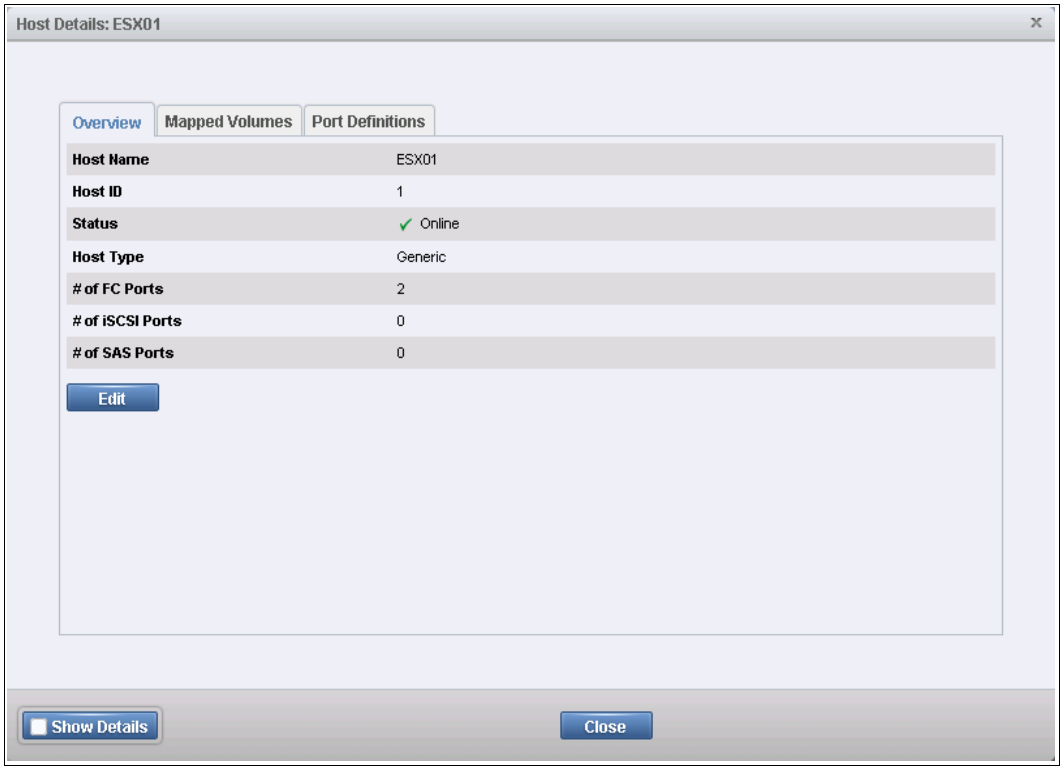


Add Host
Actions
Filter
Rename...
Modify Volume Mappings...
Duplicate Volume Mappings...
Import Volume Mappings...
Unmap All Volumes
Remove
Properties

| Name | Status | Host Type | # of Ports | Host Mappings |
|---|---|---|---|---|
| ESX01 | ✓ Online | Generic | 2 | Yes |
| ITSO_SVC_DH8 | | | 4 | Yes |
| Win2012_FC | | | 2 | Yes |

*Figure 8-62 Host properties*

The Host Details window opens (Figure 8-63).

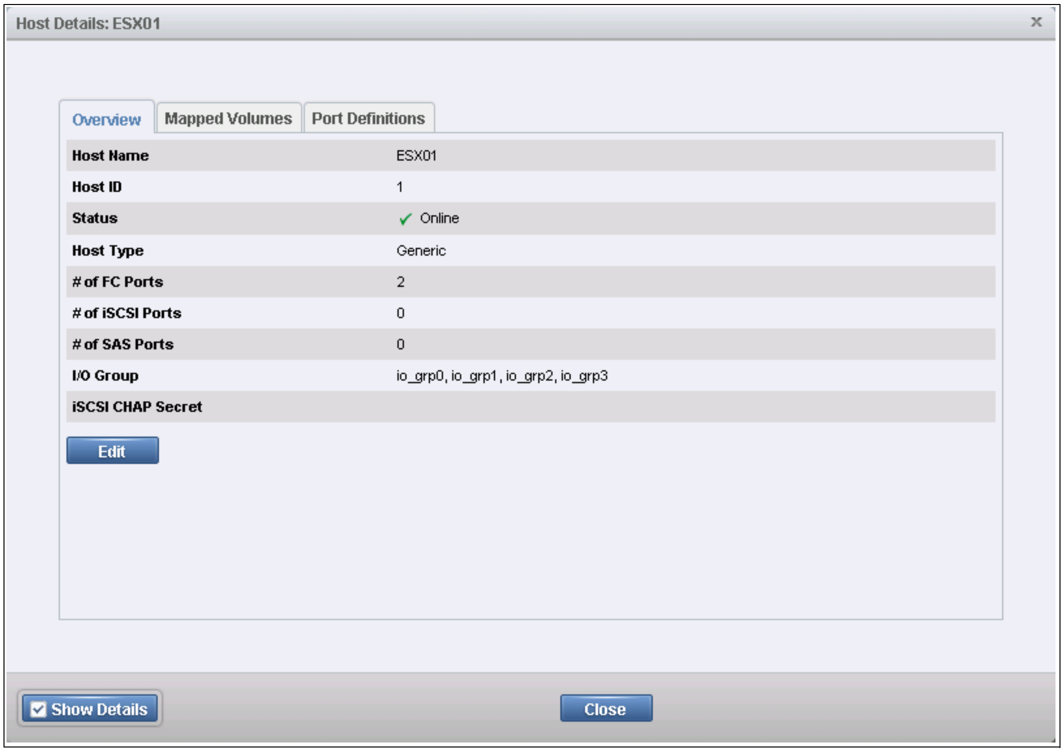


| | |
|---|---|
| Host Name | ESX01 |
| Host ID | 1 |
| Status | ✓ Online |
| Host Type | Generic |
| # of FC Ports | 2 |
| # of iSCSI Ports | 0 |
| # of SAS Ports | 0 |

*Figure 8-63 Host properties overview*

The Host Details window shows an overview of your host properties. It has three tabs: Overview, Mapped Volumes, and Port Definitions. The Overview tab is shown in Figure 8-63.

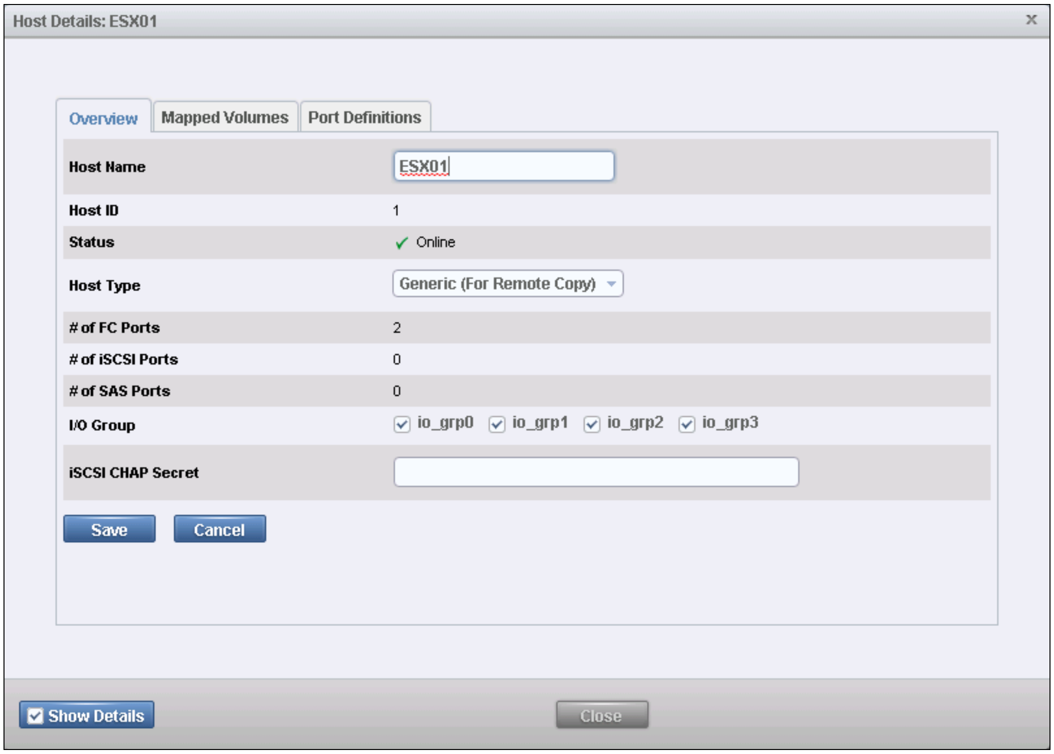
2. Select the **Show Details** check box to see more information about the host (Figure 8-64).



| | |
|---|---|
| Host Name | ESX01 |
| Host ID | 1 |
| Status | ✓ Online |
| Host Type | Generic |
| # of FC Ports | 2 |
| # of iSCSI Ports | 0 |
| # of SAS Ports | 0 |
| I/O Group | io_grp0, io_grp1, io_grp2, io_grp3 |
| iSCSI CHAP Secret | |

*Figure 8-64 Host Properties: Show details*

3. Click **Edit** if you want to change any host properties (Figure 8-65).



| | |
|---|---|
| Host Name | ESX01 |
| Host ID | 1 |
| Status | ✓ Online |
| Host Type | Generic (For Remote Copy) |
| # of FC Ports | 2 |
| # of iSCSI Ports | 0 |
| # of SAS Ports | 0 |
| I/O Group | ☑ io_grp0 ☑ io_grp1 ☑ io_grp2 ☑ io_grp3 |
| iSCSI CHAP Secret | |

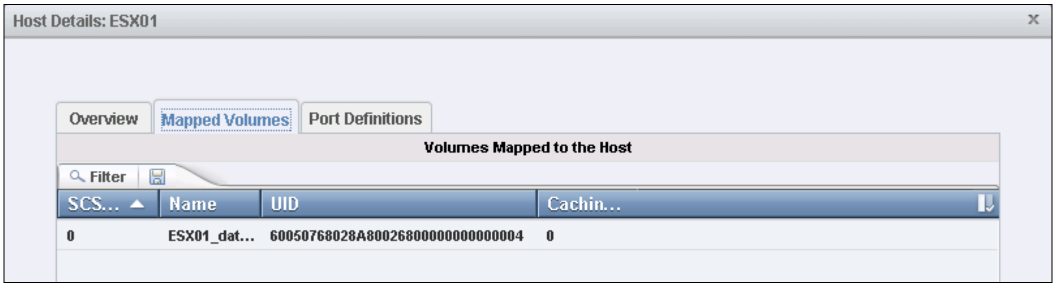
*Figure 8-65 Edit host properties*

In this window (shown in Figure 8-65 on page 324), you can modify the following items:

- Host Name. Change the host name.
- Host Type. If you are going to attach HP/UX, OpenVMS, or Target Port Group Support (TPGS) hosts, change this setting.
- I/O Group. Host has access to volumes mapped from selected I/O Groups.
- iSCSI Challenge Handshake Authentication Protocol (CHAP) Secret. Enter or change the iSCSI CHAP secret for this host.

4. When you finish making changes (if required), click **Save** to apply them (Figure 8-65 on page 324). The editing window closes.

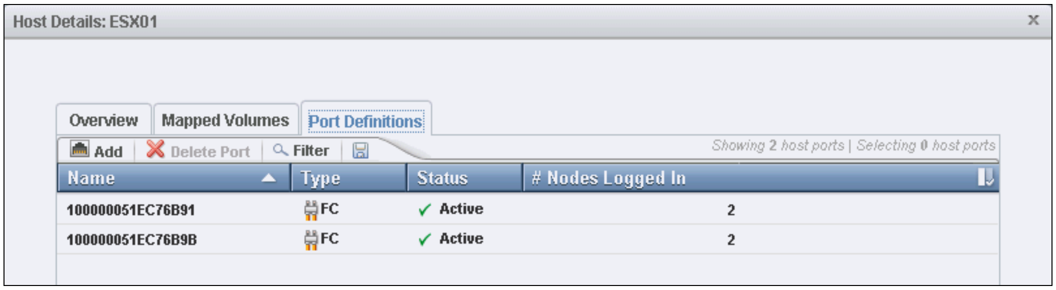
   The Mapped Volumes tab shows an overview of which volumes are currently mapped with which SCSI ID and UID to this host (Figure 8-66). The **Show Details** check box does not show any additional information.



| SCS... | Name | UID | Cachin... |
|---|---|---|---|
| 0 | ESX01_dat... | 60050768028A80026800000000000004 | 0 |

*Figure 8-66 Mapped volumes tab*

The Port Definitions tab shows the configured host ports of a host and gives you status information about them (Figure 8-67).



| Name | Type | Status | # Nodes Logged In |
|---|---|---|---|
| 100000051EC76B91 | FC | ✓ Active | 2 |
| 100000051EC76B9B | FC | ✓ Active | 2 |

*Figure 8-67 Port definitions*

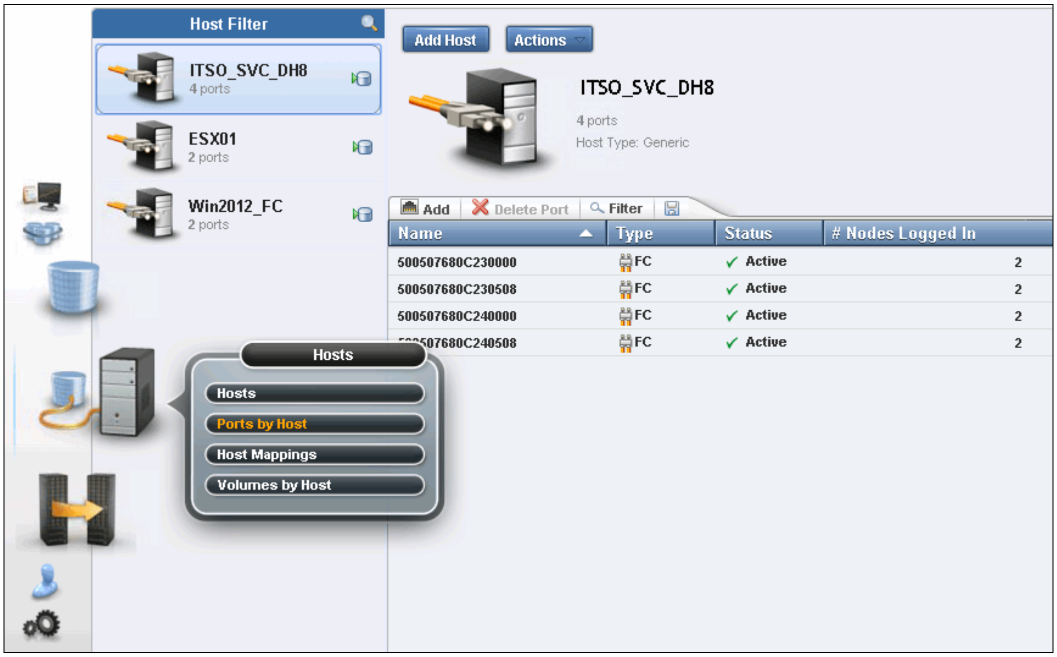
This window offers you the option to start Add and Delete Port actions, as described in 8.4.3, “Adding and deleting host ports” on page 325.

5. Click **Close** to close the Host Details window.

## 8.4.3 Adding and deleting host ports

To configure host ports, complete the following steps:

1. From the dynamic menu hover over **Hosts** and select **Ports by Host** to open the associated pane (Figure 8-68).

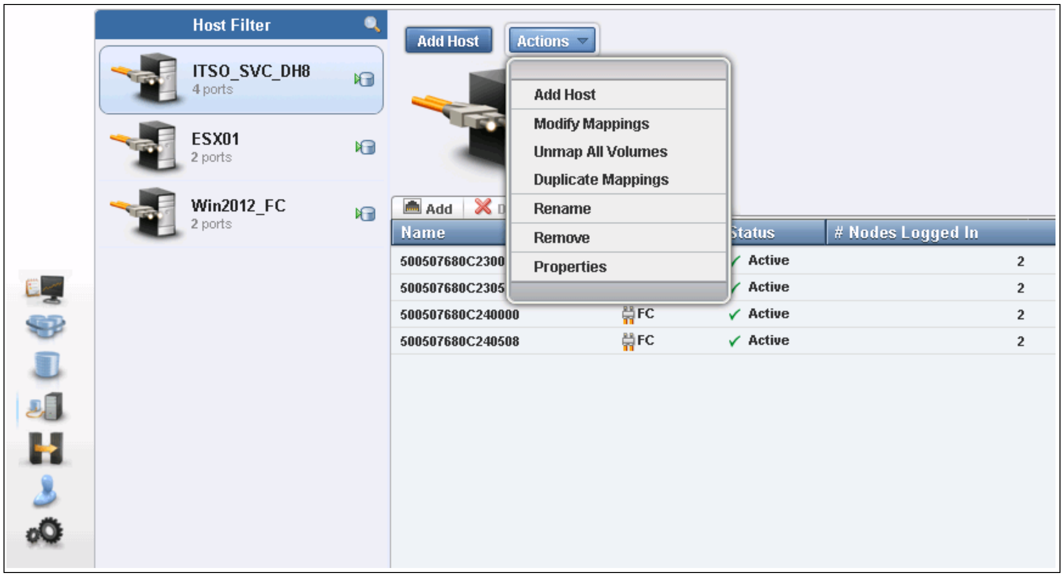


Host Type: Generic

| Name | Type | Status | # Nodes Logged In |
|---|---|---|---|
| 500507680C230000 | FC | ✓ Active | 2 |
| 500507680C230508 | FC | ✓ Active | 2 |
| 500507680C240000 | FC | ✓ Active | 2 |
| 500507680C240508 | FC | ✓ Active | 2 |

*Figure 8-68 Ports by Host*

2. The left pane lists all the hosts; the function icons indicate whether the host is Fibre Channel (*orange* cable) or iSCSI (*blue* cable). The properties of the highlighted host are shown in the right pane. If you click **New Host**, the wizard starts. If you click **Actions** (Figure 8-69), the tasks described in , “Modifying Mappings menu” on page 306 can be started.



| Name | Type | Status | # Nodes Logged In |
|---|---|---|---|
| 500507680C2300 | | ✓ Active | 2 |
| 500507680C2305 | | ✓ Active | 2 |
| 500507680C240000 | FC | ✓ Active | 2 |
| 500507680C240508 | FC | ✓ Active | 2 |

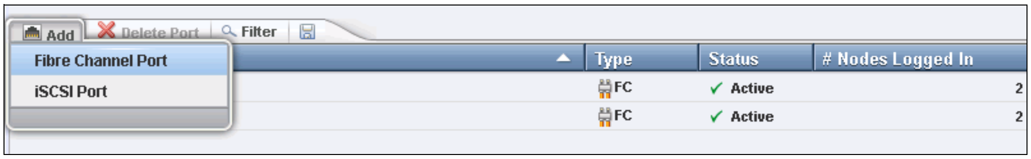
*Figure 8-69 Ports by Host actions*

## Adding a Fibre Channel or iSCSI host port

To add a host port, complete the following steps:

1. Highlight the host.
2. Click **Add** (Figure 8-70) and select one of the two options:

a. Select Fibre Channel Port (see “Adding a Fibre Channel port”).
b. Select iSCSI Port (see “Adding an iSCSI host port” on page 330).



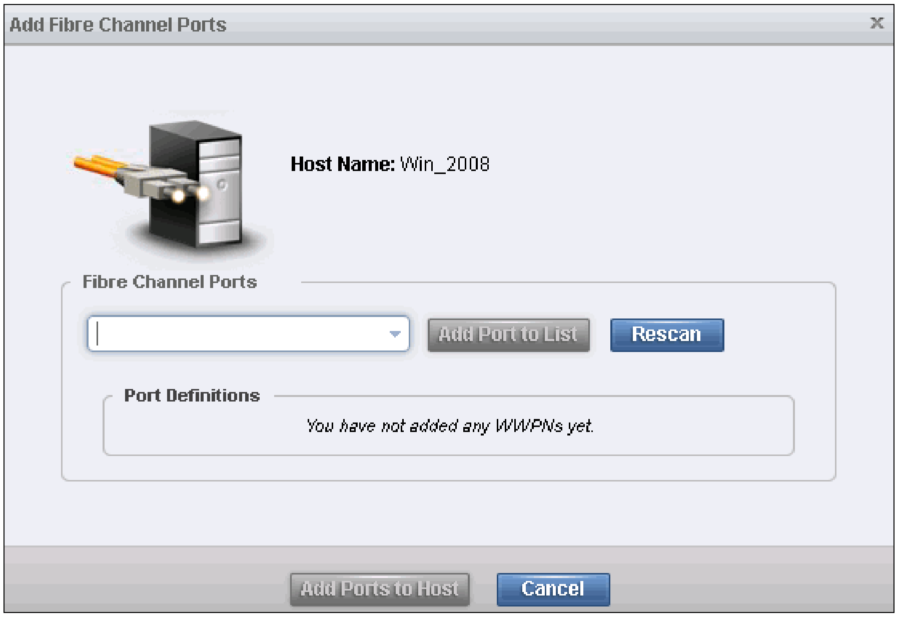
| Type | Status | # Nodes Logged In |
|---|---|---|
| FC | ✓ Active | 2 |
| FC | ✓ Active | 2 |

*Figure 8-70 Add host ports*

## Adding a Fibre Channel port

To add a Fibre Channel port, complete the following steps:

1. Click **Fibre Channel Port** (Figure 8-70). The Add Fibre Channel Ports window opens (Figure 8-71).



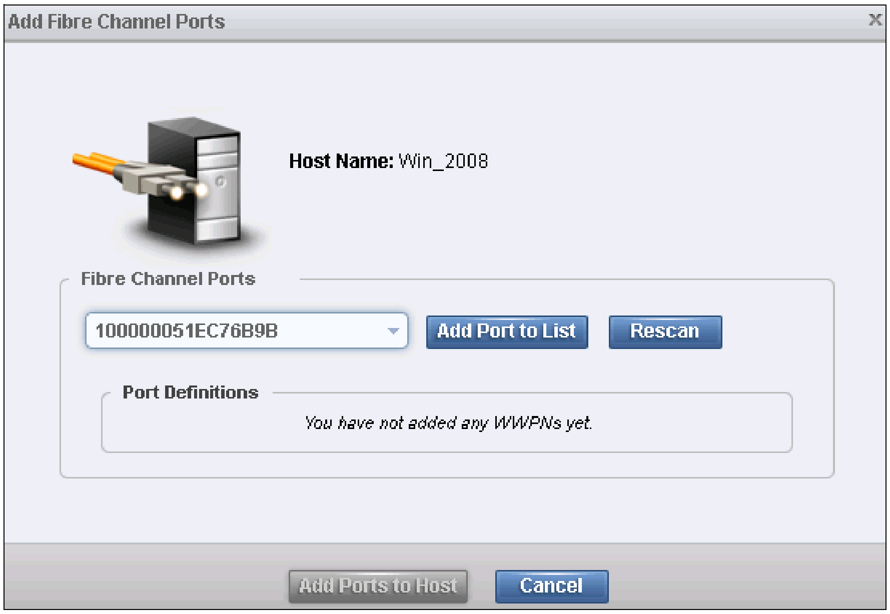
**Host Name:** Win_2008

*You have not added any WWPNs yet.*

*Figure 8-71 Add Fibre Channel Ports window*

2. If you click the drop-down menu, a list of all known Fibre Channel host ports (Figure 8-72) displays. If the worldwide port name (WWPN) of your host is not available in the menu, check your SAN zoning and rescan the SAN from the host.

3. Then, click **Rescan**. The new port is now available in the menu.

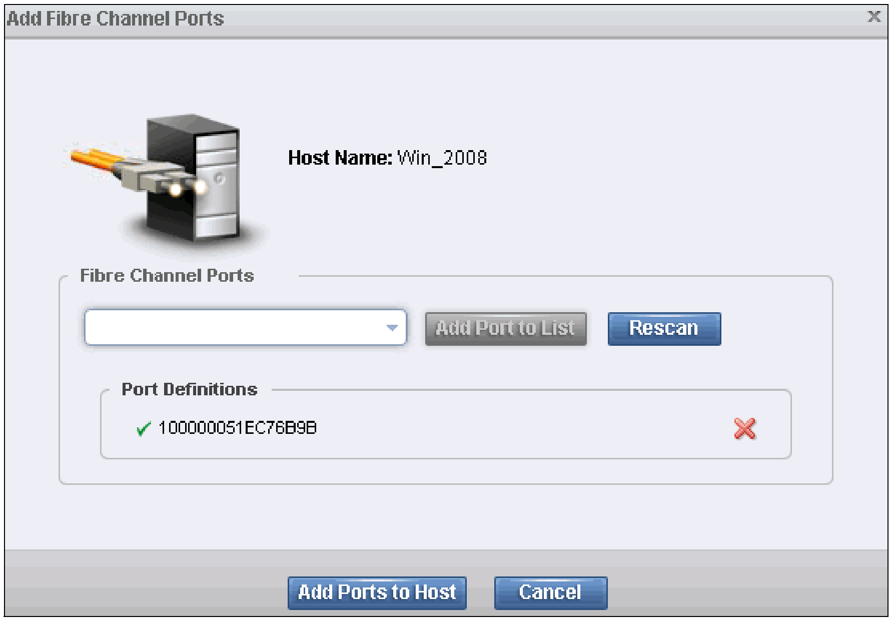


Host Name: Win_2008

You have not added any WWPNs yet.

*Figure 8-72 Add Fibre Channel Ports: Known WWPNs*

4. Select the WWPN you want to add and click **Add Port to List** (Figure 8-73).



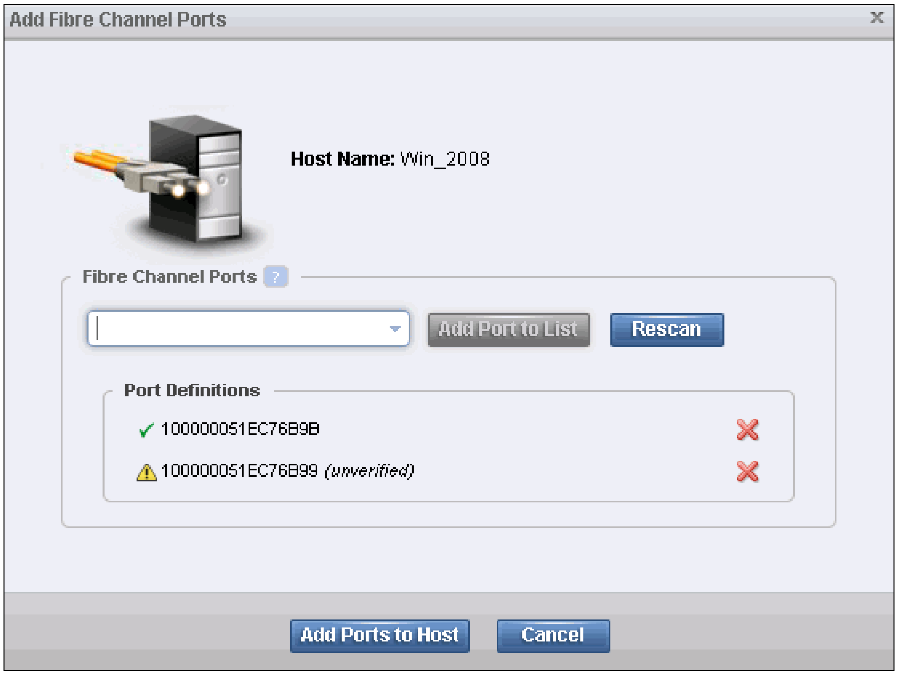
Host Name: Win_2008

✓ 100000051EC76B9B ✗

*Figure 8-73 Add a port to list*

You can repeat this step to add more ports to a host.

5. If you want to add an offline port (if the WWPN of your host is not available in the drop-down menu), manually enter the WWPN of the port into the Fibre Channel Ports field and click **Add Port to List**. The port is unverified (Figure 8-74) because it is not logged on to the Storwize V7000. The first time that it logs on, its state is automatically changed to online, and the mapping is applied to this port.



Add Fibre Channel Ports

**Host Name:** Win_2008

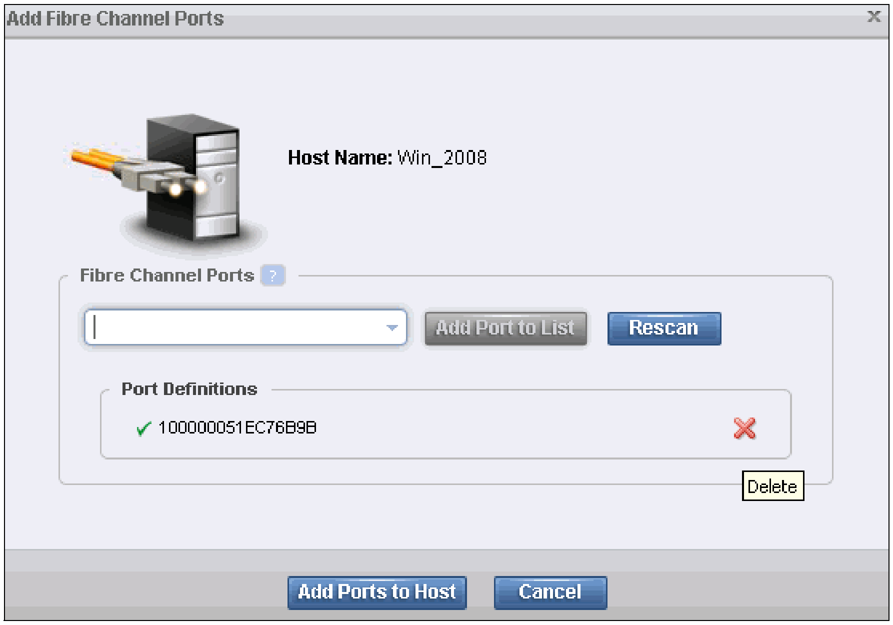
Port Definitions

✓ 100000051EC76B9B

⚠ 100000051EC76B99 *(unverified)*

*Figure 8-74 Unverified port*

6. To remove a port from the list, click the red X next to the port (Figure 8-75). In this example, delete the manually added FC port so only the detected port remains.



Add Fibre Channel Ports

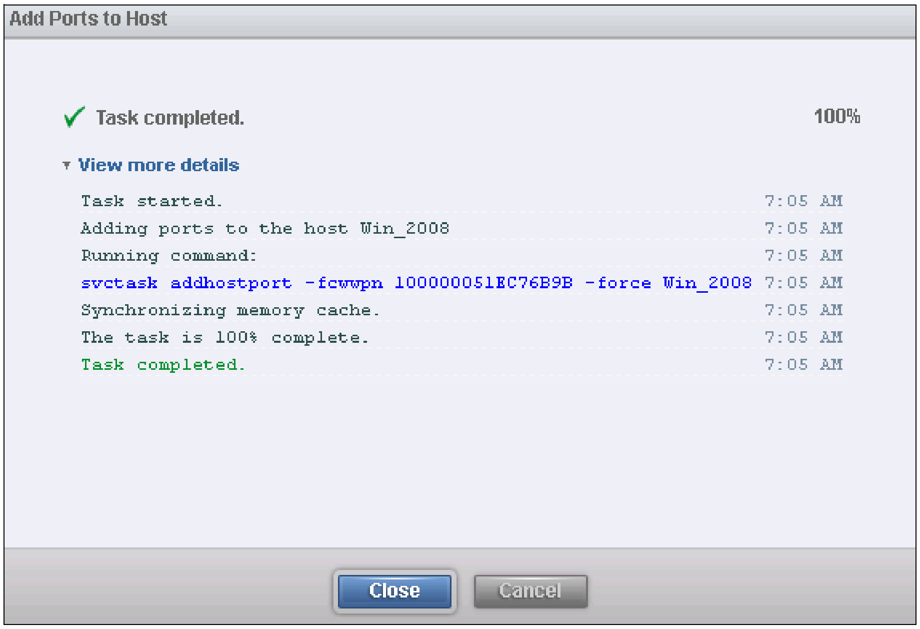
**Host Name:** Win_2008

Port Definitions

✓ 100000051EC76B9B

*Figure 8-75 Remove a port from a list*

7. Click **Add Ports to Host** and the changes are applied (Figure 8-76).



```
Task started.                                           7:05 AM
Adding ports to the host Win_2008                       7:05 AM
Running command:                                        7:05 AM
svctask addhostport -fcwwpn 10000005lEC76B9B -force Win_2008 7:05 AM
Synchronizing memory cache.                             7:05 AM
The task is 100% complete.                              7:05 AM
Task completed.                                         7:05 AM
```

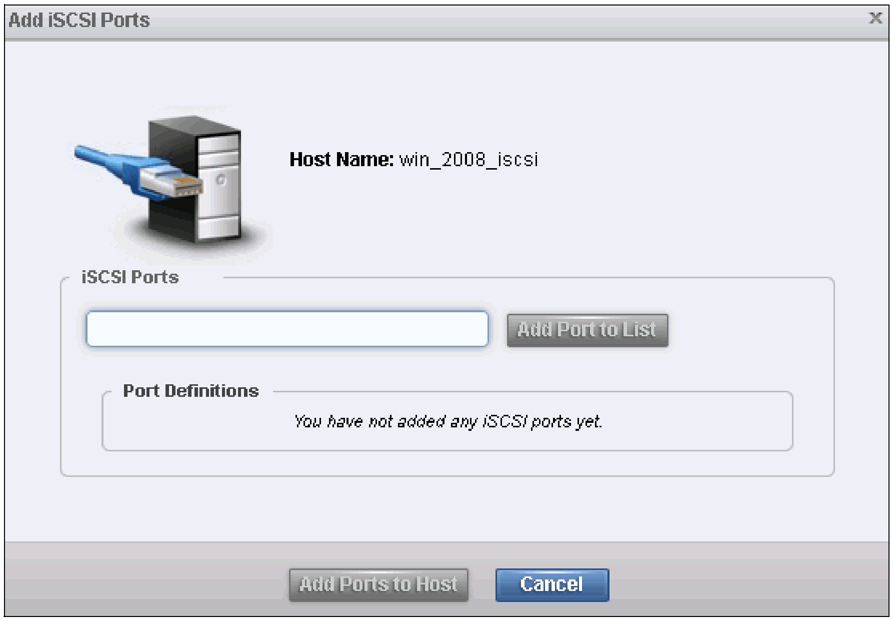
*Figure 8-76 Add Ports to Host task completion window*

8. Click **Close** to return to the Ports to Host window.

## Adding an iSCSI host port

To add an iSCSI host port, complete the following steps:

1. Click **iSCSI Port** (Figure 8-70 on page 327). The Add iSCSI Ports window opens (Figure 8-77).

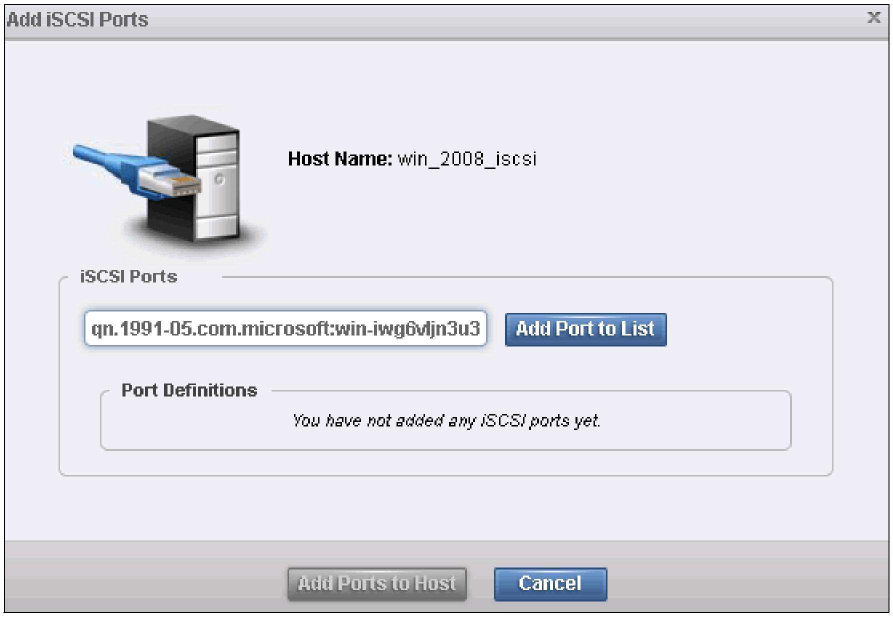


**Host Name:** win_2008_iscsi

*You have not added any iSCSI ports yet.*

*Figure 8-77 Add iSCSI host ports*

2. Enter the initiator name of your host (Figure 8-78) and click **Add Port to List**.



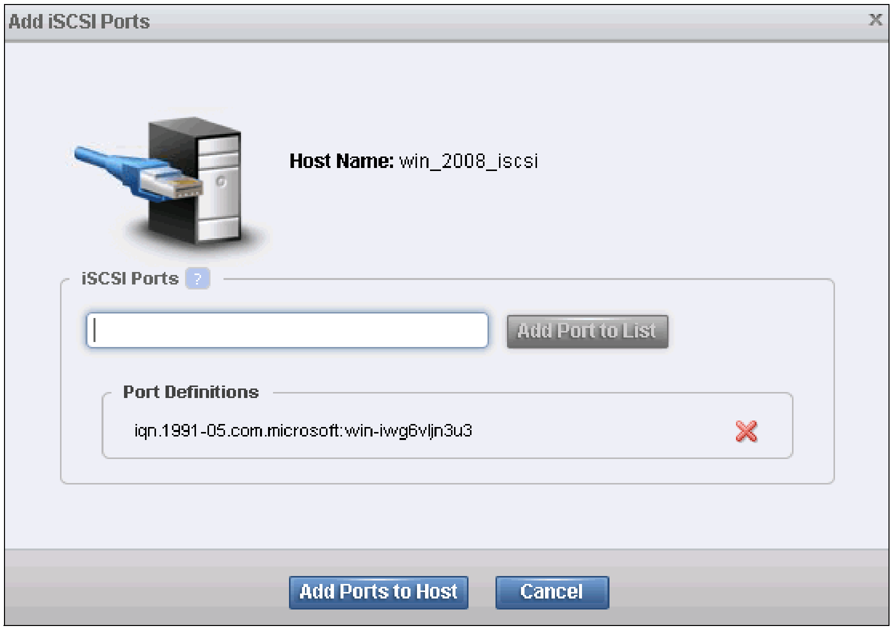
Host Name: win_2008_iscsi

qn.1991-05.com.microsoft:win-iwg6vljn3u3

You have not added any iSCSI ports yet.

*Figure 8-78 Enter the initiator name*

3. Click **Add Ports to Host** (Figure 8-79).

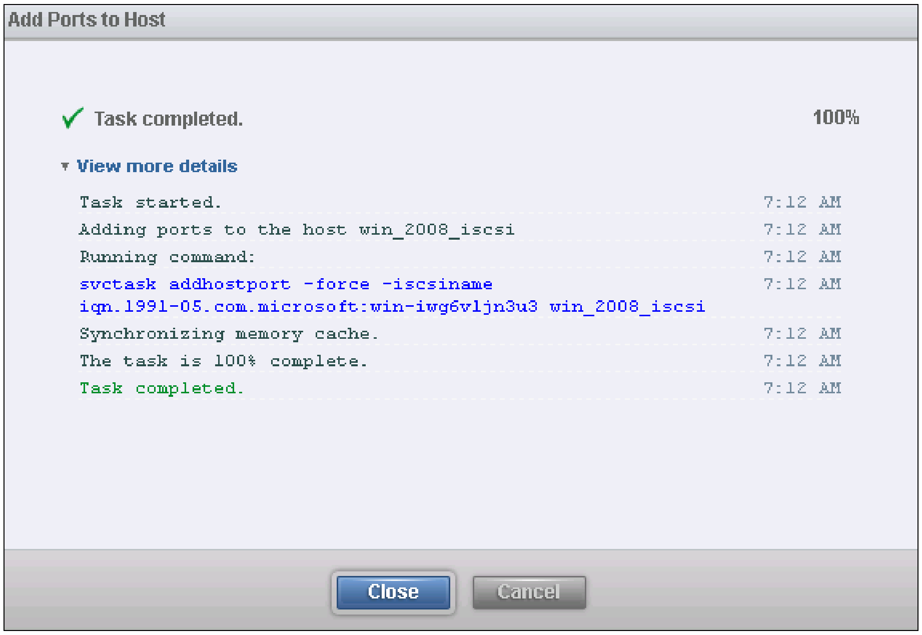


Host Name: win_2008_iscsi

iqn.1991-05.com.microsoft:win-iwg6vljn3u3

*Figure 8-79 Add ports to the ports definitions*

4. The tasks are completed and changes to the system are applied (Figure 8-80). Click **Close** to return to the Ports by Host window.



✓ Task completed. 100%

▾ View more details

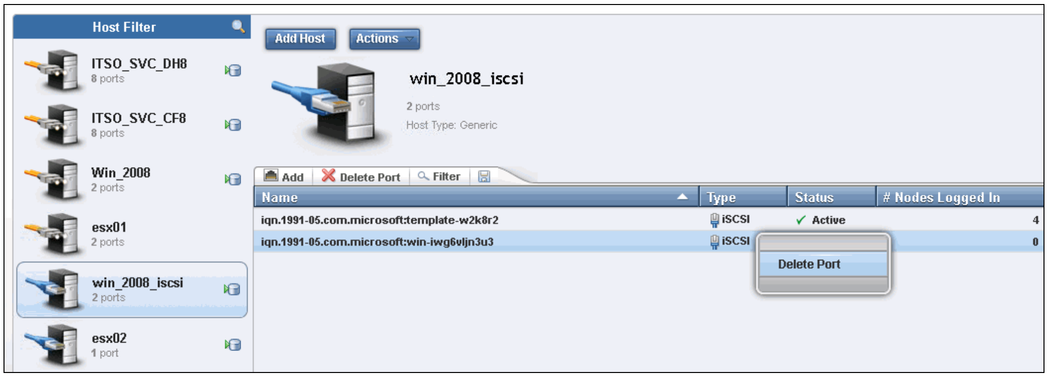
```
Task started.                                                   7:12 AM
Adding ports to the host win_2008_iscsi                         7:12 AM
Running command:                                                7:12 AM
svctask addhostport -force -iscsiname                           7:12 AM
iqn.1991-05.com.microsoft:win-iwg6vljn3u3 win_2008_iscsi
Synchronizing memory cache.                                     7:12 AM
The task is 100% complete.                                      7:12 AM
Task completed.                                                 7:12 AM
```

*Figure 8-80 Add Ports to Host task completion window*

## Deleting a host port

To delete a host port, complete the following steps:

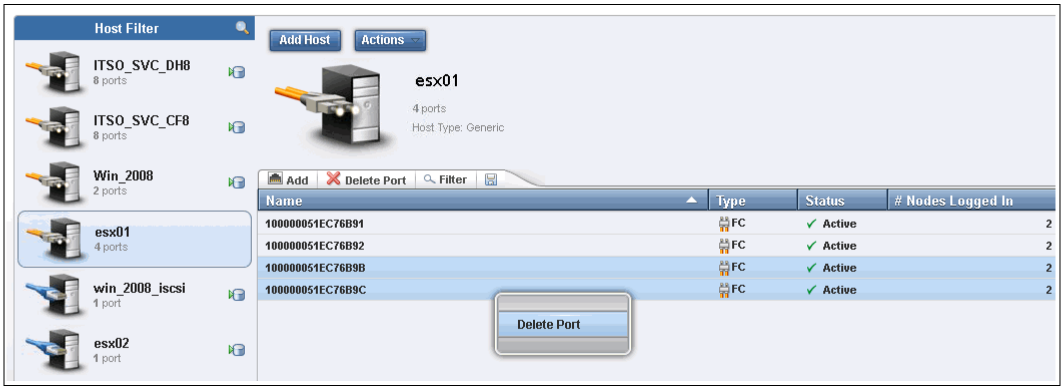
1. Highlight it and right-click it or click **Delete Port** (Figure 8-81).



| Name | Type | Status | # Nodes Logged In |
|---|---|---|---|
| iqn.1991-05.com.microsoft:template-w2k8r2 | iSCSI | ✓ Active | 4 |
| iqn.1991-05.com.microsoft:win-iwg6vljn3u3 | iSCSI | | 0 |

*Figure 8-81 Delete host port*

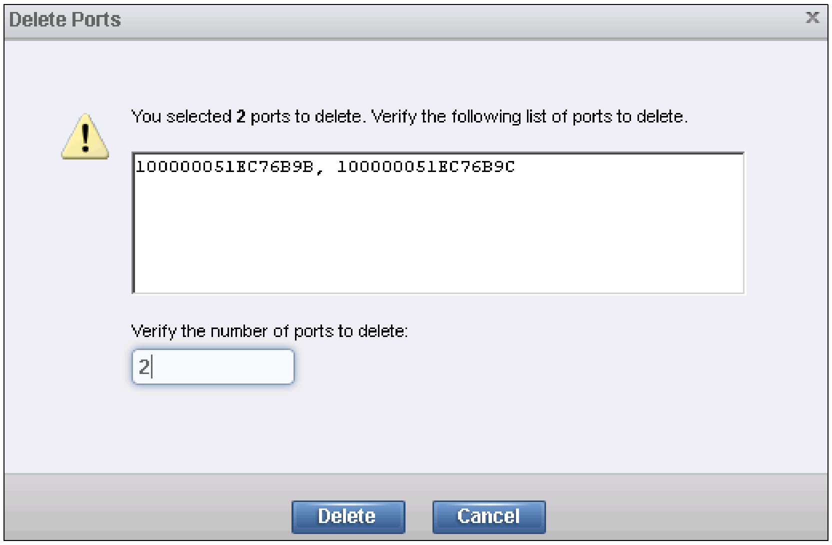
You can also press the Ctrl key while you select several host ports to delete (Figure 8-82).



| Name | Type | Status | # Nodes Logged In |
|---|---|---|---|
| 100000051EC76B91 | FC | ✓ Active | 2 |
| 100000051EC76B92 | FC | ✓ Active | 2 |
| 100000051EC76B9B | FC | ✓ Active | 2 |
| 100000051EC76B9C | FC | ✓ Active | 2 |

*Figure 8-82 Delete several host ports*

2. Click **Delete** and enter the number of host ports you want to remove (Figure 8-83).



You selected **2** ports to delete. Verify the following list of ports to delete.

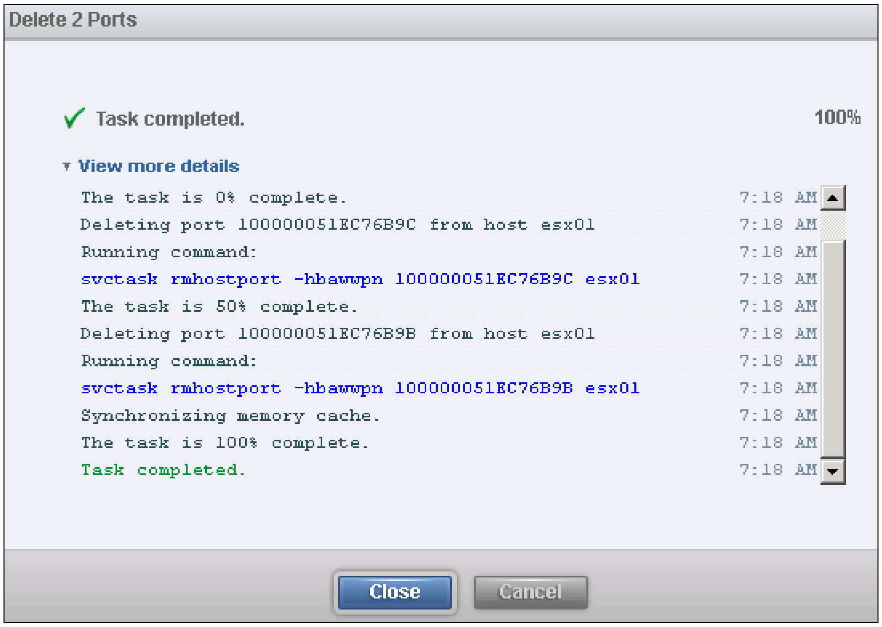
100000051EC76B9B, 100000051EC76B9C

Verify the number of ports to delete:

2

*Figure 8-83 Enter the number of host ports to delete*

3. Click **Delete** to apply the changes to the system (Figure 8-84).



✓ Task completed. 100%

▾ View more details

```
The task is 0% complete.                                    7:18 AM
Deleting port 100000051EC76B9C from host esx01              7:18 AM
Running command:                                            7:18 AM
svctask rmhostport -hbawwpn 100000051EC76B9C esx01          7:18 AM
The task is 50% complete.                                   7:18 AM
Deleting port 100000051EC76B9B from host esx01              7:18 AM
Running command:                                            7:18 AM
svctask rmhostport -hbawwpn 100000051EC76B9B esx01          7:18 AM
Synchronizing memory cache.                                 7:18 AM
The task is 100% complete.                                  7:18 AM
Task completed.                                             7:18 AM
```

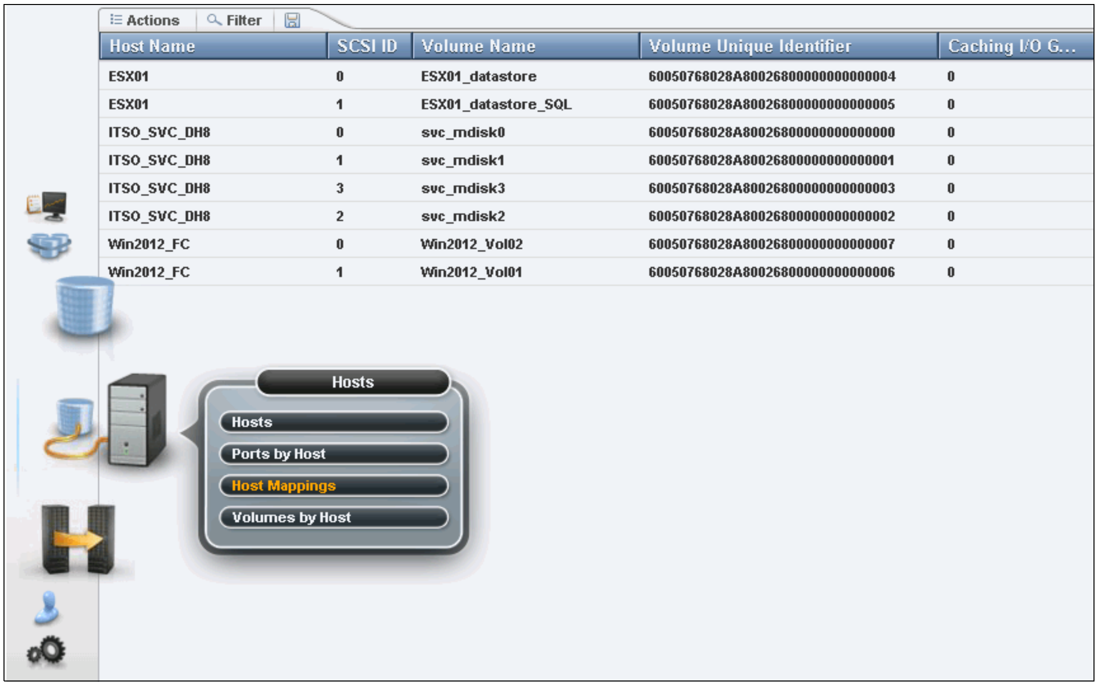
*Figure 8-84 Delete Host Ports task completion window*

4. Click **Close** to return to the Host by Ports window.

> **Note:** Deleting FC and iSCSI ports is done the same way.

## 8.4.4 Host mappings overview

Hover the mouse cursor over the host menu and select **Host Mappings** (Figure 8-25 on page 305) to open the host mappings pane (Figure 8-85).



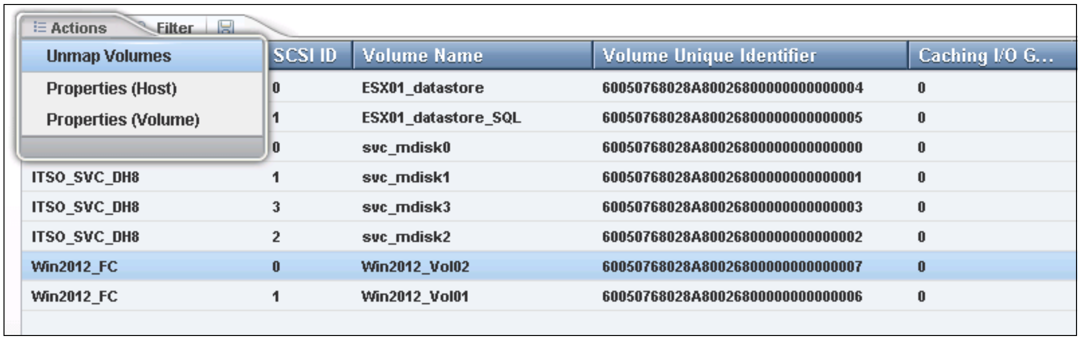
| Host Name | SCSI ID | Volume Name | Volume Unique Identifier | Caching I/O G... |
|---|---|---|---|---|
| ESX01 | 0 | ESX01_datastore | 60050768028A80026800000000000004 | 0 |
| ESX01 | 1 | ESX01_datastore_SQL | 60050768028A80026800000000000005 | 0 |
| ITSO_SVC_DH8 | 0 | svc_mdisk0 | 60050768028A80026800000000000000 | 0 |
| ITSO_SVC_DH8 | 1 | svc_mdisk1 | 60050768028A80026800000000000001 | 0 |
| ITSO_SVC_DH8 | 3 | svc_mdisk3 | 60050768028A80026800000000000003 | 0 |
| ITSO_SVC_DH8 | 2 | svc_mdisk2 | 60050768028A80026800000000000002 | 0 |
| Win2012_FC | 0 | Win2012_Vol02 | 60050768028A80026800000000000007 | 0 |
| Win2012_FC | 1 | Win2012_Vol01 | 60050768028A80026800000000000006 | 0 |

*Figure 8-85 Host mappings*

This pane lists all hosts and volumes. Our example shows that the host `Win2012_FC` has two mapped volumes, and their associated SCSI ID, Volume Name, and Volume Unique Identifier (UID). If you have more than one caching I/O group, you also see which volume is handled by which I/O group.

If you select one line and click **Actions** (Figure 8-86), the following tasks are available:

- Unmap Volumes
- Properties (Host)
- Properties (Volume)

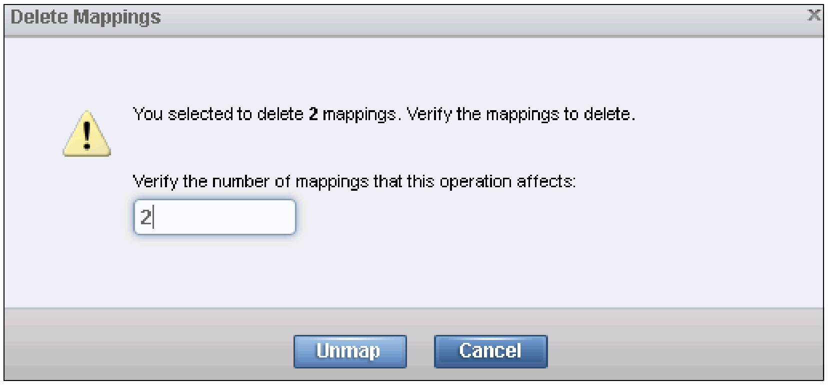


| | SCSI ID | Volume Name | Volume Unique Identifier | Caching I/O G... |
|---|---|---|---|---|
| | 0 | ESX01_datastore | 60050768028A80026800000000000004 | 0 |
| | 1 | ESX01_datastore_SQL | 60050768028A80026800000000000005 | 0 |
| | 0 | svc_mdisk0 | 60050768028A80026800000000000000 | 0 |
| ITSO_SVC_DH8 | 1 | svc_mdisk1 | 60050768028A80026800000000000001 | 0 |
| ITSO_SVC_DH8 | 3 | svc_mdisk3 | 60050768028A80026800000000000003 | 0 |
| ITSO_SVC_DH8 | 2 | svc_mdisk2 | 60050768028A80026800000000000002 | 0 |
| Win2012_FC | 0 | Win2012_Vol02 | 60050768028A80026800000000000007 | 0 |
| Win2012_FC | 1 | Win2012_Vol01 | 60050768028A80026800000000000006 | 0 |

*Figure 8-86 Host Mappings Actions menu*

## Unmapping a volume

Select one or more lines with the Ctrl key, click **Unmap Volumes**, enter the number of entries to remove (Figure 8-87), and then click **Unmap**.



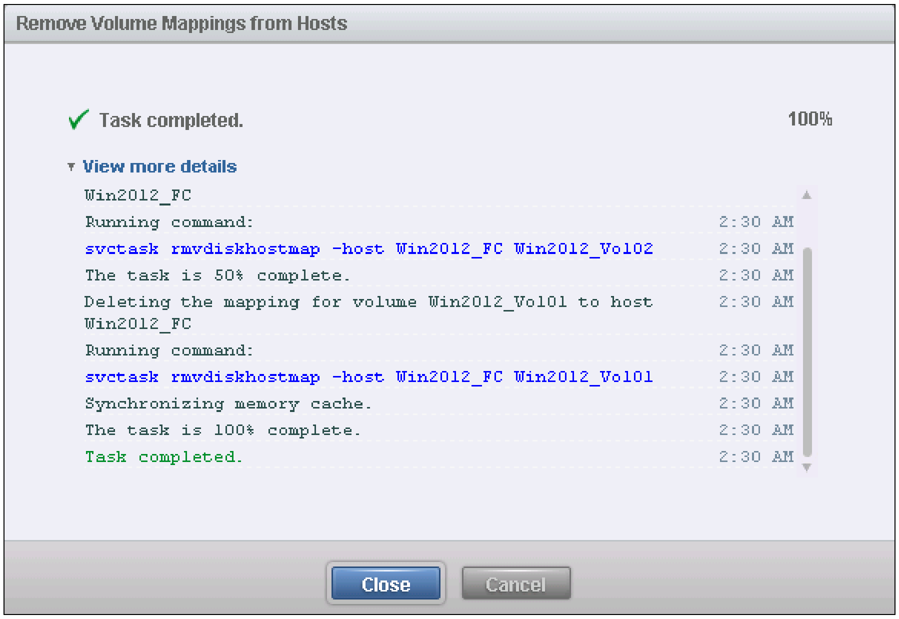
You selected to delete **2** mappings. Verify the mappings to delete.

Verify the number of mappings that this operation affects:

2

*Figure 8-87 Unmap selected volumes*

This action removes the mappings for all selected entries (Figure 8-88).



```
Win2012_FC
Running command:                                              2:30 AM
svctask rmvdiskhostmap -host Win2012_FC Win2012_Vol02         2:30 AM
The task is 50% complete.                                     2:30 AM
Deleting the mapping for volume Win2012_Vol01 to host         2:30 AM
Win2012_FC
Running command:                                              2:30 AM
svctask rmvdiskhostmap -host Win2012_FC Win2012_Vol01         2:30 AM
Synchronizing memory cache.                                   2:30 AM
The task is 100% complete.                                    2:30 AM
Task completed.                                               2:30 AM
```

*Figure 8-88 Remove Volume Mappings from Hosts task completion window*

### Properties (Host)

Selecting an entry and clicking **Properties (Host)**, as shown in Figure 8-86 on page 335, opens the host properties window. The contents of this window are described in , “Host properties” on page 322.

### Properties (Volume)

Selecting an entry and clicking **Properties (Volume)**, as shown in Figure 8-86 on page 335, opens the volume properties view. The contents of this window are described in Chapter 7, “Volumes” on page 213.

# 8.5 Hosts operations using CLI

This section describes some of the host related actions that can be taken within the Storwize V7000 system using the command line interface.

## 8.5.1 Create a host using CLI

This section describes how to create Fibre Channel and iSCSI hosts using the IBM Spectrum Virtualize CLI. It is assumed that hosts are prepared for attachment, as described in the host attachment section of IBM Storwize V7000 V7.8 knowledge center at:

`https://ibm.biz/BdsKhC`

### Creating Fibre Channel hosts

To create a Fibre Channel host, complete the described steps:

1. Rescan the SAN on Storwize V7000 using **`detectmdisk`** `command` as shown in Example 8-13.

*Example 8-13 Rescanning the SAN*

```
IBM_Storwize:ITSO:superuser>detectmdisk
```

Provided that the zoning has been implemented appropriately, the new WWPNs should be visible to the Storwize V7000 system after running the **detectmdisk** command.

2. List the candidate WWPNs and identify the WWPNs belonging to the new host as shown in Example 8-14.

*Example 8-14 Available WWPNs*

```
IBM_Storwize:ITSO:superuser>lshbaportcandidate
id
2101001B32BA36B4
2100001B329A36B4
```

3. Execute the **mkhost** command with the required parameters as shown in Example 8-15.

*Example 8-15 Host creation*

```
IBM_Storwize:ITSO:superuser>mkhost -name RHEL_HOST -fcwwpn
2100001B329A36B4:2101001B32BA36B4
Host, id [7], successfully created
```

## Creating iSCSI hosts

Prior to creating an iSCSI host in Storwize V7000, the iSCSI Qualified Name (IQN) address of the host needs to be known. Refer to the host operating system specific documentation to derive the IQN of the host.

Create the iSCSI host using the **mkhost** command as in Example 8-16.

*Example 8-16 Creating an iSCSI host using mkhost*

```
IBM_Storwize:ITSO:superuser>mkhost -iscsiname iqn.1994-05.com.redhat:e6dd277b58 -name
iSCSI_RHEL_HOST
Host, id [8], successfully created
```

The iSCSI host can be verified using the **lshost** command as shown in Example 8-17.

*Example 8-17 Verifying iSCSI host via lshost command*

```
IBM_Storwize:ITSO:superuser>lshost 8
id 8
name iSCSI_RHEL_HOST
port_count 1
type generic
mask 1111111111111111111111111111111111111111111111111111111111111111
iogrp_count 4
status offline
site_id
site_name
host_cluster_id
host_cluster_name
iscsi_name iqn.1994-05.com.redhat:e6dd277b58
node_logged_in_count 0
state offline
```

**Note:** When the host is initially configured, the default authentication method is set to no authentication and no Challenge Handshake Authentication Protocol (CHAP) secret is set. To set a CHAP secret for authenticating the iSCSI host with the Storwize V7000 system, use the **chhost** command with the **chapsecret** parameter. If you must display a CHAP secret for a defined server, use the **lsiscsiauth** command. The **lsiscsiauth** command lists the CHAP secret that is configured for authenticating an entity to the Storwize V7000 system.

**Note:** FC hosts and iSCSI hosts are handled in the same way operationally after they are created.

## 8.5.2 Advanced host administration using CLI

This section describes following advanced host operations that can be carried out using the CLI

- Mapping a volume to a host
- Mapping a volume already mapped to a different host
- Unmapping a volume from a host
- Renaming a host
- Host properties

### Mapping a volume to a host

To map an existing volume to a host the **mkvdiskhostmap** command can be issued as shown in Example 8-18.

*Example 8-18 Mapping a volume*

```
IBM_Storwize:ITSO:superuser>mkvdiskhostmap -host RHEL_HOST -scsi 0 RHEL_VOLUME
Virtual Disk to Host map, id [0], successfully created
```

The volume mapping can then be checked by issuing the **lshostvdiskmap** command against that particular host as shown in Example 8-19.

*Example 8-19 Checking the mapped volume*

```
IBM_Storwize:ITSO:superuser>lshostvdiskmap RHEL_HOST
id name      SCSI_id vdisk_id vdisk_name  vdisk_UID                        IO_group_id
IO_group_name mapping_type host_cluster_id host_cluster_name
7  RHEL_HOST 0       109      RHEL_VOLUME 600507680C81825B0000000000000154 0
io_grp0       private
```

### Mapping a volume already mapped to a different host

To map a volume to another host that has already been mapped to one host, issue **mkvdiskhost -force** as shown in Example 8-20.

*Example 8-20 Mapping the same volume to a second host*

```
IBM_Storwize:ITSO:superuser>mkvdiskhostmap -force -host iSCSI_RHEL_HOST -scsi 0 RHEL_VOLUME
Virtual Disk to Host map, id [0], successfully created
```

**Note:** The volume RHEL_VOLUME is mapped to both the hosts using the same SCSI ID. Typically that is the requirement for most host based clustering software, such as Microsoft Clustering Service (MSCS), IBM PowerHA and so on.

Subsequently, the volume “RHEL_VOLUME” is mapped to two hosts, namely, RHEL_HOST and iSCSI_RHEL_HOST and can be seen by issuing **lsvdiskhostmap** as shown in Example 8-21.

*Example 8-21 Ensuring the same volume mapped to multiple hosts*

```
IBM_Storwize:ITSO:superuser>lsvdiskhostmap RHEL_VOLUME
id  name        SCSI_id host_id host_name       vdisk_UID
IO_group_id IO_group_name mapping_type host_cluster_id host_cluster_name
109 RHEL_VOLUME 0       7      RHEL_HOST       600507680C81825B0000000000000154 0
io_grp0       private
109 RHEL_VOLUME 0       8      iSCSI_RHEL_HOST 600507680C81825B0000000000000154 0
io_grp0       private
```

## Unmapping a volume from a host

To unmap a volume from the host **rmvdiskhostmap** is used as shown in Example 8-22 on page 339.

*Example 8-22 Unmapping a volume from a host*

```
IBM_Storwize:ITSO:superuser>rmvdiskhostmap -host iSCSI_RHEL_HOST RHEL_VOLUME
```

**Note:** Prior to unmapping a volume from a host on Storwize V7000, ensure that the host side action is completed on that volume using the respective host operating system platform commands, such as, unmounting the file system, removing the volume and/or volume group, otherwise it could potentially result in data corruption.

## Renaming a host

To rename an existing host definition issue **chhost -name** as shown in Example 8-23.

*Example 8-23 Rename a host*

```
IBM_Storwize:ITSO:superuser>chhost -name FC_RHEL_HOST RHEL_HOST
```

In Example 8-23, the host “RHEL_HOST” has now been renamed to “FC_RHEL_HOST”.

## Removing a host

To remove a host from the Storwize V7000 the **rmhost** command is used as shown in Example 8-24.

*Example 8-24 Remove a host*

```
IBM_Storwize:ITSO:superuser>rmhost iSCSI_RHEL_HOST
```

**Note:** Prior to removing a host from Storwize V7000, ensure that all the volumes are unmapped from that host as described in Example 8-22.

## Host properties

To get more details on a particular host, **lshost** command can be used with the hostname or host id as a parameter as shown in Example 8-25

*Example 8-25 Host details*

```
IBM_Storwize:ITSO:superuser>lshost FC_RHEL_HOST
id 7
name FC_RHEL_HOST
port_count 2
type generic
mask 1111111111111111111111111111111111111111111111111111111111111111
iogrp_count 4
status online
site_id
site_name
host_cluster_id
host_cluster_name
WWPN 2101001B32BA36B4
node_logged_in_count 2
state inactive
WWPN 2100001B329A36B4
node_logged_in_count 2
state inactive
```

## 8.5.3 Adding and deleting a host port using CLI

This section describes adding and deleting a host port to/from Storwize V7000.

### Adding ports to a defined host

If an HBA is added or a network interface controller (NIC) is to be added to a server that is defined within the Storwize V7000, the command **addhostport** can be used to add the new port definitions to the host configuration.

If the host is connected through SAN with FC, and if the WWPN is zoned to the Storwize V7000 system, issue the **lshbaportcandidate** command to compare with the information is available from the server administrator, as shown in Example 8-26

*Example 8-26 Listing new available WWPN*

```
IBM_Storwize:ITSO:superuser>lshbaportcandidate
id
210000E08B054CAA
```

Use host or SAN switch utilities to verify whether the WWPN matches the information for the new WWPN. If the WWPN matches, use the **addhostport** command to add the port to the host as shown in Example 8-27

*Example 8-27 Adding the newly discovered WWPN to the host definition*

```
IBM_Storwize:ITSO:superuser>addhostport -hbawwpn 210000E08B054CAA FC_RHEL_HOST
```

This command adds the WWPN of `210000E08B054CAA` to the host FC_RHEL_HOST.

> **Note:** Multiple ports can be added at one time by using the separator or colon (:) between WWPNs, as shown in the following example:
>
> **`addhostport -hbawwpn 210000E08B054CAA:210000E08B89C1CD FC_RHEL_HOST`**

If the new HBA is not connected or zoned, the **lshbaportcandidate** command does not display your WWPN. In this case, you can manually enter the WWPN of your HBA or HBAs and use the **-force** flag to create the host, as shown in Example 8-28.

*Example 8-28 Adding a WWPN to the host definition using -force option*

```
IBM_Storwize:ITSO:superuser>addhostport -hbawwpn 210000E08B054CAA -force FC_RHEL_HOST
```

This command forces the addition of the WWPN that is named 210000E08B054CAA to the host called FC_RHEL_HOST.

**Note:** WWPNs are not case-sensitive within the CLI.

The host port count can be verified by running the **lshost** command again. The host FC_RHEL_HOST has an updated port count of 3, as shown in Example 8-29 on page 341.

*Example 8-29 Host with updated port count*

```
IBM_Storwize:ITSO:superuser>lshost
id name                      port_count iogrp_count status   site_id site_name
host_cluster_id host_cluster_name
5  ARCHX513HT6_RHEV_H_HOST_1 2          4           online
6  ARCHX513HT7_RHEV_H_HOST_2 2          4           online
7  FC_RHEL_HOST              3          4           online
11 ZJ_ARCSUN42KD0603         1          4           online
12 ZJ_ARCSUN42KD0629         1          4           online
13 lxia_YF                   1          4           online
17 ZJ_ARCX36506V8XM          2          4           online
18 ZJ_ARCX36506V8YV          2          4           online
```

If the host uses iSCSI as a connection method, the new iSCSI IQN ID should be used to add the port. Unlike FC-attached hosts, with iSCSI, available candidate ports cannot be checked.

After getting the other iSCSI IQN issue the **addhostport** command as shown in Example 8-30.

*Example 8-30 Adding an iSCSI port to the defined host*

```
IBM_Storwize:ITSO:superuser>addhostport -iscsiname iqn.1994-05.com.redhat:e6dd277b58
iSCSI_RHEL_HOST
```

### Deleting ports from a defined host

If a mistake is made while adding a port or if an HBA is removed from a server that is defined within the Storwize V7000, use the **rmhostport** command to remove WWPN definitions from an existing host.

Prior to removing the WWPN, ensure that it is the correct WWPN by issuing the **lshost** command, as shown in Example 8-31.

*Example 8-31 lshost command to check the WWPNs*

```
IBM_Storwize:ITSO:superuser>lshost FC_RHEL_HOST
id 7
name FC_RHEL_HOST
port_count 3
type generic
mask 1111111111111111111111111111111111111111111111111111111111111111
iogrp_count 4
```

```
status degraded
site_id
site_name
host_cluster_id
host_cluster_name
WWPN 210000E08B054CAA
node_logged_in_count 0
state offline
WWPN 2101001B32BA36B4
node_logged_in_count 2
state active
WWPN 2100001B329A36B4
node_logged_in_count 2
state active
```

Once the WWPN or iSCSI IQN is known that needs to be deleted, use the **rmhostport** command to delete a host port, as shown in Example 8-32.

*Example 8-32 mhostport to remove a WWPN*

```
IBM_Storwize:ITSO:superuser>rmhostport -fcwwpn 210000E08B054CAA FC_RHEL_HOST
```

Use the command to remove the iSCSI IQN as shown in Example 8-33.

*Example 8-33 Removing iSCSI port from the host*

```
IBM_Storwize:ITSO:superuser>rmhostport -iscsiname iqn.1994-05.com.redhat:e6dd277b58
iSCSI_RHEL_HOST
```

This command removes the WWPN of 210000E08B054CAA from the FC_RHEL_HOST host and the iSCSI IQN iqn.1994-05.com.redhat:e6dd277b58 from the host iSCSI_RHEL_HOST.

**Note:** Multiple ports can be removed at one time by using the separator or colon (:) between the port names, as shown in the following example:

```
rmhostport -hbawwpn 210000E08B054CAA:210000E08B892BCD Angola
```

## 8.5.4 Host clusters

Spectrum Virtualize V7.7.1 introduced the concept of host cluster. Host cluster allows a user to create a group of hosts to form a cluster, which will be treated as one entity instead of dealing with all the hosts individually in the cluster.

The host cluster is quite useful for hosts which are participating in a cluster at host operating system levels. The examples are Microsoft Clustering Server, IBM PowerHA, Red Hat Cluster Suite and such. By defining a host cluster, the user can map one or more volumes to the host cluster object. As a result the volume, or set of volumes, in turn gets assigned to each individual host object that is part of the host cluster and each of the volumes gets mapped with the same SCSI ID to all the hosts that are part of the host cluster with just one command.

A host cluster is made up of individual hosts, and volumes can also be assigned to individual hosts that make up the cluster.

Even though a host is part of host cluster, volumes can still be assigned to a particular host in a non-shared manner. A policy can be devised which may pre-assign a standard set of SCSI

IDs for volumes to be assigned to the host cluster, and another set of SCSI IDs to be used for individual assignments to hosts.

> **Note:** For example, SCSI ID 0-100 for individual host assignment, and SCSI ID above 100 can be used for host cluster. By employing such a policy, desired volumes will not be shared and others will be. For example, the boot volume of each host can be kept private, while data and application volumes can be shared.

## 8.5.5 Host Cluster Operations

This section describes following host cluster operations using the CLI:

- Creating a host cluster (**mkhostcluster**)
- Adding a member to the host cluster (**addhostclustermember**)
- Listing a host cluster (**lshostcluster**)
- Listing a host cluster members (**lshostclustermember**)
- Assigning a volume to the host cluster (**mkvolumehostclustermap**)
- Listing a host cluster for mapped volumes (**lshostclustervolumemap**)
- Unmapping a volume from the host cluster (**rmvolumehostclustermap**)
- Removing a host cluster member (**rmhostclustermember**)
- Remvoing the host cluster (**rmhostcluster**)

### Creating a host cluster

To create a host cluster, the command **mkhostcluster** can be used as shown in Example 8-34.

*Example 8-34 Create a host cluster using mkhostcluster*

```
IBM_Storwize:ITSO:superuser>mkhostcluster -name ITSO_HOST_CLUSTER
Host cluster, id [0], successfully created.
```

> **Note:** While creating the host cluster, if it is desired to inherit the volumes mapped to a particular host, then **-seedfromhost** flag option can be used. Any volume mapping that does not need to be shared can be kept private by using **-ignoreseedvolume** flag option.

### Adding a host to a host cluster

After creating a host cluster, a host or a list of hosts can be added using the **addhostclustermember** command as shown in Example 8-35.

*Example 8-35 Adding a host or hosts to host cluster*

```
IBM_Storwize:ITSO:superuser>addhostclustermember -host ITSO_HOST_1:ITSO_HOST_2
ITSO_HOST_CLUSTER
```

In Example 8-35, the hosts ITSO_HOST_1 and ITSO_HOST_2 are added as part of host cluster ITSO_HOST_CLUSTER.

### Listing the host cluster member

To list the host members that are part of a particular host cluster, **lshostclustermember** command can be used as shown in Example 8-36.

*Example 8-36 Listing host cluster members with lshostclustermember*

```
IBM_Storwize:ITSO:superuser>lshostclustermember ITSO_HOST_CLUSTER
host_id host_name   status  type    site_id site_name
```

```
4       ITSO_HOST_1 offline generic
5       ITSO_HOST_2 offline generic
```

### Map volume to a host cluster

To map a volume to a host cluster so that it automatically gets mapped to member hosts, **mkvolumehostclustermap** command can be used as shown in Example 8-37.

*Example 8-37 Map volume to host cluster*

```
IBM_Storwize:ITSO:superuser>mkvolumehostclustermap -hostcluster ITSO_HOST_CLUSTER ITSO_VD_1
Volume to Host Cluster map, id [0], successfully created
```

**Note:** When a volume is mapped to a host cluster, that volume gets mapped to all the members of the host cluster with the same SCSI_ID.

### Listing the mapped volumes to host cluster

To list the volumes mapped to a host cluster the **lshostclustervolumemap** command can be used as shown in Example 8-38.

*Example 8-38 Listing volumes mapped to a host cluster using lshostclustervolumemap*

```
IBM_Storwize:ITSO:superuser>lshostclustervolumemap ITSO_HOST_CLUSTER
id name              SCSI_id volume_id volume_name volume_UID
IO_group_id IO_group_name
0  ITSO_HOST_CLUSTER 0      86       ITSO_VD_1   60050768018786C188000000000001E1 0
io_grp0
0  ITSO_HOST_CLUSTER 1      87       ITSO_VD_2   60050768018786C188000000000001E2 0
io_grp0
0  ITSO_HOST_CLUSTER 2      88       ITSO_VD_3   60050768018786C188000000000001E3 0
io_grp0
```

**Note:** **lshostvdiskmap** command can be executed against each host that is part of host cluster, to ensure the mapping type for the shared volume is **shared** and is **private** for the non-shared volume.

### Removing a volume mapping from a host cluster

To remove a volume mapping to a host cluster use the **rmvolumehostclustermap** command can be used as shown in Example 8-39.

*Example 8-39 Removing a volume mapping*

```
IBM_Storwize:ITSO:superuser>rmvolumehostclustermap -hostcluster ITSO_HOST_CLUSTER ITSO_VD_1
```

In the Example 8-39, volume ITSO_VD_1 has been unmapped from the host cluster ITSO_HOST_CLUSTER. The current volume mapping can be checked to ensure this as shown in Example 8-38.

**Note:** To specify the host or hosts that acquire private mappings from the volume that is being removed from the host cluster, specify **-makeprivate** flag.

### Removing a host cluster member

To remove a host cluster member, **rmhostclustermember** can be used as shown in Example 8-40.

*Example 8-40 Removing a host cluster member*

```
IBM_Storwize:ITSO:superuser>rmhostclustermember -host ITSO_HOST_2 -removemappings
ITSO_HOST_CLUSTER
```

In the Example 8-40, the host ITSO_HOST_2 has been removed as member from the host cluster ITSO_HOST_CLUSTER, along with the associated volume mappings due to the **`-removemappings`** flag being specified.

### Removing a host cluster

To remove a host cluster use the **`rmhostcluster`** command can be used as shown in Example 8-41.

*Example 8-41 Removing a host cluster.*

```
IBM_Storwize:ITSO:superuser>rmhostcluster -removemappings ITSO_HOST_CLUSTER
```

The **`-removemappings`** flag also causes the system to remove any host mappings to volumes that are shared. The mappings are deleted before the host cluster is deleted.

> **Note:** If it is desired to keep the volumes mapped to the host objects even after the host cluster is deleted, then specify the **`-keepmappings`** flag instead of **`-removemappings`** for the **`rmhostcluster`** command. When **`-keepmappings`** is specified, the host cluster is deleted but the volume mapping to the host becomes **`private`** instead of **`shared`**.

# 9

# Storage migration

This chapter describes the steps involved in migrating data from an existing external storage system to the capacity of the IBM Storwize V7000 using the storage migration wizard. Migrating data from other storage systems to the Storwize V7000 consolidates storage and allows for IBM Spectrum Virtualize features, such as Easy Tier, thin provisioning, compression, encryption, storage replication, and the easy-to-use graphical user interface to be realized across all volumes. Storage migration leverages the volume mirroring functionality to allow read and writes during the migration, minimizing disruption and downtime. After the migration is complete, the existing system can be retired. Storwize V7000 supports migration through Fibre Channel and Internet Small Computer Systems Interface (iSCSI) connections. Storage migration can be used to migrate data from other Storwize systems.

This chapter includes the following topics:

- Storage migration overview
- Storage migration wizard

**Note:** This chapter does not cover migration outside of the storage migration wizard. To migrate data outside of the wizard you must use **Import**. For information about the Import action see Chapter 6, “Storage pools” on page 165.

# 9.1 Storage migration overview

To migrate data from an existing storage system to the IBM Storwize V7000, it is necessary to use the built-in external virtualization capability of the Storwize V7000. This capability places external connected Logical Units (LUs) under the control of the Storwize V7000. After volumes are virtualized, hosts continue to access them but do so through the Storwize V7000, which acts as a proxy.

> **Attention:** The external virtualization capability requires an external virtualization license for each enclosure being virtualized. However, data can be migrated from existing systems within 45 days of purchase of the new system without purchasing a license.
>
> Set the license temporarily during the migration process to prevent messages indicating that you are in violation of the license agreement from being sent. When the migration is complete, or after 45 days, either reset the license to its original limit or purchase a new license.

The following topics give an overview of the storage migration process:

- Typically, storage systems divide storage into many Small Computer System Interface LUs that are presented to hosts.
- I/O to the LUs must be stopped and changes made to the mapping of the storage system LUs and to the fabric configuration so that the original LUs are presented directly to the Storwize V7000 and not to the hosts. The system discovers the external LUs as *unmanaged* MDisks.
- The unmanaged MDisks are *imported* to the Storwize V7000 as *image-mode volumes* and placed into a storage pool. This storage pool is now a logical container for the LUs.
- Each MDisk has a one-to-one mapping with an image-mode volume. From a data perspective, the image-mode volumes represent the LUs exactly as they were before the import operation. The image-mode volumes are on the same physical drives of the external storage system and the data remains unchanged. The Storwize V7000 is presenting active images of the LUs and is acting as a proxy.
- The hosts must have the existing storage system multipath device driver removed and are then configured for Storwize V7000 attachment. The Storwize V7000 hosts are defined with Worldwide Port Names (WWPNs) or iSCSI Qualified Names (IQNs) and the volumes are mapped to the hosts. After the volumes are mapped, the hosts discover the Storwize V7000 volumes through a host re-scan or reboot operation.
- Storwize V7000 volume mirror operations are then initiated. The image-mode volumes are mirrored to generic volumes. The mirrors are online migration tasks, which means a host can still access and use the volumes during the mirror synchronization process.
- After the mirror operations are complete, the image-mode volumes are removed. The other storage system LUs are now migrated and the now redundant storage can be retired or re-used elsewhere.

## 9.1.1 Interoperability and compatibility

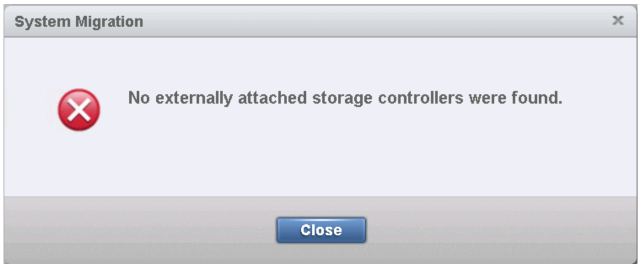
Interoperability is an important consideration when a new storage system is set up in an environment that contains existing storage infrastructure. Before attaching any external storage systems to the Storwize V7000, see the IBM Storwize V7000 and IBM Spectrum Virtualize 7.8 support matrix:

http://www.ibm.com/support/docview.wss?uid=ssg1S1009559

### 9.1.2 Prerequisites

Before the storage migration wizard can be started the external storage system must be visible to the Storwize V7000. To ensure the external system is configured properly, follow the steps mentioned in Chapter 6, “Storage pools” on page 165.

If the external storage system is not detected, the warning message shown in Figure 9-1is displayed when you attempt to start the migration wizard. Click **Close** and correct the problem before trying to start the migration wizard again.



System Migration

No externally attached storage controllers were found.

*Figure 9-1 No external storage detected error message*

## 9.2 Storage migration wizard

The storage migration wizard simplifies the migration task. The wizard features easy-to-follow panels that guide users through the entire process.

**Attention:** The risk of losing data when the storage migration wizard is used correctly is low. However, it is prudent to avoid potential data loss by creating a backup of all the data that is stored on the hosts, the existing storage systems, and the Storwize V7000 before the wizard is used.

Follow the instructions below to complete the migration using the storage migration wizard:

1. Navigate to **Pools** → **System Migration**, as shown in Figure 9-2. The System Migration panel provides access to the storage migration wizard and displays information about the migration progress.

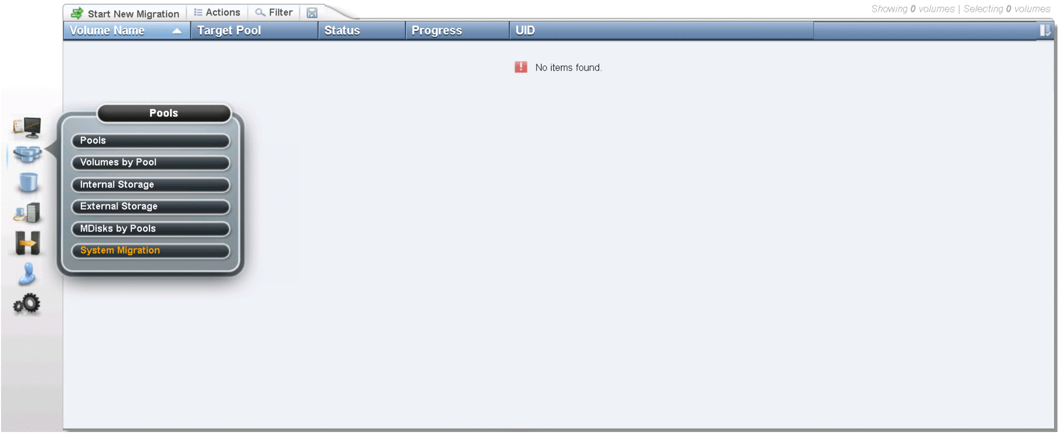


Figure 9-2 Accessing the System Migration panel

2. Click **Start New Migration** to begin the storage migration wizard, as shown in Figure 9-3.

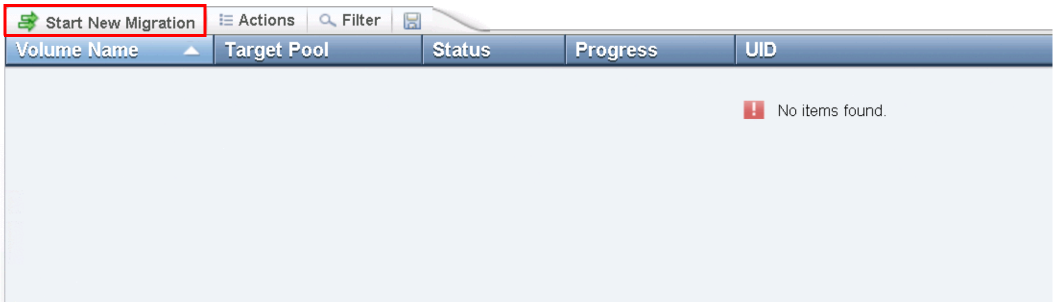
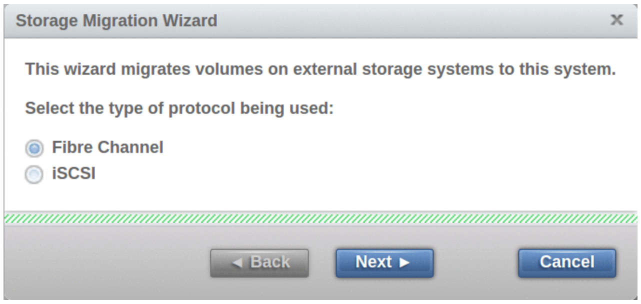


Figure 9-3 Starting a new migration

3. If both Fibre Channel and iSCSI external systems are detected, the dialog in Figure 9-4 is shown. Select the type of attachment between the Storwize V7000 and the external system from which you want to migrate volumes and click **Next**. If only one type of attachment is detected, this dialog is not displayed.



Storage Migration Wizard

This wizard migrates volumes on external storage systems to this system.

Select the type of protocol being used:

● Fibre Channel
○ iSCSI

Figure 9-4 Storage migration wizard: type of protocol

> **Note:** The steps required are the same for both protocols. There are only minimal variations on the text displayed on the wizard. This example assumes that only one type of attachment was detected. When that is the case the text displayed is generic and valid for both types.

4. When the wizard starts you are prompted to verify the restrictions and prerequisites listed in Figure 9-5 on page 352. The restrictions and prerequisites to be addressed are the following:

   – Restrictions:

     - You are not using the storage migration wizard to migrate cluster hosts, including clusters of VMware hosts and Virtual I/O Servers (VIOS).
     - You are not using the storage migration wizard to migrate SAN boot images.

     If you have either of these two environments, the migration must be performed outside of the wizard because more steps are required.

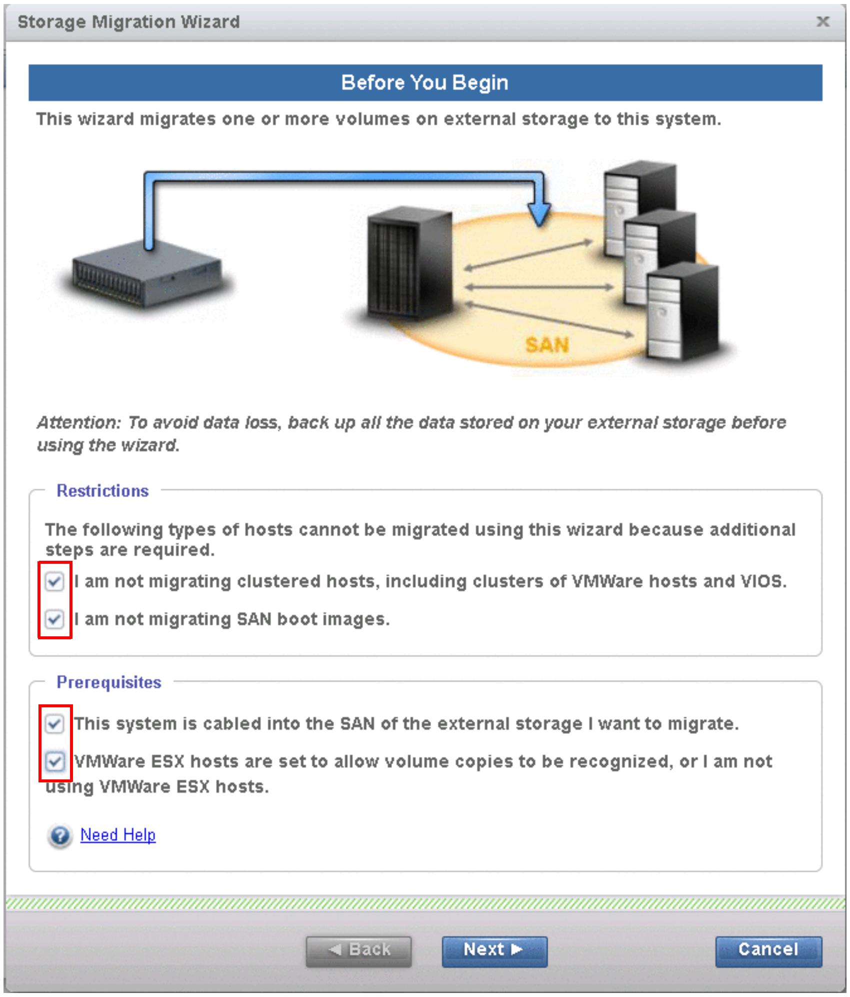
     The VMware vSphere Storage vMotion feature might be an alternative for migrating VMware clusters. For information about this topic, see this website:

     `http://www.vmware.com/products/vsphere/features/storage-vmotion.html`

   – Prerequisites:

     - The Storwize V7000 and the external storage system are connected to the same SAN fabric.
     - If there are VMware ESX hosts involved in the data migration, the VMware ESX hosts are set to allow volume copies to be recognized.

   If all restrictions are satisfied and prerequisites met, check all the boxes and click **Next**, as shown in Figure 9-5 on page 352.



Storage Migration Wizard

Before You Begin

This wizard migrates one or more volumes on external storage to this system.

*Attention: To avoid data loss, back up all the data stored on your external storage before using the wizard.*

Restrictions

The following types of hosts cannot be migrated using this wizard because additional steps are required.

☑ I am not migrating clustered hosts, including clusters of VMWare hosts and VIOS.

☑ I am not migrating SAN boot images.

Prerequisites

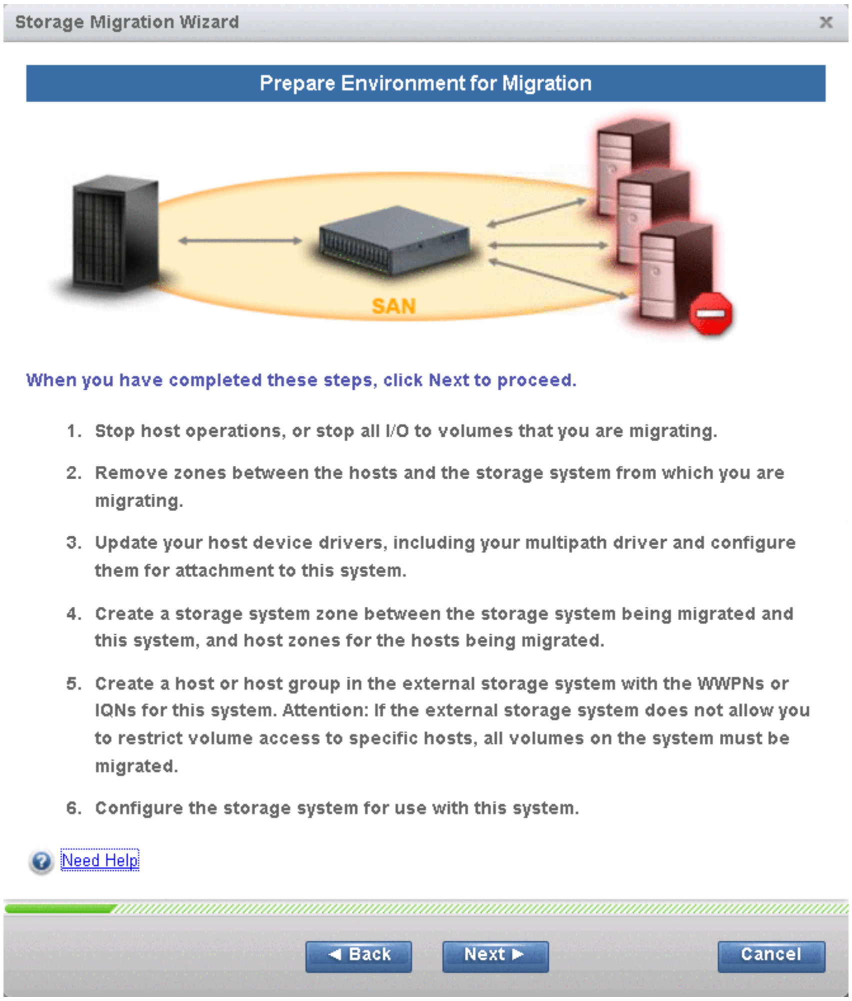
☑ This system is cabled into the SAN of the external storage I want to migrate.

☑ VMWare ESX hosts are set to allow volume copies to be recognized, or I am not using VMWare ESX hosts.

Need Help

*Figure 9-5 Storage migration wizard: confirming restrictions and prerequisites*

5. Prepare the environment migration following the on-screen instructions shown in Figure 9-6 on page 353. When all of the required tasks are complete, click **Next**.



Prepare Environment for Migration

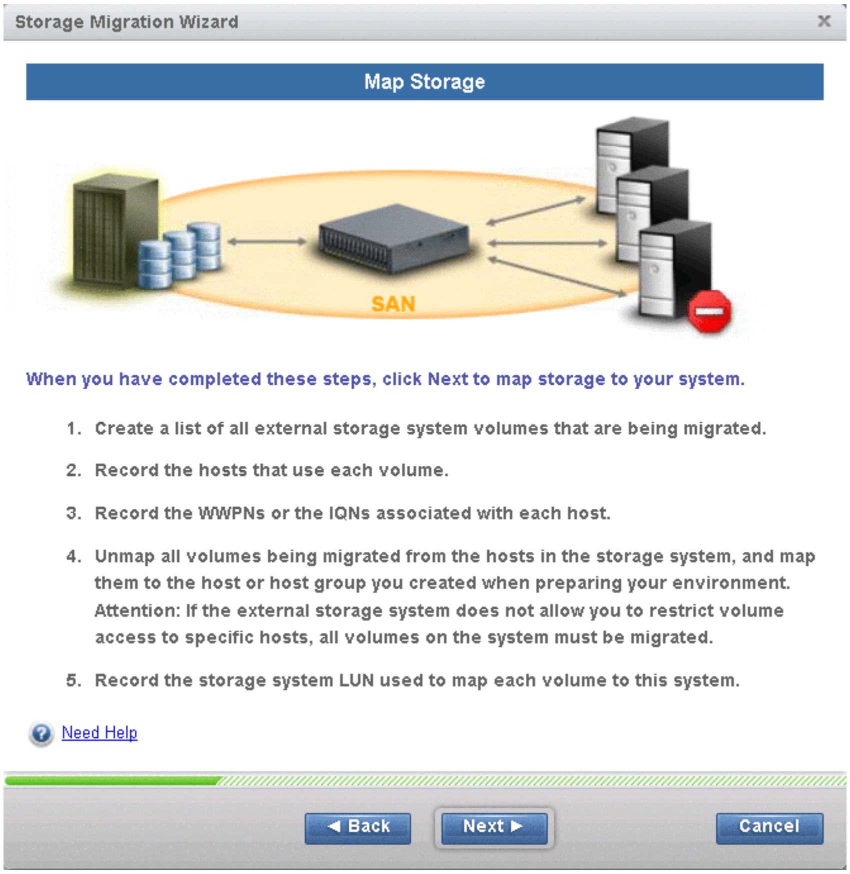
When you have completed these steps, click Next to proceed.

1. Stop host operations, or stop all I/O to volumes that you are migrating.
2. Remove zones between the hosts and the storage system from which you are migrating.
3. Update your host device drivers, including your multipath driver and configure them for attachment to this system.
4. Create a storage system zone between the storage system being migrated and this system, and host zones for the hosts being migrated.
5. Create a host or host group in the external storage system with the WWPNs or IQNs for this system. Attention: If the external storage system does not allow you to restrict volume access to specific hosts, all volumes on the system must be migrated.
6. Configure the storage system for use with this system.

Need Help

*Figure 9-6 Storage migration wizard: preparing the environment*

6. Map the external storage system following the on-screen instructions shown in Figure 9-7 on page 354.



When you have completed these steps, click Next to map storage to your system.

1. Create a list of all external storage system volumes that are being migrated.
2. Record the hosts that use each volume.
3. Record the WWPNs or the IQNs associated with each host.
4. Unmap all volumes being migrated from the hosts in the storage system, and map them to the host or host group you created when preparing your environment. Attention: If the external storage system does not allow you to restrict volume access to specific hosts, all volumes on the system must be migrated.
5. Record the storage system LUN used to map each volume to this system.

*Figure 9-7 Storage migration wizard: mapping external storage to your system*

Be sure to record the information mentioned because it is required for later steps. Table 9-1 shows an example of a table used to capture information that relates to the external storage system LUs.

*Table 9-1 Example table for capturing external LU information*

| LU name | SCSI ID | Host name | Capacity |
|---|---|---|---|
| V3700external0 | 0 | host-001 | 50 GiB |
| V3700external1 | 1 | host-001 | 50 GiB |
| V3700external2 | 0 | host-002 | 50 GiB |
| V3700external3 | 1 | host-002 | 50 GiB |

**Note:** Make sure to record the SCSI ID of the LUs to which the host is originally mapped. Some operating systems do not support changing the SCSI ID during the migration.

Table 9-2 shows an example table for capturing information that relates to a Fibre Channel host.

*Table 9-2 Example table for capturing host information*

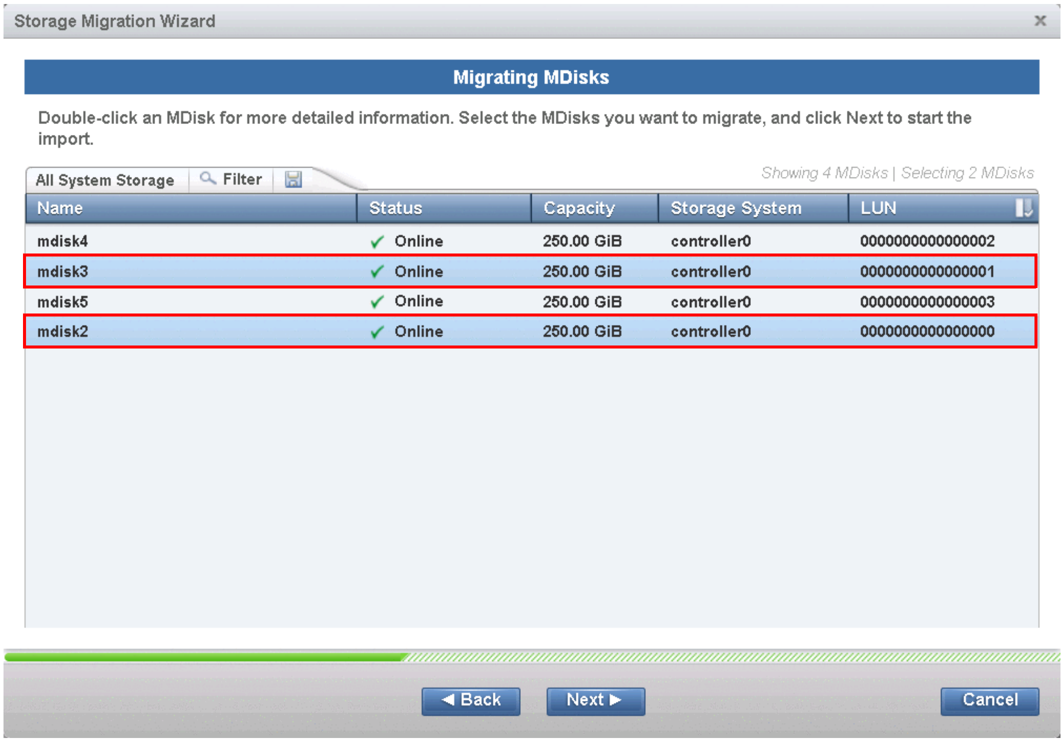
| Host Name | Adapter / Slot / Port | WWPN |
|---|---|---|
| host-001 | QLE2562 / 2 / 1 | 21000024FF2D076C |
| host-001 | QLE2562 / 2 / 2 | 21000024FF2D076D |

Click **Next** and wait for the system to discover external devices.

7. The next panel shows all the MDisks found. If the MDisks to be migrated are not in the list, check your zoning or IP configuration, as applicable, and your LU mappings. Repeat the previous step to trigger the discovery procedure again.

   Select the MDisks you want to migrate, as shown in Figure 9-8 on page 355. In this example only `mdisk2` and `mdisk3` are being migrated. Detailed information about an MDisk is visible by double-clicking it.

> **Note:** Select only the MDisks that are applicable to the current migration plan. After step 11 of the current migration completes, another migration can be started to migrate any remaining MDisks.



Storage Migration Wizard

Migrating MDisks

Double-click an MDisk for more detailed information. Select the MDisks you want to migrate, and click Next to start the import.

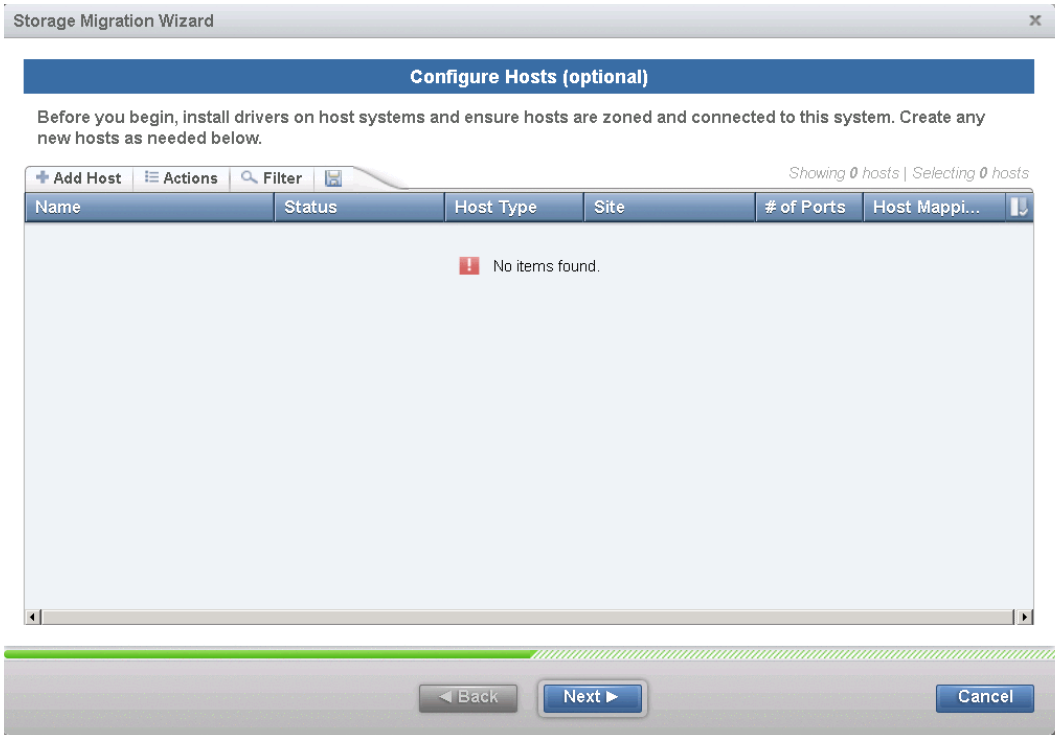
Showing 4 MDisks | Selecting 2 MDisks

| Name | Status | Capacity | Storage System | LUN |
|---|---|---|---|---|
| mdisk4 | ✓ Online | 250.00 GiB | controller0 | 0000000000000002 |
| mdisk3 | ✓ Online | 250.00 GiB | controller0 | 0000000000000001 |
| mdisk5 | ✓ Online | 250.00 GiB | controller0 | 0000000000000003 |
| mdisk2 | ✓ Online | 250.00 GiB | controller0 | 0000000000000000 |

*Figure 9-8 Storage migration wizard: selecting the MDisks to migrate*

Click **Next** and wait for the MDisks to be imported. During this task the system creates a new storage pool and adds the imported MDisks to the storage pool as image-mode volumes.

8. The next panel lists all the hosts configured on the system and allows you to configure new hosts. This step is optional and can be bypassed by clicking **Next**. In this example no hosts are configured, as shown in Figure 9-9 on page 356.



Storage Migration Wizard
Configure Hosts (optional)

Before you begin, install drivers on host systems and ensure hosts are zoned and connected to this system. Create any new hosts as needed below.

Add Host
Actions
Filter
Showing 0 hosts | Selecting 0 hosts

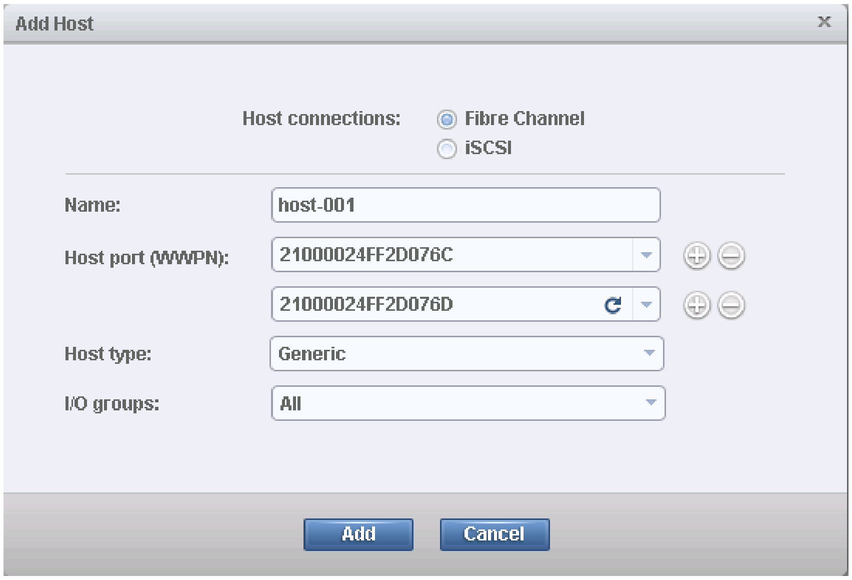
| Name | Status | Host Type | Site | # of Ports | Host Mappi... |
|---|---|---|---|---|---|

No items found.

Figure 9-9 Storage migration wizard: listing configured hosts

If the host that needs access to the migrated data is not configured, select **Add Host** to begin the Add Host wizard. Enter the host connection type, name, and connection details. Optionally, click **Advanced** to modify the host type and I/O group assignment. Figure 9-10 on page 356 shows the Add Host wizard with the details filled in.

For more information about the Add Host wizard see Chapter 8, “Hosts” on page 283.



Host connections: ● Fibre Channel ○ iSCSI

Name: host-001

Host port (WWPN): 21000024FF2D076C

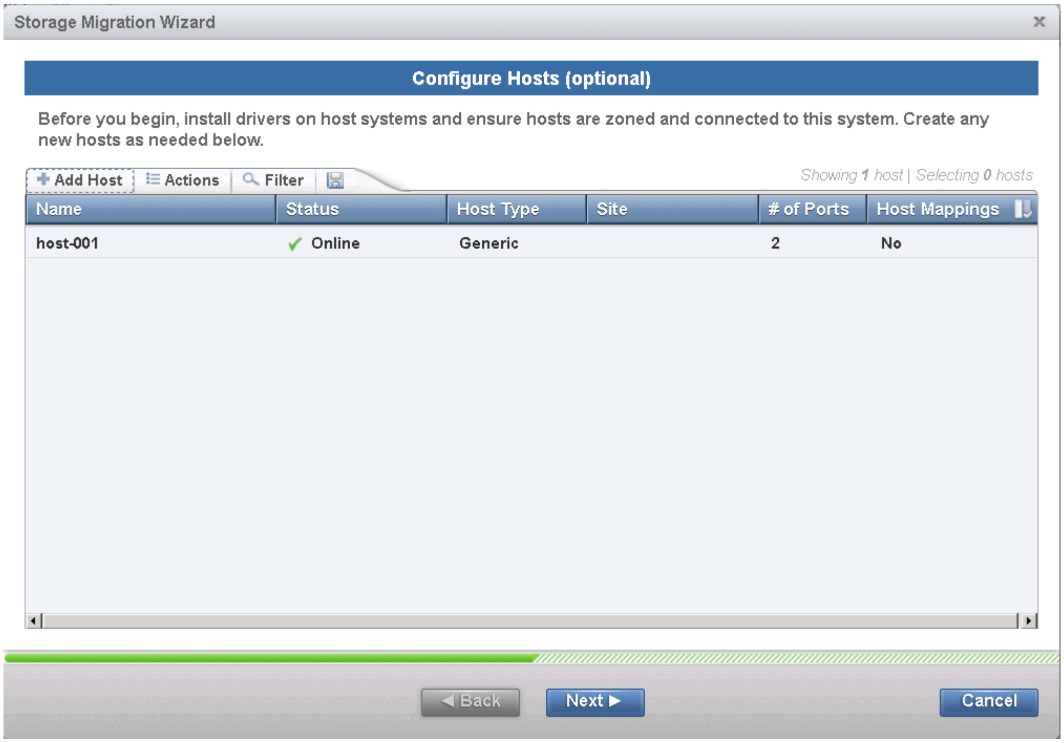
21000024FF2D076D

Host type: Generic

I/O groups: All

Figure 9-10 Storage migration wizard: adding a new host

Click **Add**. The host is created and is now listed in the Configure Hosts dialog, as shown in Figure 9-11. Click **Next** to proceed.



Before you begin, install drivers on host systems and ensure hosts are zoned and connected to this system. Create any new hosts as needed below.

| Name | Status | Host Type | Site | # of Ports | Host Mappings |
|---|---|---|---|---|---|
| host-001 | ✓ Online | Generic | | 2 | No |

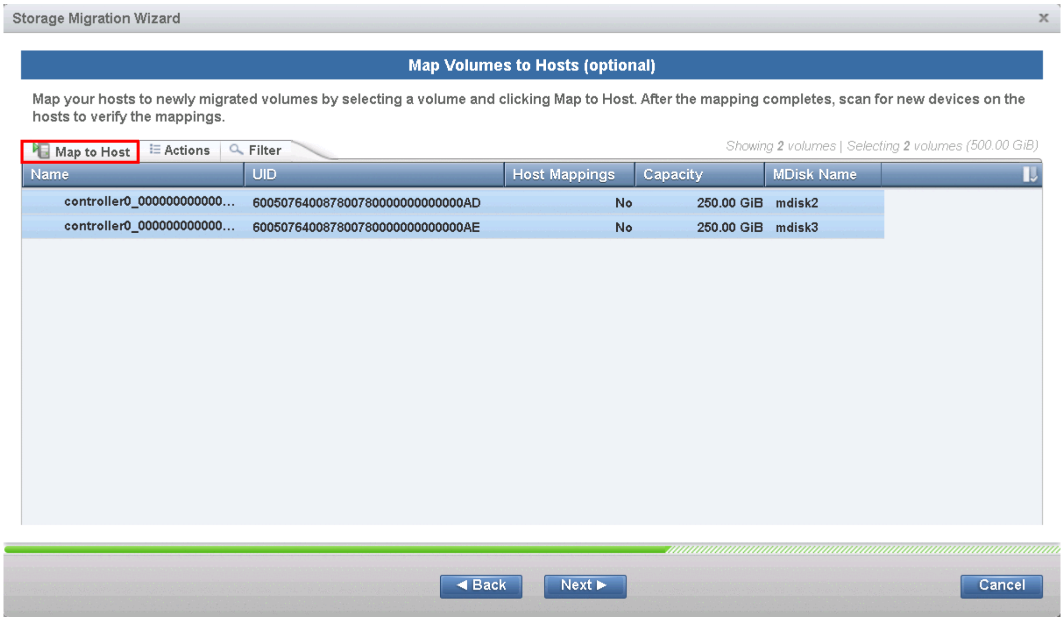
*Figure 9-11 Storage migration wizard: listing new configured host*

9. The next panel lists the new volumes and allows you to map them to hosts. The volumes are listed with names that were automatically assigned by the system. The names can be changed to reflect something more meaningful to the user by selecting the volume and clicking **Rename** in the Actions menu.

> **Note:** During the migration process a generic copy is added to the image-mode volume. This starts a mirroring operation between the two copies. Both image-mode and generic copies are presented to the hosts as the same volume and share the same volume properties. This means that the name and mapping properties set for the image-mode copy now are valid for the generic (or migrated) copy created later.

Map the volumes to hosts by selecting the volumes and clicking **Map to Host**, as shown in Figure 9-12. This step is optional and can be bypassed by clicking **Next**.

> **Note:** If your host requires the SCSI ID of a volume to remain unchanged after the migration do not use the wizard to map the volume to the host. Use the command line interface instead and run the command **`svctask mkvdiskhostmap`** with the **`-scsi`** parameter set.
>
> For example, for a volume named `controller0_0000000000000001`, a host named `host-001`, and a SCSI ID of `2`, enter the following command:
>
> ```
> svctask mkvdiskhostmap -host host-001 -scsi 2 controller0_0000000000000001
> ```



Storage Migration Wizard

Map Volumes to Hosts (optional)

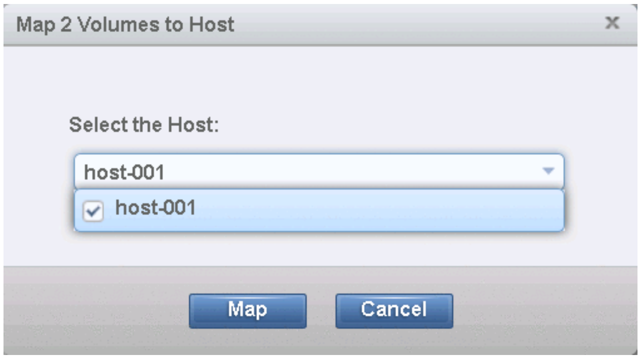
Map your hosts to newly migrated volumes by selecting a volume and clicking Map to Host. After the mapping completes, scan for new devices on the hosts to verify the mappings.

*Showing 2 volumes | Selecting 2 volumes (500.00 GiB)*

| Name | UID | Host Mappings | Capacity | MDisk Name |
|---|---|---|---|---|
| controller0_000000000000... | 600507640087800780000000000000AD | No | 250.00 GiB | mdisk2 |
| controller0_000000000000... | 600507640087800780000000000000AE | No | 250.00 GiB | mdisk3 |

*Figure 9-12 Storage migration wizard: selecting the volumes to map*

A dialog box with a drop-down list of the hosts available is displayed. Select the host you want to map the volumes to and click **Map**, as shown in Figure 9-13 on page 358.



Map 2 Volumes to Host

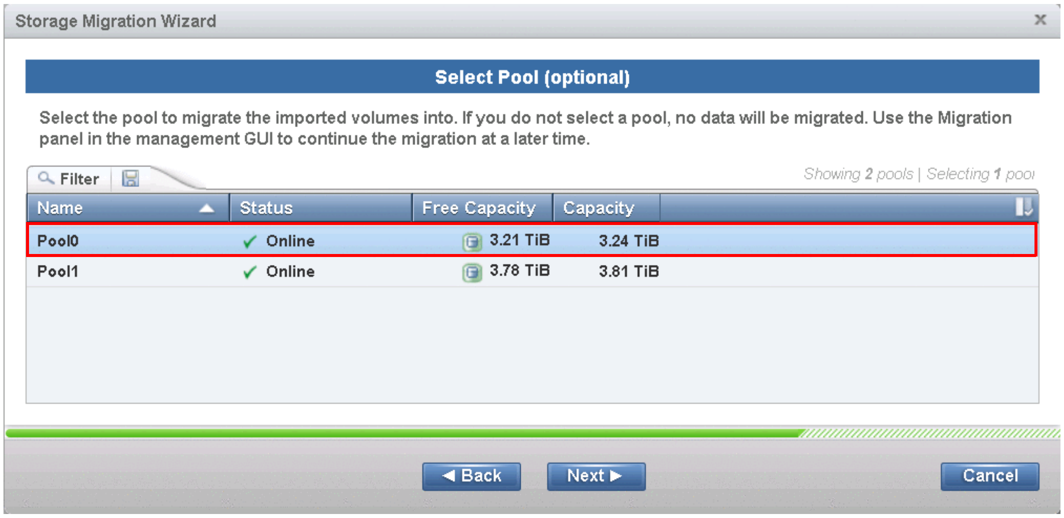
Select the Host: host-001

☑ host-001

*Figure 9-13 Storage migration wizard: selecting the host*

Wait for the mappings to be created. Click **Next**.

10. Select the storage pool you want to migrate the imported volumes into. Ensure that the selected storage pool has enough space to accommodate the migrated volumes before continuing. Click **Next**, as shown in Figure 9-14 on page 359.



Storage Migration Wizard

Select Pool (optional)

Select the pool to migrate the imported volumes into. If you do not select a pool, no data will be migrated. Use the Migration panel in the management GUI to continue the migration at a later time.

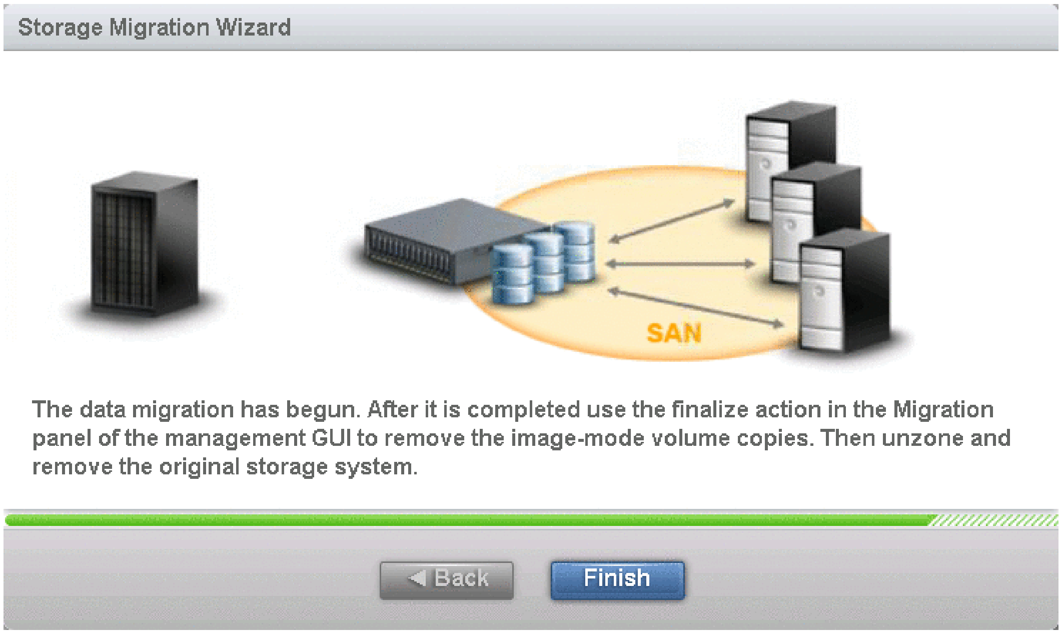
Showing 2 pools | Selecting 1 pool

| Name | Status | Free Capacity | Capacity |
|---|---|---|---|
| Pool0 | ✓ Online | 3.21 TiB | 3.24 TiB |
| Pool1 | ✓ Online | 3.78 TiB | 3.81 TiB |

Figure 9-14 Storage migration wizard: selecting a storage pool

The migration starts. This task continues running in the background and uses the volume mirroring function to place a generic copy of the image-mode volumes in the selected storage pool.

11.Click **Finish** to end the storage migration wizard, as shown in Figure 9-15.

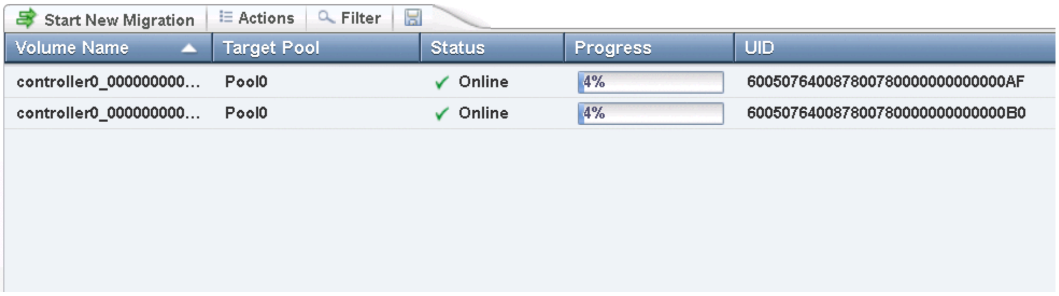


Storage Migration Wizard

The data migration has begun. After it is completed use the finalize action in the Migration panel of the management GUI to remove the image-mode volume copies. Then unzone and remove the original storage system.

Figure 9-15 Storage migration wizard: finishing the wizard

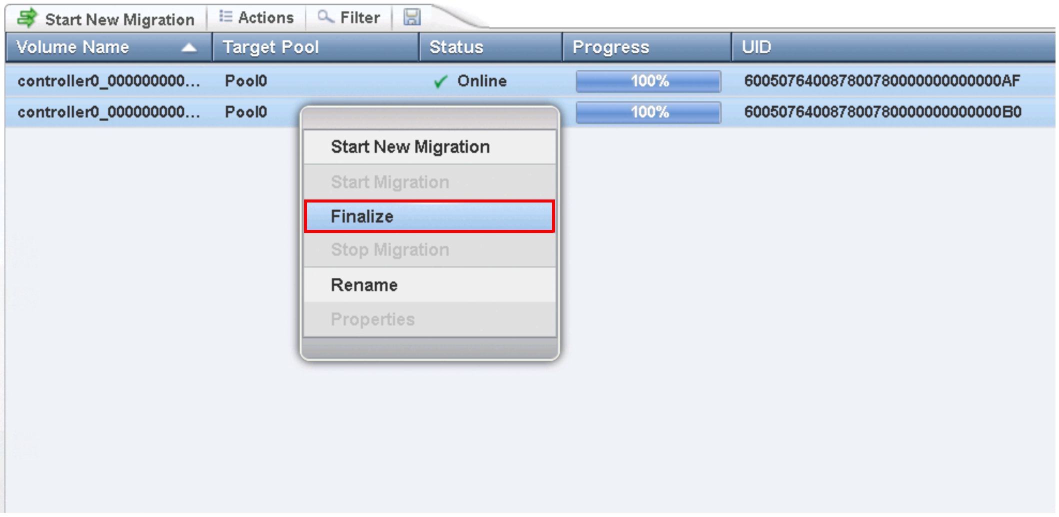
The end of the wizard is not the end of the migration task. You can find the progress of the migration in the Storage Migration panel, as shown in Figure 9-16 on page 360. The destination storage and the status of the volumes is also displayed there.



| Volume Name | Target Pool | Status | Progress | UID |
|---|---|---|---|---|
| controller0_000000000... | Pool0 | ✓ Online | 4% | 600507640087800780000000000000AF |
| controller0_000000000... | Pool0 | ✓ Online | 4% | 600507640087800780000000000000B0 |

*Figure 9-16 Migration progress*

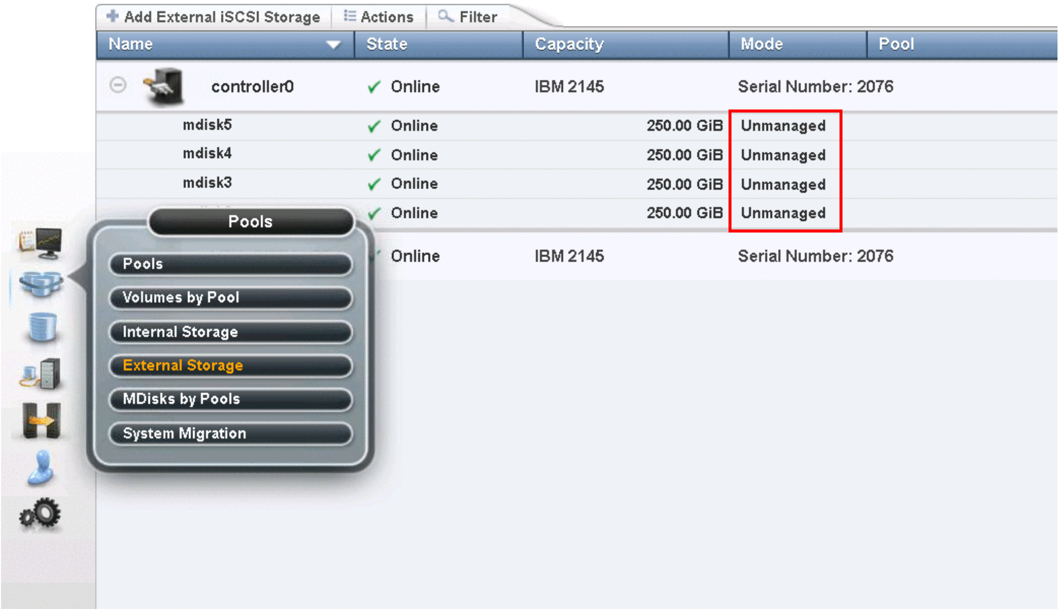
12. When the migration completes, select all the migrations you want to finalize, right-click the selection and click **Finalize**, as shown in Figure 9-17. Alternatively, select **Actions** and then **Finalize**.



| Volume Name | Target Pool | Status | Progress | UID |
|---|---|---|---|---|
| controller0_000000000... | Pool0 | ✓ Online | 100% | 600507640087800780000000000000AF |
| controller0_000000000... | Pool0 | | 100% | 600507640087800780000000000000B0 |

*Figure 9-17 Finalizing the migration*

The image-mode copies of the volumes are deleted and the associated MDisks are removed from the migration pool. The status of those MDisks returns to `unmanaged`. You can verify the status of the MDisks by navigating to **Pools** → **External Storage**, as shown in Figure 9-18.



| Name | State | Capacity | Mode | Pool |
|---|---|---|---|---|
| controller0 | ✓ Online | IBM 2145 | Serial Number: 2076 | |
| mdisk5 | ✓ Online | 250.00 GiB | Unmanaged | |
| mdisk4 | ✓ Online | 250.00 GiB | Unmanaged | |
| mdisk3 | ✓ Online | 250.00 GiB | Unmanaged | |
| | ✓ Online | 250.00 GiB | Unmanaged | |
| | Online | IBM 2145 | Serial Number: 2076 | |

*Figure 9-18 Unmanaged external MDisks*

# 10

# Advanced features for storage efficiency

In this chapter, we introduce the basic concepts of dynamic data relocation and storage optimization features. The IBM Spectrum Virtualize software running inside IBM Storwize V7000 offers IBM Easy Tier, Thin Provisioning, and IBM Real-time Compression functions for storage efficiency. We provide only a basic technical overview and benefits of each feature. For more information about planning and configuration, see the following IBM Redbooks publications:

- Easy Tier:
  - *Implementing IBM Easy Tier with IBM Real-time Compression*, TIPS1072
  - *IBM System Storage SAN Volume Controller and Storwize V7000 Best Practices and Performance Guidelines*, SG24-7521
  - *IBM DS8000 Easy Tier*, REDP-4667 (this concept is similar to IBM SAN Volume Controller Easy Tier)
- Thin Provisioning:
  - *Thin Provisioning in an IBM SAN or IP SAN Enterprise Environment*, REDP-4265
  - *DS8000 Thin Provisioning*, REDP-4554 (similar concept to IBM Storwize V7000 thin provisioning)
- Real-Time Compression:
  - *IBM Real-time Compression in IBM SAN Volume Controller and IBM Storwize V7000*, REDP-4859
  - *Implementing IBM Real-time Compression in SAN Volume Controller and IBM Storwize V7000*, TIPS1083
  - *Implementing IBM Easy Tier with IBM Real-time Compression*, TIPS1072

Specifically, this chapter provides information about the following topics:

- 10.1, “Introduction” on page 364
- 10.2, “Easy Tier” on page 365
- 10.3, “Thin provisioning” on page 392
- 10.4, “Real-time Compression software” on page 398

## 10.1 Introduction

In modern and complex application environments, the increasing and often unpredictable demands for storage capacity and performance lead to issues of planning and optimization of storage resources.

Consider the following typical storage management issues:

- Usually when a storage system is implemented, only a portion of the configurable physical capacity is deployed. When the storage system runs out of the installed capacity and more capacity is requested, a hardware upgrade is implemented to add physical resources to the storage system. This new physical capacity can hardly be configured to keep an even spread of the overall storage resources.

  Typically, the new capacity is allocated to fulfill only new storage requests. The existing storage allocations do not benefit from the new physical resources. Similarly, the new storage requests do not benefit from the existing resources; only new resources are used.

- In a complex production environment, it is not always possible to optimize storage allocation for performance. The unpredictable rate of storage growth and the fluctuations in throughput requirements, which are input/output (I/O) operations per second (IOPS), often lead to inadequate performance.

  Furthermore, the tendency to use even larger volumes to simplify storage management works against the granularity of storage allocation, and a cost-efficient storage tiering solution becomes difficult to achieve. With the introduction of high-performing technologies, such as solid-state drives (SSD) or all-flash arrays, this challenge becomes even more important.

- The move to larger and larger physical disk drive capacities means that previous access densities that were achieved with low-capacity drives can no longer be sustained.

- Any business has applications that are more critical than others, and a need exists for specific application optimization. Therefore, the ability to relocate specific application data to a faster storage media is needed.

- Although more servers are purchased with local SSDs attached for better application response time, the data distribution across these direct-attached SSDs and external storage arrays must be carefully addressed. An integrated and automated approach is crucial to achieve performance improvement without compromise to data consistency, especially in a disaster recovery (DR) situation.

All of these issues deal with data placement and relocation capabilities, or data volume reduction. Most of these challenges can be managed by having spare resources available, by moving data, and by the use of data mobility tools or operating systems features (such as host level mirroring) to optimize storage configurations.

However, all of these corrective actions are expensive in terms of hardware resources, labor, and service availability. Relocating data among the physical storage resources that dynamically or effectively reduces the amount of data, transparently to the attached host systems, is becoming increasingly important.

# 10.2 Easy Tier

In today's storage market, SSDs and flash arrays are emerging as an attractive alternative to hard disk drives (HDDs). Because of their low response times, high throughput, and IOPS-energy-efficient characteristics, SSDs and flash arrays have the potential to enable your storage infrastructure to achieve significant savings in operational costs.

However, the current acquisition cost per gibibyte (GiB) for SSDs or flash arrays is higher than for HDDs. SSD and flash array performance depends greatly on workload characteristics; therefore, they should be used with HDDs for optimal performance.

Choosing the correct mix of drives and the correct data placement is critical to achieve optimal performance at low cost. Maximum value can be derived by placing "hot" data with high I/O density and low response time requirements on SSDs or flash arrays, while targeting HDDs for "cooler" data that is accessed more sequentially and at lower rates.

Easy Tier automates the placement of data among different storage tiers, and it can be enabled for internal and external storage. This IBM Spectrum Virtualize feature boosts your storage infrastructure performance to achieve optimal performance through a software, server, and storage solution.

Additionally, the new, no-charge feature called *storage pool balancing,* introduced in V7.3, automatically moves extents within the same storage tier, from overloaded to less loaded managed disks (MDisks). Storage pool balancing ensures that your data is optimally placed among all disks within storage pools.

## 10.2.1 Easy Tier concepts

IBM Spectrum Virtualize implements Easy Tier enterprise storage functions, which were originally available on IBM DS8000 and IBM XIV enterprise class storage systems. It enables automated subvolume data placement throughout different or within the same storage tiers. This intelligently aligns the system with current workload requirements and optimizes the usage of SSDs or flash arrays.

This functionality includes the ability to automatically and non-disruptively relocate data (at the extent level) from one tier to another tier, or even within the same tier, in either direction to achieve the best available storage performance for your workload in your environment. Easy Tier reduces the I/O latency for hot spots, but it does not replace storage cache.

Both Easy Tier and storage cache solve a similar access latency workload problem, but these two methods weigh differently in the algorithmic construction that is based on *locality of reference*, recency, and frequency. Because Easy Tier monitors I/O performance from the device end (after cache), it can pick up the performance issues that cache cannot solve, and complement the overall storage system performance.

Figure 10-1 shows placement of the Easy Tier engine within the IBM Spectrum Virtualize software stack.

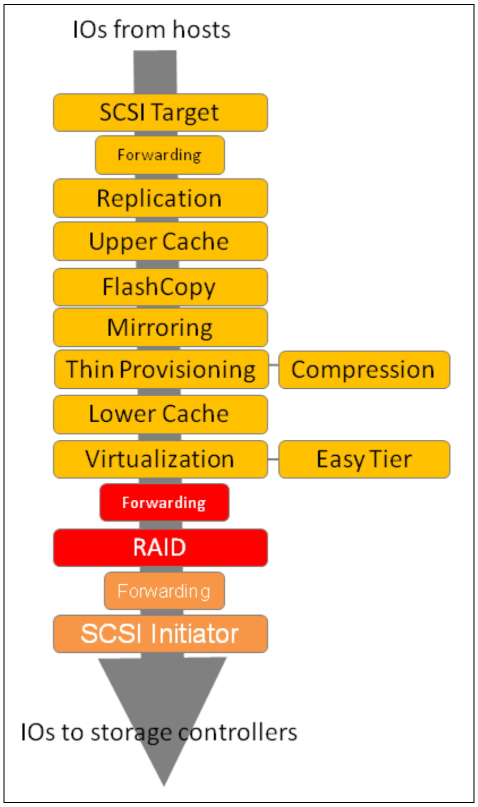


*Figure 10-1 Easy Tier in the IBM Spectrum Virtualize software stack*

In general, the storage environment's I/O is monitored at a volume level, and the entire volume is always placed inside one appropriate storage tier. Determining the amount of I/O, moving part of the underlying volume to an appropriate storage tier, and reacting to workload changes is too complex for manual operation. This is where the Easy Tier feature can be used.

Easy Tier is a performance optimization function, because it automatically migrates (or moves) extents that belong to a volume between different storage tiers (see Figure 10-2 on page 367) or the same storage tier (see Figure 10-6 on page 370). As this migration works at the extent level, it is often referred to as *sub-logical unit number (LUN) migration*. Movement of the extents is done online and is unnoticed from the host point of view. As a result of extent movement, the volume no longer has all its data in one tier, but rather in two or three tiers.

Figure 10-2 on page 367 shows the basic Easy Tier principle of operation.

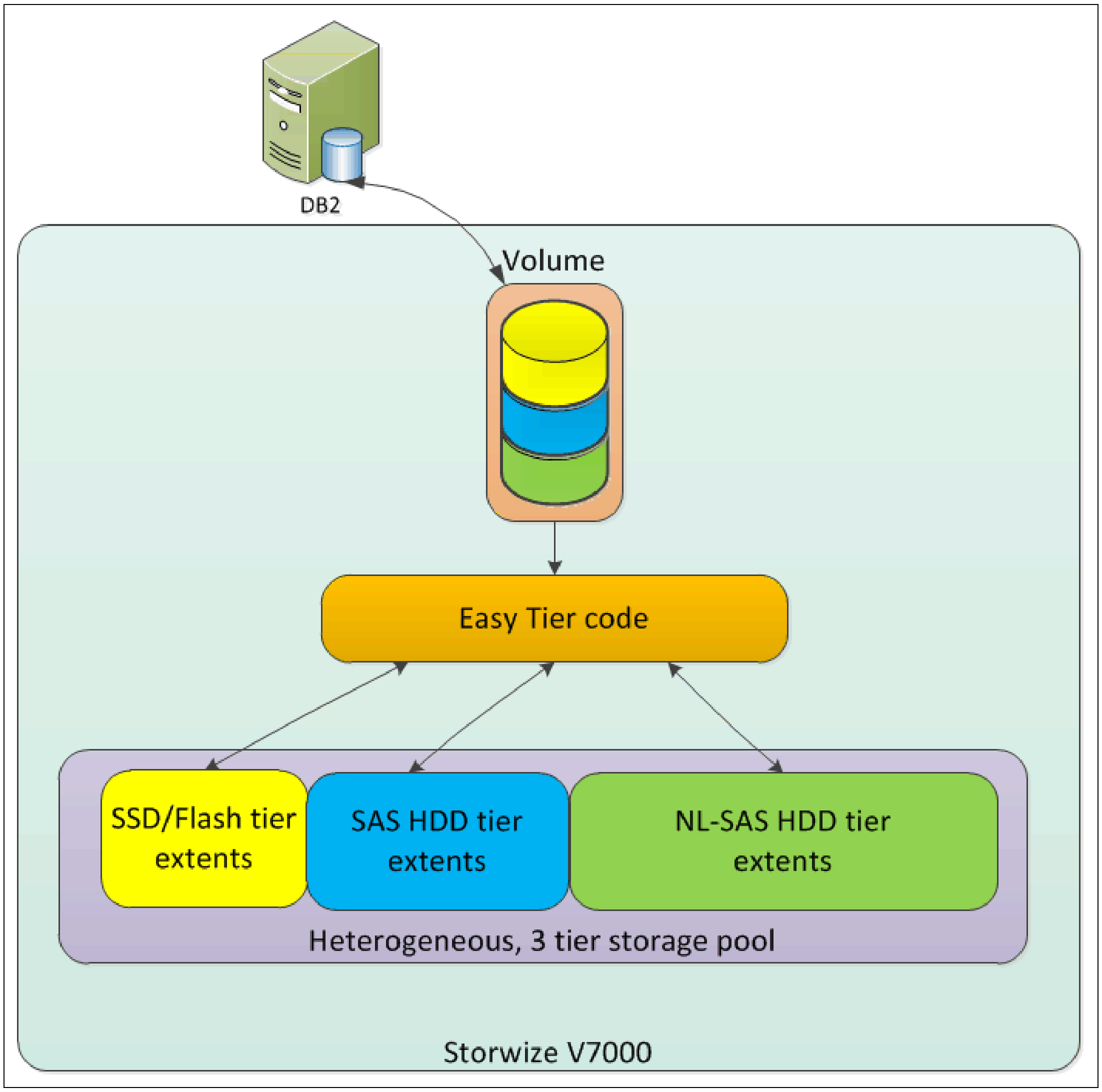


*Figure 10-2 Easy Tier*

You can enable Easy Tier on a volume basis. It monitors the I/O activity and latency of the extents on all Easy Tier enabled volumes over a 24-hour period. Based on the performance log, Easy Tier creates an extent migration plan and dynamically moves (promotes) high activity or hot extents to a higher disk tier within the same storage pool.

It also moves (demotes) extents whose activity dropped off, or cooled, from higher disk tier MDisks back to a lower tier MDisk. When Easy Tier runs in a storage pool rebalance mode, it moves extents from busy MDisks to less busy MDisks of the same type.

## 10.2.2 SSD arrays and flash MDisks

The SSDs or flash arrays are treated no differently by the IBM Storwize V7000 than normal HDDs regarding Redundant Array of Independent Disks (RAID) arrays or MDisks. The individual SSDs in the storage enclosures are combined into an array, usually in RAID 10 or RAID 5 format. It is unlikely that RAID6 SSD arrays are used, because of the double parity resource requirements, with two logical SSDs used for parity only. As with usual HDDs, RAID is an MDisk of an array type and after creation is then managed the same way the HDD MDisks are.

As is the case for HDDs, the SSD RAID array format helps to protect against individual SSD failures. Depending on your requirements, you can achieve more high availability (HA) protection above the RAID level by using volume mirroring.

The internal storage configuration of flash arrays can differ depending on an array vendor. Regardless of the methods used to configure flash-based storage, the flash system maps a volume to a host, in this case, to the Storwize V7000. From the IBM Storwize V7000 perspective, a volume presented from a flash storage is also seen as a normal managed disk.

After creation of an SSD RAID array it appears as a usual MDisk but with a tier of *flash*, which differs from MDisks presented from external storage systems or RAID arrays made of HDDs. Because Storwize does not know what kind of physical disks external MDisks are formed of, the default MDisk tier that Storwize adds to each external MDisk is *enterprise*. It is up to the user or administrator to change the tier of MDisks to *flash*, *enterprise*, or *nearline*.

> **Note:** It is possible to change the tier of MDisks made of internal Storwize drives even if the tier of MDisk does not fit the tier of physical drives that the MDisk is made of. Storwize knows the tier of drives it has in its disk enclosures, and selects the MDisk tier according to the drive tier. However, this selection can be overridden by a user or administrator. The only way to change the internal drive's tier is using the command-line interface (CLI).

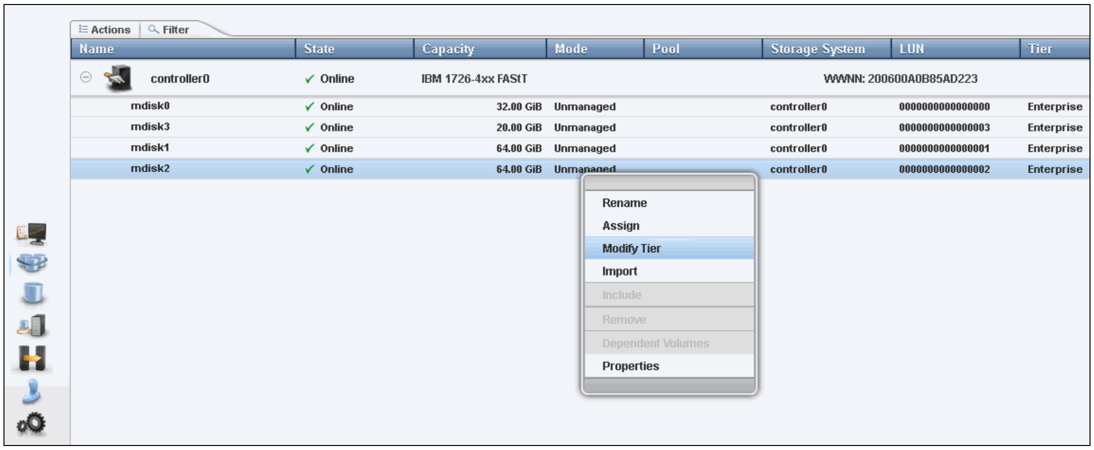
To change a tier of an MDisk in the CLI, use the **chmdisk** command as in Example 10-1.

*Example 10-1 Changing MDisk tier*

```
IBM_Storwize:V7000 Gen 1 EXTSTG2:superuser>lsmdisk -delim " "
id name status mode mdisk_grp_id mdisk_grp_name capacity ctrl_LUN_#
controller_name UID tier encrypt site_id site_name
0 MDisk_01 online array 0 test_pool_1 2.7TB    ssd no
1 MDisk_02 online array 1 test_pool_2 2.4TB    enterprise no
2 mdisk0 online unmanaged   32.0GB 0000000000000000 controller0
600a0b80005ad22300000371527a29b200000000000000000000000000000000 enterprise no
3 mdisk1 online unmanaged   64.0GB 0000000000000001 controller0
600a0b80005ad22300000372527a29cf00000000000000000000000000000000 enterprise no
4 mdisk2 online unmanaged   64.0GB 0000000000000002 controller0
600a0b80005ad22300000373527a29ea00000000000000000000000000000000 enterprise no
5 mdisk3 online unmanaged   20.0GB 0000000000000003 controller0
600a0b80005ad223000004b952d3693e00000000000000000000000000000000 enterprise no


IBM_Storwize:V7000 Gen 1 EXTSTG2:superuser>chmdisk -tier ssd mdisk3

IBM_Storwize:V7000 Gen 1 EXTSTG2:superuser>lsmdisk -delim " "
id name status mode mdisk_grp_id mdisk_grp_name capacity ctrl_LUN_#
controller_name UID tier encrypt site_id site_name
0 MDisk_01 online array 0 test_pool_1 2.7TB    ssd no
1 MDisk_02 online array 1 test_pool_2 2.4TB    enterprise no
2 mdisk0 online unmanaged   32.0GB 0000000000000000 controller0
600a0b80005ad22300000371527a29b200000000000000000000000000000000 enterprise no
3 mdisk1 online unmanaged   64.0GB 0000000000000001 controller0
600a0b80005ad22300000372527a29cf00000000000000000000000000000000 enterprise no
4 mdisk2 online unmanaged   64.0GB 0000000000000002 controller0
600a0b80005ad22300000373527a29ea00000000000000000000000000000000 enterprise no
5 mdisk3 online unmanaged   20.0GB 0000000000000003 controller0
600a0b80005ad223000004b952d3693e00000000000000000000000000000000 ssd no
```

It is also possible to change the MDisk tier from the graphical user interface (GUI), but this only applies to external MDisks. To change the tier, complete the following steps:

1. Go to **Pools** → **External Storage** and click the Plus sign (+) next to the controller that owns the MDisks for which you want to change the tier.
2. Then, right-click the wanted MDisk and select **Modify Tier** (Figure 10-3).



| Name | State | Capacity | Mode | Pool | Storage System | LUN | Tier |
|---|---|---|---|---|---|---|---|
| controller0 | ✓ Online | IBM 1726-4xx FAStT | | | WWNN: 200600A0B85AD223 | | |
| mdisk0 | ✓ Online | 32.00 GiB | Unmanaged | | controller0 | 0000000000000000 | Enterprise |
| mdisk3 | ✓ Online | 20.00 GiB | Unmanaged | | controller0 | 0000000000000003 | Enterprise |
| mdisk1 | ✓ Online | 64.00 GiB | Unmanaged | | controller0 | 0000000000000001 | Enterprise |
| mdisk2 | ✓ Online | 64.00 GiB | Unmanaged | | controller0 | 0000000000000002 | Enterprise |

Figure 10-3 Change the MDisk tier

3. The new window opens with options to change the tier (Figure 10-4).

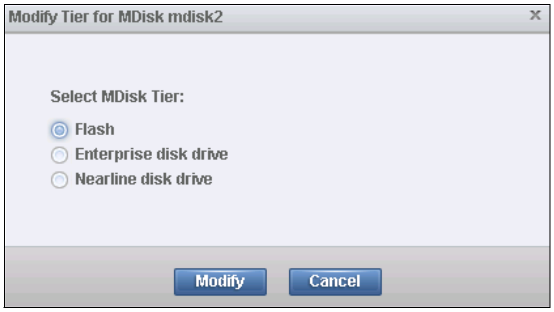
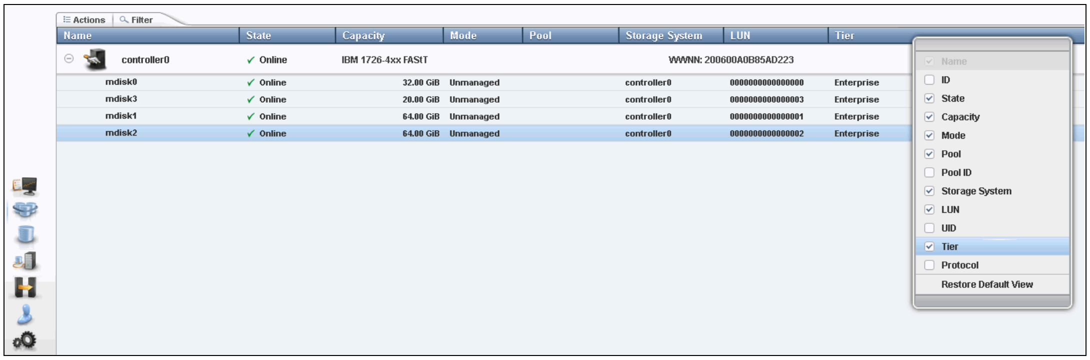


Figure 10-4 Select wanted MDisk tier

This change happens online and has no effect on hosts or availability of the volumes.

4. If you do not see the *Tier* column, right-click the blue title row and select the **Tier** check box, as shown in Figure 10-5.



| Name | State | Capacity | Mode | Pool | Storage System | LUN | Tier |
|---|---|---|---|---|---|---|---|
| controller0 | ✓ Online | IBM 1726-4xx FAStT | | | WWNN: 200600A0B85AD223 | | |
| mdisk0 | ✓ Online | 32.00 GiB | Unmanaged | | controller0 | 0000000000000000 | Enterprise |
| mdisk3 | ✓ Online | 20.00 GiB | Unmanaged | | controller0 | 0000000000000003 | Enterprise |
| mdisk1 | ✓ Online | 64.00 GiB | Unmanaged | | controller0 | 0000000000000001 | Enterprise |
| mdisk2 | ✓ Online | 64.00 GiB | Unmanaged | | controller0 | 0000000000000002 | Enterprise |

Figure 10-5 Customizing the title row to show the tier column

## 10.2.3 Disk tiers

The internal or external MDisks (LUNs) are likely to have different performance attributes because of the type of disk or RAID array on which they reside. The MDisks can be created on 15,000 revolutions per minute (RPM) Fibre Channel (FC) or serial-attached SCSI (SAS) disks, nearline SAS (NL-SAS) or Serial Advanced Technology Attachment (SATA), or even SSDs or flash storage systems.

As mentioned in 10.2.2, “SSD arrays and flash MDisks” on page 367, Storwize V7000 does not automatically detect the type of external MDisks. Instead, all external MDisks initially are put into the enterprise tier by default. Then, the administrator must manually change the tier of MDisks and add them to storage pools. Depending on what type of disks are gathered to form a storage pool, we distinguish two types of storage pools:

- Single-tier
- Multitier

### Single-tier storage pools

Figure 10-6 shows a scenario in which a single storage pool is populated with MDisks that are presented by an external storage controller. In this solution, the striped volumes can be measured by Easy Tier, and can benefit from *Storage Pool Balancing* mode, which moves extents between MDisks of the same type.

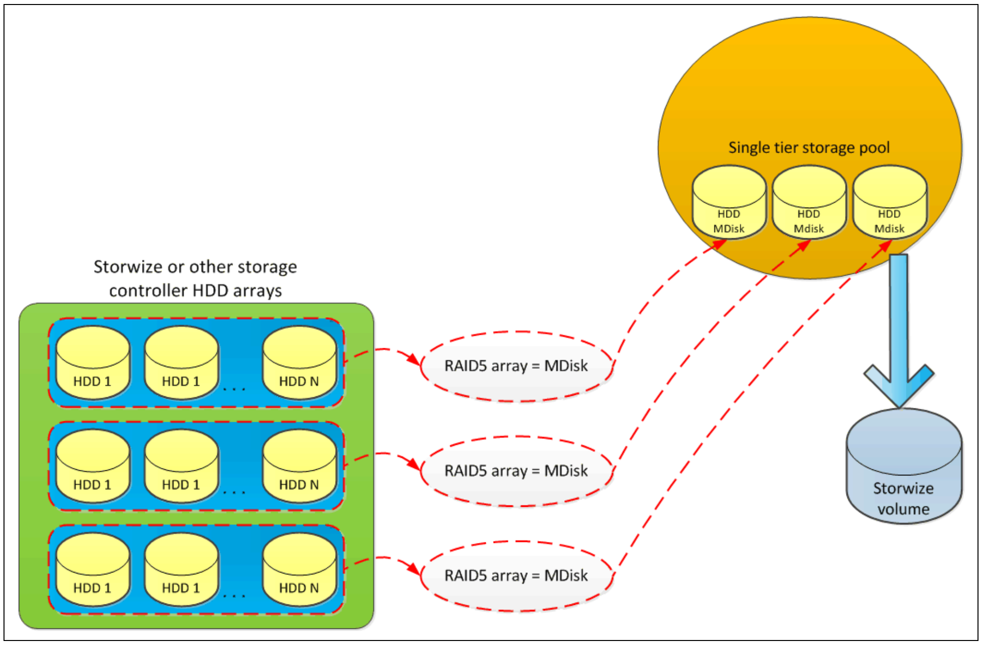


*Figure 10-6 Single tier storage pool with striped volume*

MDisks that are used in a single-tier storage pool should have the same hardware characteristics, for example, the same RAID type, RAID array size, disk type, disk RPM, and controller performance characteristics.

### Multitier storage pools

A multitier storage pool has a mix of MDisks with more than one type of disk tier attribute, for example, a storage pool that contains a mix of enterprise and SSD MDisks or enterprise and NL-SAS MDisks.

Figure 10-7 shows a scenario in which a storage pool is populated with three different MDisk types (one belonging to an SSD array, one belonging to an SAS HDD array, and one belonging to an NL-SAS HDD array). Although this example shows RAID 5 arrays, other RAID types can be used as well.

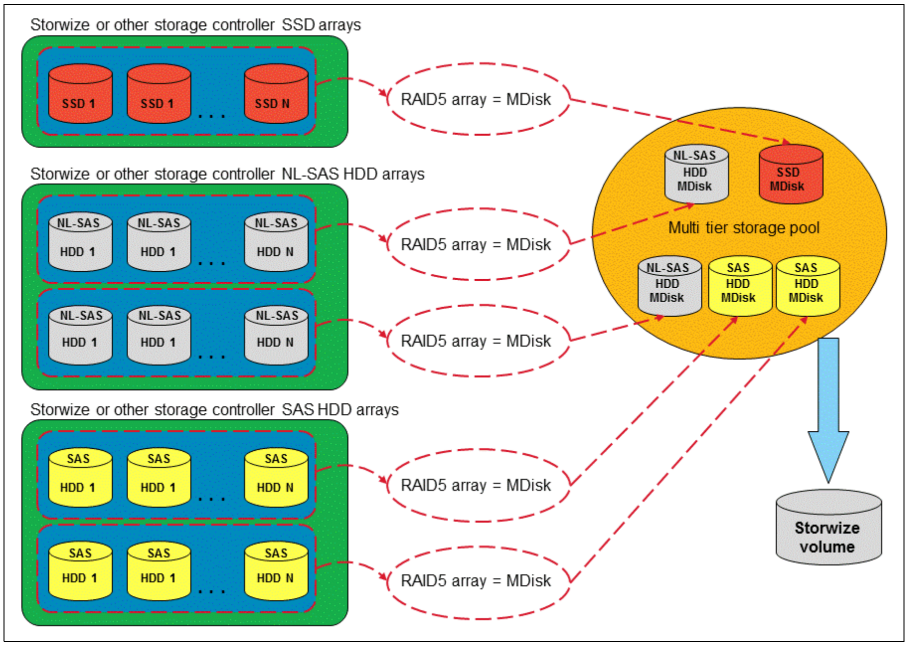


*Figure 10-7 Multitier storage pool with striped volume*

Adding SSDs to the pool also means that more space is now available for new volumes or volume expansion.

**Note:** Image mode and sequential volumes are not candidates for Easy Tier automatic data placement, because all extents for those types of volumes must be on one specific MDisk, and cannot be moved.

The Easy Tier setting can be changed on a storage pool and volume level. Depending on the Easy Tier setting and the number of tiers in the storage pool, Easy Tier services might function in a different way. Table 10-1 shows possible combinations of Easy Tier setting.

*Table 10-1 EasyTier settings*

| Storage pool Easy Tier setting | Number of tiers in the storage pool | Volume copy Easy Tier setting | Volume copy Easy Tier status |
|---|---|---|---|
| Off | One | off | inactive (see note 2) |
| Off | One | on | inactive (see note 2) |
| Off | Two or three | off | inactive (see note 2) |
| Off | Two or three | on | inactive (see note 2) |
| Measure | One | off | measured (see note 3) |
| Measure | One | on | measured (see note 3) |
| Measure | Two or three | off | measured (see note 3) |
| Measure | Two or three | on | measured (see note 3) |

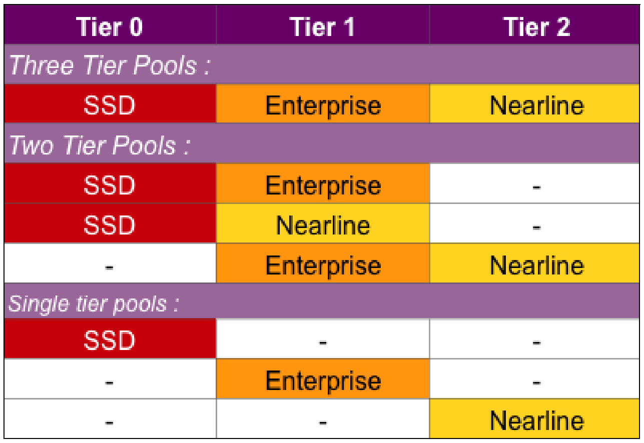
| Storage pool Easy Tier setting | Number of tiers in the storage pool | Volume copy Easy Tier setting | Volume copy Easy Tier status |
|---|---|---|---|
| Auto | One | off | measured (see note 3) |
| Auto | One | on | balanced (see note 4) |
| Auto | Two or three | off | measured (see note 3) |
| Auto | Two or three | on | active (see note 5) |
| On | One | off | measured (see note 3) |
| On | One | on | balanced (see note 4) |
| On | Two or three | off | measured (see note 3) |
| On | Two or three | on | active (see note 5) |

**Table notes:**

1. If the volume copy is in image or sequential mode, or is being migrated, the volume copy Easy Tier status is measured rather than active.
2. When the volume copy status is inactive, no Easy Tier functions are enabled for that volume copy.
3. When the volume copy status is measured, the Easy Tier function collects usage statistics for the volume, but automatic data placement is not active.
4. When the volume copy status is balanced, the Easy Tier function enables performance-based pool balancing for that volume copy.
5. When the volume copy status is active, the Easy Tier function operates in automatic data placement mode for that volume.

The default Easy Tier setting for a storage pool is Auto, and the default Easy Tier setting for a volume copy is On. Therefore, Easy Tier functions, except pool performance balancing, are disabled for storage pools with a single tier. Automatic data placement mode is enabled by default for all striped volume copies in a storage pool with two or more tiers.

Figure 10-8 shows the naming convention and all supported combinations of storage tiering used by Easy Tier.



| Tier 0 | Tier 1 | Tier 2 |
|---|---|---|
| *Three Tier Pools :* | | |
| SSD | Enterprise | Nearline |
| *Two Tier Pools :* | | |
| SSD | Enterprise | - |
| SSD | Nearline | - |
| - | Enterprise | Nearline |
| *Single tier pools :* | | |
| SSD | - | - |
| - | Enterprise | - |
| - | - | Nearline |

*Figure 10-8 Easy Tier supported storage pools*

### 10.2.4 Read Intensive flash drive and Easy Tier

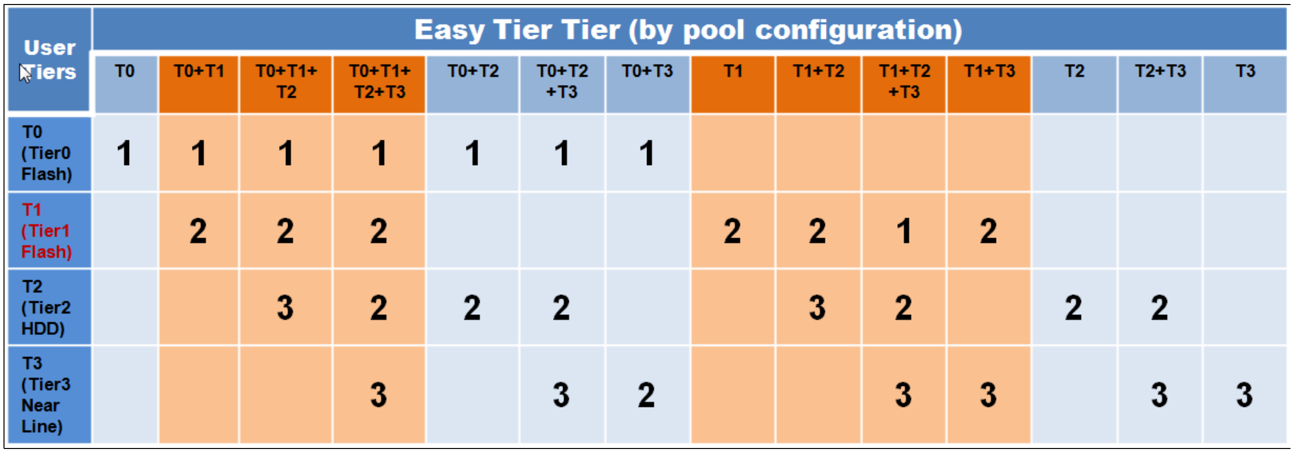
One of the reasons why flash technology is still quite expensive when compared to traditional HDD, is that an over provisioning of the physical memory is provided to mitigate the Write Amplification issue (https://en.wikipedia.org/wiki/Write_amplification). Read-Intensive (RI) flash drives are lower-cost flash drives with the cost reduction being achieved by having less redundant flash material.

Read Intensive flash drive support for Spectrum Virtualize/Storwize systems has been initially introduced with V7.7 and then enhanced with V7.8 introducing, among other things, Easy Tier support for RI MDisks.

Even though Easy Tier still remains a three tier storage architecture, 7.8 added a new “user” tier specifically for the RI MDisks. From a user perspective then there are now four tiers:

- T0 or *Tier0_flash* that represent the enterprise flash technology
- T1 or *Tier1_flash* that represent the RI flash technology
- T2 or *Tier2_HDD* that represent the enterprise HDD technology
- T3 or *Tier3_Nearline* that represent the nearline HDD technology

These user tiers are mapped to Easy Tier tiers depending on the pool configuration. The table in Figure 10-9 shows the possible combinations for the pool configuration with regard to the four user tiers (in orange we have highlighted the configurations containing the RI user tier).



| User Tiers | Easy Tier Tier (by pool configuration) | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | T0 | T0+T1 | T0+T1+ T2 | T0+T1+ T2+T3 | T0+T2 | T0+T2 +T3 | T0+T3 | T1 | T1+T2 | T1+T2 +T3 | T1+T3 | T2 | T2+T3 | T3 |
| T0 (Tier0 Flash) | 1 | 1 | 1 | 1 | 1 | 1 | 1 | | | | | | | |
| T1 (Tier1 Flash) | | 2 | 2 | 2 | | | | 2 | 2 | 1 | 2 | | | |
| T2 (Tier2 HDD) | | | 3 | 2 | 2 | 2 | | | 3 | 2 | | 2 | 2 | |
| T3 (Tier3 Near Line) | | | | 3 | | 3 | 2 | | | 3 | 3 | | 3 | 3 |

*Figure 10-9 Easy Tier mapping policy*

The table columns represent all the possible pool configurations, while the rows reports in which Easy Tier tier each user tier is mapped. For example, consider a pool with all the possible tiers configured that corresponds with the T0+T1+T2+T3 configuration in the table. With this configuration the T1 and T2 are mapped to the same Easy Tier tier (tier 2). Note that the Tier1_flash tier is only mapped to Easy Tier 1 or 2 tier.

For more information about planning and configuration considerations or best practices see *IBM System Storage SAN Volume Controller and Storwize V7000 Best Practices and Performance Guidelines*, SG24-7521 and *Implementing IBM Easy Tier with IBM Real-time Compression*, TIPS1072.

We discuss RI SSD at a 7.7 level in:

*Read Intensive Flash Drives*, REDP-5380

### 10.2.5 Easy Tier process

The Easy Tier function includes the following four main processes:

- I/O Monitoring

  This process operates continuously and monitors volumes for host I/O activity. It collects performance statistics for each extent, and derives averages for a rolling 24-hour period of I/O activity.

  Easy Tier makes allowances for large block I/Os; therefore, it considers only I/Os of up to 64 kibibytes (KiB) as migration candidates.

  This process is efficient and adds negligible processing resource use to the IBM SAN Volume Controller nodes.

- Data Placement Advisor

  The Data Placement Advisor uses workload statistics to make a cost benefit decision as to which extents are to be candidates for migration to a higher performance tier.

  This process also identifies extents that must be migrated back to a lower tier.

- Data Migration Planner (DMP)

  By using the extents that were previously identified, the DMP builds the extent migration plans for the storage pool. The DMP builds two plans:

  - Automatic Data Relocation (ADR mode) plan to migrate extents across adjacent tiers
  - Rebalance (RB mode) plan to migrate extents within the same tier

- Data Migrator

  This process involves the actual movement or migration of the volume’s extents up to, or down from, the higher disk tier. The extent migration rate is capped so that a maximum of up to 30 megabytes per second (MBps) is migrated, which equates to approximately 3 terabytes (TB) per day that is migrated between disk tiers.

When enabled, Easy Tier performs the following actions between three tiers presented in Figure 10-8 on page 372:

- Promote:

  Moves the relevant hot extents to higher performing tier.

- Swap:

  Exchanges cold extent in upper tier with hot extent in lower tier.

- Warm Demote:

  - Prevents performance overload of a tier by demoting a warm extent to the lower tier.
  - Triggered when bandwidth or IOPS exceeds predefined threshold.

- Demote or Cold Demote:

  Coldest data is moved to lower HDD tier. Only supported between HDD tiers.

- Expanded Cold Demote:

  Demotes appropriate sequential workloads to the lowest tier to better use nearline disk bandwidth.

- Storage Pool Balancing:

  - Redistributes extents within a tier to balance usage across MDisks for maximum performance.
  - Moves hot extents from high used MDisks to low used MDisks.

   - Exchanges extents between high used MDisks and low used MDisks.
- Easy Tier attempts to migrate the most active volume extents up to SSD first.
- A previous migration plan and any queued extents that are not yet relocated are abandoned.

> **Note:** Extent migration occurs only between adjacent tiers. In a three-tiered storage pool, Easy Tier will not move extents from SSDs directly to NL-SAS and vice versa without moving them first to SAS drives.

Easy Tier extent migration types are presented in Figure 10-10.

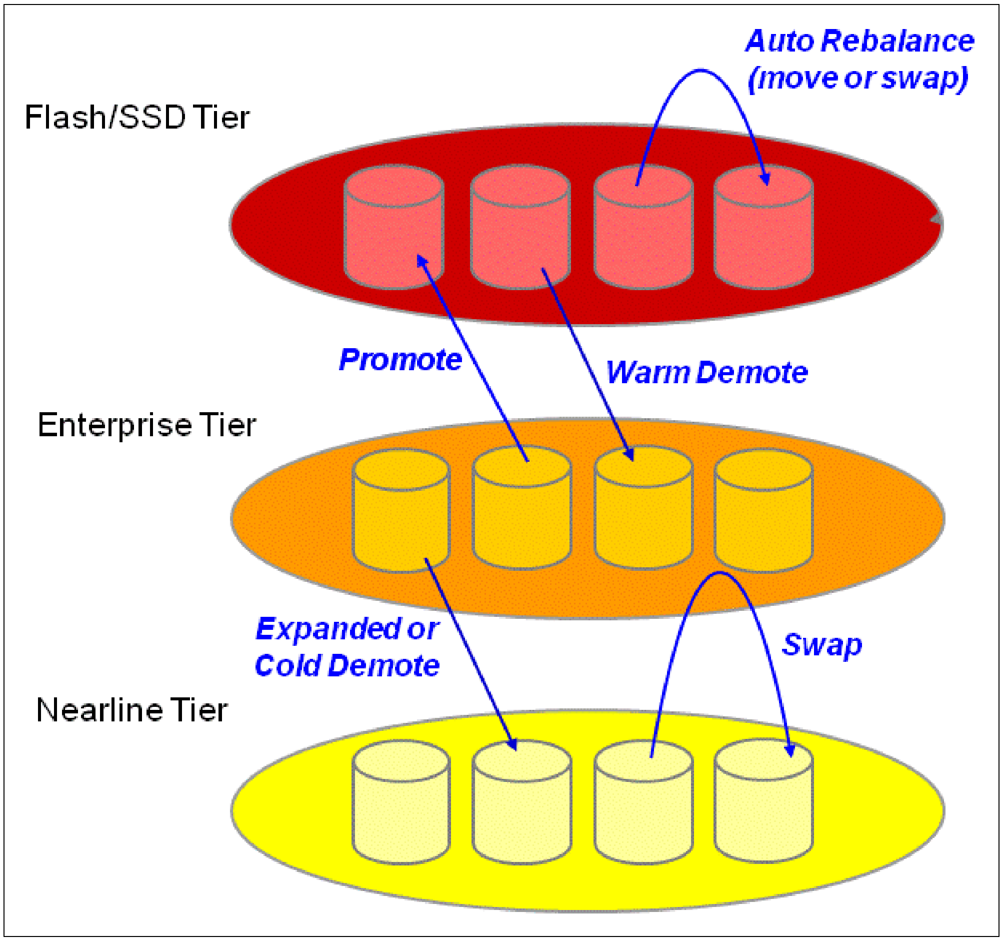


*Figure 10-10 Easy Tier extent migration types*

## 10.2.6 Easy Tier operating modes

Easy Tier includes the following main operating modes:

- Off
- Evaluation or measurement only
- Automatic data placement or extent migration
- Storage pool balancing

### Easy Tier off mode

With Easy Tier turned off, no statistics are recorded, and no cross-tier extent migration occurs.

### Evaluation or measurement only mode

Easy Tier Evaluation or measurement-only mode collects usage statistics for each extent in a single-tier storage pool where the Easy Tier value is set to On for both the volume and the pool. This collection is typically done for a single-tier pool that contains only HDDs so that the benefits of adding SSDs to the pool can be evaluated before any major hardware acquisition.

A `dpa_heat.nodeid.yymmdd.hhmmss.data` statistics summary file is created in the `/dumps` directory of the Storwize V7000 node canisters. This file can be offloaded from the Storwize node canisters with PuTTY Secure Copy Client (PSCP) **`-load`** command or by using the GUI, as described in *IBM System Storage SAN Volume Controller and Storwize V7000 Best Practices and Performance Guidelines*, SG24-7521. A web browser is used to view the report that is created by the tool.

### Automatic Data Placement or extent migration mode

In Automatic data placement or extent migration operating mode, the storage pool parameter **`-easytier on`** or **`auto`** must be set, and the volumes in the pool must have **`-easytier on`**. The storage pool must also contain MDisks with different disk tiers, which makes it a multitier storage pool.

Dynamic data movement is not apparent to the host server and application users of the data, other than providing improved performance. Extents are automatically migrated, as explained in “Implementation rules” on page 377.

The statistic summary file is also created in this mode. This file can be offloaded for input to the advisor tool. The tool produces a report on the extents that are moved to a higher tier, and a prediction of performance improvement that can be gained if more higher tier disks are available.

**Options:** The Easy Tier function can be turned on or off at the storage pool level *and* at the volume level.

### Storage Pool Balancing

Although storage pool balancing is associated with Easy Tier, it operates independently of Easy Tier, and does not require an Easy Tier license. This feature assesses the extents that are written in a pool, and balances them automatically across all MDisks within the pool. This process works along with Easy Tier when multiple classes of disks exist in a single pool. In such a case, Easy Tier moves extents between the different tiers, and storage pool balancing moves extents within the same tier, to better use MDisks.

The process automatically balances existing data when new MDisks are added into an existing pool, even if the pool only contains a single type of drive. This does not mean that the process migrates extents from existing MDisks to achieve even extent distribution among all, old and new, MDisks in the storage pool. The Easy Tier rebalancing process within a tier migration plan is based on performance, not on the capacity of underlying MDisks.

**Note:** Storage pool balancing can be used to balance extents when mixing different size disks of the same performance tier. For example, when adding larger capacity drives to a pool with smaller capacity drives of the same class, Storage pool balancing redistributes the extents to take advantage of the additional performance of the new MDisks.

## 10.2.7 Implementation considerations

Easy Tier is a licensed feature, except for storage pool balancing, which is a no-charge feature that is enabled by default. Easy Tier comes as part of the IBM Spectrum Virtualize code. For Easy Tier to migrate extents between different tier disks, you must have disk storage available that offers different tiers (for example, a mix of SSD and HDD). Easy Tier uses Storage Pool Balancing if you have only single tier pool.

### Implementation rules

Remember the following implementation and operational rules when you use the IBM System Storage Easy Tier function on the IBM Storwize V7000:

- Easy Tier automatic data placement is not supported on image mode or sequential volumes. I/O monitoring for such volumes is supported, but you cannot migrate extents on these volumes unless you convert image or sequential volume copies to striped volumes.
- Automatic data placement and extent I/O activity monitors are supported on each copy of a mirrored volume. Easy Tier works with each copy independently of the other copy.

> **Volume mirroring consideration:** Volume mirroring can have different workload characteristics on each copy of the data because reads are normally directed to the primary copy and writes occur to both copies. Therefore, the number of extents that Easy Tier migrates between the tiers might be different for each copy.

- If possible, the IBM Storwize V7000 creates volumes or expands volumes by using extents from MDisks from the HDD tier. However, if necessary, it uses extents from MDisks from the SSD tier.

When a volume is migrated out of a storage pool that is managed with Easy Tier, Easy Tier automatic data placement mode is no longer active on that volume. Automatic data placement is also turned off while a volume is being migrated, even when it is between pools that both have Easy Tier automatic data placement enabled. Automatic data placement for the volume is re-enabled when the migration is complete.

### Limitations

When you use Easy Tier on the IBM Storwize V7000, remember the following limitations:

- Removing an MDisk by using the `-force` parameter

  When an MDisk is deleted from a storage pool with the `-force` parameter, extents in use are migrated to MDisks in the same tier as the MDisk that is being removed, if possible. If insufficient extents exist in that tier, extents from the other tier are used.

- Migrating extents

  When Easy Tier automatic data placement is enabled for a volume, you cannot use the `svctask migrateexts` CLI command on that volume.

- Migrating a volume to another storage pool

  When IBM Storwize V7000 migrates a volume to a new storage pool, Easy Tier automatic data placement between the two tiers is temporarily suspended. After the volume is migrated to its new storage pool, Easy Tier automatic data placement between the generic SSD tier and the generic HDD tier resumes for the moved volume, if appropriate.

  When the IBM Storwize V7000 migrates a volume from one storage pool to another, it attempts to migrate each extent to an extent in the new storage pool from the same tier as the original extent. In several cases, such as where a target tier is unavailable, the other tier is used. For example, the generic SSD tier might be unavailable in the new storage pool.

- Migrating a volume to an image mode

  Easy Tier automatic data placement does not support image mode. When a volume with active Easy Tier automatic data placement mode is migrated to an image mode, Easy Tier automatic data placement mode is no longer active on that volume.

- Image mode and sequential volumes cannot be candidates for automatic data placement; however, Easy Tier supports evaluation mode for image mode volumes.

## 10.2.8 Modifying the Easy Tier setting

The Easy Tier setting for storage pools and volumes can only be changed from the command-line interface. All of the changes are done online without any effect on hosts or data availability.

### Turning Easy Tier on and off

Use the **chvdisk** command to turn off or turn on Easy Tier on selected volumes. Use the **chmdiskgrp** to change status of Easy Tier on selected storage pools as shown in Example 10-2.

*Example 10-2 Changing Easy Tier setting*

```
IBM_Storwize:V7000 Gen 2:superuser>lsvdisk test_vol_2
id 11
name test_vol_2
IO_group_id 0
IO_group_name io_grp0
status online
mdisk_grp_id 0
mdisk_grp_name test_pool_1
capacity 5.00GB
type striped
formatted no
formatting yes
mdisk_id
mdisk_name
FC_id
FC_name
RC_id
RC_name
vdisk_UID 600507680283818B300000000000000D
throttling 0
preferred_node_id 1
fast_write_state not_empty
cache readwrite
udid
fc_map_count 0
sync_rate 50
copy_count 1
se_copy_count 0
filesystem
mirror_write_priority latency
RC_change no
compressed_copy_count 0
access_IO_group_count 1
last_access_time
```

```
parent_mdisk_grp_id 0
parent_mdisk_grp_name test_pool_1
owner_type none
owner_id
owner_name
encrypt no
volume_id 11
volume_name test_vol_2
function

copy_id 0
status online
sync yes
auto_delete no
primary yes
mdisk_grp_id 0
mdisk_grp_name test_pool_1
type striped
mdisk_id
mdisk_name
fast_write_state not_empty
used_capacity 5.00GB
real_capacity 5.00GB
free_capacity 0.00MB
overallocation 100
autoexpand
warning
grainsize
se_copy no
easy_tier off
easy_tier_status measured
tier ssd
tier_capacity 5.00GB
tier enterprise
tier_capacity 0.00MB
tier nearline
tier_capacity 0.00MB
compressed_copy no
uncompressed_used_capacity 5.00GB
parent_mdisk_grp_id 0
parent_mdisk_grp_name test_pool_1
encrypt no

IBM_Storwize:V7000 Gen 2:superuser>chvdisk -easytier on test_vol_2

IBM_Storwize:V7000 Gen 2:superuser>lsvdisk test_vol_2
id 11
name test_vol_2
IO_group_id 0
IO_group_name io_grp0
status online
mdisk_grp_id 0
mdisk_grp_name test_pool_1
capacity 5.00GB
type striped
```

```
formatted no
formatting yes
mdisk_id
mdisk_name
FC_id
FC_name
RC_id
RC_name
vdisk_UID 600507680283818B300000000000000D
throttling 0
preferred_node_id 1
fast_write_state not_empty
cache readwrite
udid
fc_map_count 0
sync_rate 50
copy_count 1
se_copy_count 0
filesystem
mirror_write_priority latency
RC_change no
compressed_copy_count 0
access_IO_group_count 1
last_access_time
parent_mdisk_grp_id 0
parent_mdisk_grp_name test_pool_1
owner_type none
owner_id
owner_name
encrypt no
volume_id 11
volume_name test_vol_2
function

copy_id 0
status online
sync yes
auto_delete no
primary yes
mdisk_grp_id 0
mdisk_grp_name test_pool_1
type striped
mdisk_id
mdisk_name
fast_write_state not_empty
used_capacity 5.00GB
real_capacity 5.00GB
free_capacity 0.00MB
overallocation 100
autoexpand
warning
grainsize
se_copy no
easy_tier on
easy_tier_status balanced
```

```
tier ssd
tier_capacity 5.00GB
tier enterprise
tier_capacity 0.00MB
tier nearline
tier_capacity 0.00MB
compressed_copy no
uncompressed_used_capacity 5.00GB
parent_mdisk_grp_id 0
parent_mdisk_grp_name test_pool_1
encrypt no

IBM_Storwize:V7000 Gen 2:superuser>lsmdiskgrp test_pool_1
id 0
name test_pool_1
status online
mdisk_count 1
vdisk_count 10
capacity 2.70TB
extent_size 1024
free_capacity 1.52TB
virtual_capacity 185.00GB
used_capacity 185.00GB
real_capacity 185.00GB
overallocation 6
warning 80
easy_tier auto
easy_tier_status balanced
tier ssd
tier_mdisk_count 1
tier_capacity 2.70TB
tier_free_capacity 2.52TB
tier enterprise
tier_mdisk_count 0
tier_capacity 0.00MB
tier_free_capacity 0.00MB
tier nearline
tier_mdisk_count 0
tier_capacity 0.00MB
tier_free_capacity 0.00MB
compression_active no
compression_virtual_capacity 0.00MB
compression_compressed_capacity 0.00MB
compression_uncompressed_capacity 0.00MB
site_id
site_name
parent_mdisk_grp_id 0
parent_mdisk_grp_name test_pool_1
child_mdisk_grp_count 1
child_mdisk_grp_capacity 1.00TB
type parent
encrypt no
owner_type none
owner_id
owner_name
```

```
IBM_Storwize:V7000 Gen 2:superuser>chmdiskgrp -easytier off test_pool_1


IBM_Storwize:V7000 Gen 2:superuser>lsmdiskgrp test_pool_1
id 0
name test_pool_1
status online
mdisk_count 1
vdisk_count 10
capacity 2.70TB
extent_size 1024
free_capacity 1.52TB
virtual_capacity 185.00GB
used_capacity 185.00GB
real_capacity 185.00GB
overallocation 6
warning 80
easy_tier off
easy_tier_status inactive
tier ssd
tier_mdisk_count 1
tier_capacity 2.70TB
tier_free_capacity 2.52TB
tier enterprise
tier_mdisk_count 0
tier_capacity 0.00MB
tier_free_capacity 0.00MB
tier nearline
tier_mdisk_count 0
tier_capacity 0.00MB
tier_free_capacity 0.00MB
compression_active no
compression_virtual_capacity 0.00MB
compression_compressed_capacity 0.00MB
compression_uncompressed_capacity 0.00MB
site_id
site_name
parent_mdisk_grp_id 0
parent_mdisk_grp_name test_pool_1
child_mdisk_grp_count 1
child_mdisk_grp_capacity 1.00TB
type parent
encrypt no
owner_type none
owner_id
owner_name
```

### Tuning Easy Tier

It is also possible to change more advanced parameters of Easy Tier. These parameters should be used with caution, because changing the default values can affect system performance.

### Easy Tier acceleration

The first setting is called *Easy Tier acceleration*. This is a system-wide setting, and is disabled by default. Turning on this setting makes Easy Tier move extents up to four times faster than when in default setting. In accelerate mode Easy Tier can move up to 48 GiB per 5 minutes while in normal mode it moves up to 12 GiB. Enabling Easy Tier acceleration is advised only during periods of low system activity. The following two use cases for acceleration are the most probable:

- When adding new capacity to the pool, accelerating Easy Tier can quickly spread existing volumes onto the new MDisks.
- When migrating the volumes between the storage pools in cases where the target storage pool has more tiers than the source storage pool, accelerating Easy Tier can quickly promote or demote extents in the target pool.

This setting can be changed online, without any effect on host or data availability. To turn Easy Tier acceleration mode on or off, use the **chsystem** command, as shown in bold in Example 10-3.

*Example 10-3 chsystem command*

```
IBM_Storwize:V7000 Gen 2:superuser>lssystem
id 000001002140020E
name ITSO Gen2
location local
partnership
total_mdisk_capacity 1.5TB
space_in_mdisk_grps 1.1TB
space_allocated_to_vdisks 18.00MB
total_free_space 1.5TB
total_vdiskcopy_capacity 2.00GB
total_used_capacity 0.16MB
total_overallocation 0
total_vdisk_capacity 2.00GB
total_allocated_extent_capacity 1.00GB
statistics_status on
statistics_frequency 15
cluster_locale en_US
time_zone 520 US/Pacific
code_level 7.8.0.0 (build 133.17.1612071632000)
console_IP 10.18.228.71:443
id_alias 000001002140020E
gm_link_tolerance 300
gm_inter_cluster_delay_simulation 0
gm_intra_cluster_delay_simulation 0
gm_max_host_delay 5
email_reply ITSO@ITSO.COM
email_contact ITSO
email_contact_primary 123456789
email_contact_alternate
email_contact_location
email_contact2
email_contact2_primary
email_contact2_alternate
email_state running
inventory_mail_interval 7
cluster_ntp_IP_address
```

```
cluster_isns_IP_address
iscsi_auth_method none
iscsi_chap_secret 1111
auth_service_configured no
auth_service_enabled no
auth_service_url
auth_service_user_name
auth_service_pwd_set no
auth_service_cert_set no
auth_service_type tip
relationship_bandwidth_limit 25
tier ssd
tier_capacity 0.00MB
tier_free_capacity 0.00MB
tier enterprise
tier_capacity 1.08TB
tier_free_capacity 1.08TB
tier nearline
tier_capacity 0.00MB
tier_free_capacity 0.00MB
easy_tier_acceleration off
has_nas_key no
layer storage
rc_buffer_size 48
compression_active yes
compression_virtual_capacity 2.00GB
compression_compressed_capacity 0.16MB
compression_uncompressed_capacity 0.00MB
cache_prefetch on
email_organization
email_machine_address
email_machine_city
email_machine_state XX
email_machine_zip
email_machine_country
total_drive_raw_capacity 4.91TB
compression_destage_mode off
local_fc_port_mask
1111111111111111111111111111111111111111111111111111111111111111
partner_fc_port_mask
1111111111111111111111111111111111111111111111111111111111111111
high_temp_mode off
topology standard
topology_status
rc_auth_method chap
vdisk_protection_time 15
vdisk_protection_enabled no
product_name IBM Storwize V7000
odx off
max_replication_delay 0

IBM_Storwize:ITSO Gen2:superuser>chsystem -easytieracceleration on

IBM_Storwize:ITSO Gen2:superuser>lssystem
id 000001002140020E
```

```
name ITSO Gen2
location local
partnership
total_mdisk_capacity 1.5TB
space_in_mdisk_grps 1.1TB
space_allocated_to_vdisks 18.00MB
total_free_space 1.5TB
total_vdiskcopy_capacity 2.00GB
total_used_capacity 0.16MB
total_overallocation 0
total_vdisk_capacity 2.00GB
total_allocated_extent_capacity 1.00GB
statistics_status on
statistics_frequency 15
cluster_locale en_US
time_zone 520 US/Pacific
code_level 7.8.0.0 (build 133.17.1612071632000)
console_IP 10.18.228.71:443
id_alias 000001002140020E
gm_link_tolerance 300
gm_inter_cluster_delay_simulation 0
gm_intra_cluster_delay_simulation 0
gm_max_host_delay 5
email_reply ITSO@ITSO.COM
email_contact ITSO
email_contact_primary 123456789
email_contact_alternate
email_contact_location
email_contact2
email_contact2_primary
email_contact2_alternate
email_state running
inventory_mail_interval 7
cluster_ntp_IP_address
cluster_isns_IP_address
iscsi_auth_method none
iscsi_chap_secret 1111
auth_service_configured no
auth_service_enabled no
auth_service_url
auth_service_user_name
auth_service_pwd_set no
auth_service_cert_set no
auth_service_type tip
relationship_bandwidth_limit 25
tier ssd
tier_capacity 0.00MB
tier_free_capacity 0.00MB
tier enterprise
tier_capacity 1.08TB
tier_free_capacity 1.08TB
tier nearline
tier_capacity 0.00MB
tier_free_capacity 0.00MB
easy_tier_acceleration on
```

```
has_nas_key no
layer storage
rc_buffer_size 48
compression_active yes
compression_virtual_capacity 2.00GB
compression_compressed_capacity 0.16MB
compression_uncompressed_capacity 0.00MB
cache_prefetch on
email_organization
email_machine_address
email_machine_city
email_machine_state XX
email_machine_zip
email_machine_country
total_drive_raw_capacity 4.91TB
compression_destage_mode off
local_fc_port_mask
1111111111111111111111111111111111111111111111111111111111111111
partner_fc_port_mask
1111111111111111111111111111111111111111111111111111111111111111
high_temp_mode off
topology standard
topology_status
rc_auth_method chap
vdisk_protection_time 15
vdisk_protection_enabled no
product_name IBM Storwize V7000
odx off
max_replication_delay 0
```

### MDisk Easy Tier load

The second setting is called *MDisk Easy Tier load*. This setting is set per MDisk basis, and indicates how much load Easy Tier can put on the particular MDisk. There are several different values that can be set to each MDisk:

- Default
- Low
- Medium
- High
- Very high

The system uses a default setting based on the storage tier of the presented MDisks, either flash, nearline, or enterprise. If the disk drives are internal, the tier is known. However, an external MDisk tier should be changed by the user to align it with underlying storage.

Change the default setting to any other value only when you are certain that a particular MDisk is under used and can handle more load, or that the MDisk is overutilized and the load should be lowered. Change this setting to *very high* only for SDD and flash MDisks.

This setting can be changed online, without any effect on the hosts or data availability.

To change this setting use command **chmdisk** as seen in Example 10-4.

*Example 10-4 The chmdisk command*

```
IBM_Storwize:V7000 Gen 2:superuser>lsmdisk mdisk0
id 2
name mdisk0
status online
mode unmanaged
mdisk_grp_id
mdisk_grp_name
capacity 32.0GB
quorum_index
block_size 512
controller_name controller0
ctrl_type 4
ctrl_WWNN 200600A0B85AD223
controller_id 0
path_count 2
max_path_count 2
ctrl_LUN_# 0000000000000000
UID 600a0b80005ad22300000371527a29b200000000000000000000000000000000
preferred_WWPN 202600A0B85AD223
active_WWPN 202600A0B85AD223
fast_write_state empty
raid_status
raid_level
redundancy
strip_size
spare_goal
spare_protection_min
balanced
tier enterprise
slow_write_priority
fabric_type fc
site_id
site_name
easy_tier_load medium
encrypt no
distributed no
drive_class_id
drive_count 0
stripe_width 0
rebuild_areas_total
rebuild_areas_available
rebuild_areas_goal

IBM_Storwize:V7000 Gen 2:superuser>chmdisk -easytierload high mdisk0

IBM_Storwize:V7000 Gen 2:superuser>lsmdisk mdisk0
id 2
name mdisk0
status online
mode unmanaged
mdisk_grp_id
mdisk_grp_name
```

```
capacity 32.0GB
quorum_index
block_size 512
controller_name controller0
ctrl_type 4
ctrl_WWNN 200600A0B85AD223
controller_id 0
path_count 2
max_path_count 2
ctrl_LUN_# 0000000000000000
UID 600a0b80005ad22300000371527a29b200000000000000000000000000000000
preferred_WWPN 202600A0B85AD223
active_WWPN 202600A0B85AD223
fast_write_state empty
raid_status
raid_level
redundancy
strip_size
spare_goal
spare_protection_min
balanced
tier enterprise
slow_write_priority
fabric_type fc
site_id
site_name
easy_tier_load high
encrypt no
distributed no
drive_class_id
drive_count 0
stripe_width 0
rebuild_areas_total
rebuild_areas_available
rebuild_areas_goal
```

## 10.2.9 Monitoring tools

The IBM Storage Tier Advisor Tool (STAT) is a Microsoft Windows console application that analyzes heat data files produced by Easy Tier. STAT creates a graphical display of the amount of “hot” data per volume. It predicts, by storage pool, how more flash drives (or SSD capacity), enterprise drives, and nearline drives might improve system performance.

Heat data files are produced approximately once a day (that is, every 24 hours) when Easy Tier is active on one or more storage pools. These files summarize the activity per volume since the prior heat data file was produced. On the IBM Storwize V7000, the heat data file is in the /dumps directory on the configuration node, and is named dpa_heat.<*node_name*>.<*time_stamp*>.data.

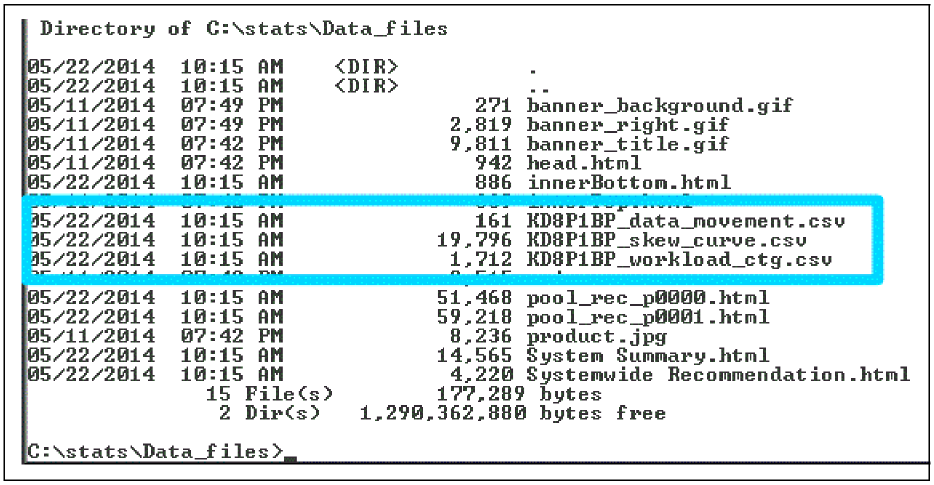
Any existing heat data file is erased after seven days. The file must be offloaded by the user and STAT must be started from a Windows command prompt console with the file specified as a parameter. The user can also specify the output directory. STAT creates a set of Hypertext Markup Language (HTML) files, and the user can then open the index.html in a browser to view the results.

Updates to the STAT for IBM Spectrum Virtualize V7.3 have introduced more capability for reporting. As a result, when the STAT tool is run on a heat map file, an extra three comma-separated values (CSV) files are created and placed in the `Data_files` directory.

The IBM STAT tool can be downloaded from the IBM Support website:

http://www.ibm.com/support/docview.wss?uid=ssg1S4000935

Figure 10-11 shows the CSV files highlighted in the `Data_files` directory after running the stat tool over an IBM Storwize V7000 heatmap.



```
 Directory of C:\stats\Data_files

05/22/2014  10:15 AM    <DIR>          .
05/22/2014  10:15 AM    <DIR>          ..
05/11/2014  07:49 PM               271 banner_background.gif
05/11/2014  07:49 PM             2,819 banner_right.gif
05/11/2014  07:42 PM             9,811 banner_title.gif
05/11/2014  07:42 PM               942 head.html
05/22/2014  10:15 AM               886 innerBottom.html
05/11/2014  07:42 PM       [illegible] innerTop.html
05/22/2014  10:15 AM               161 KD8P1BP_data_movement.csv
05/22/2014  10:15 AM            19,796 KD8P1BP_skew_curve.csv
05/22/2014  10:15 AM             1,712 KD8P1BP_workload_ctg.csv
05/11/2014  07:42 PM       [illegible]
05/22/2014  10:15 AM            51,468 pool_rec_p0000.html
05/22/2014  10:15 AM            59,218 pool_rec_p0001.html
05/11/2014  07:42 PM             8,236 product.jpg
05/22/2014  10:15 AM            14,565 System Summary.html
05/22/2014  10:15 AM             4,220 Systemwide Recommendation.html
              15 File(s)        177,289 bytes
               2 Dir(s)   1,290,362,880 bytes free

C:\stats\Data_files>_
```

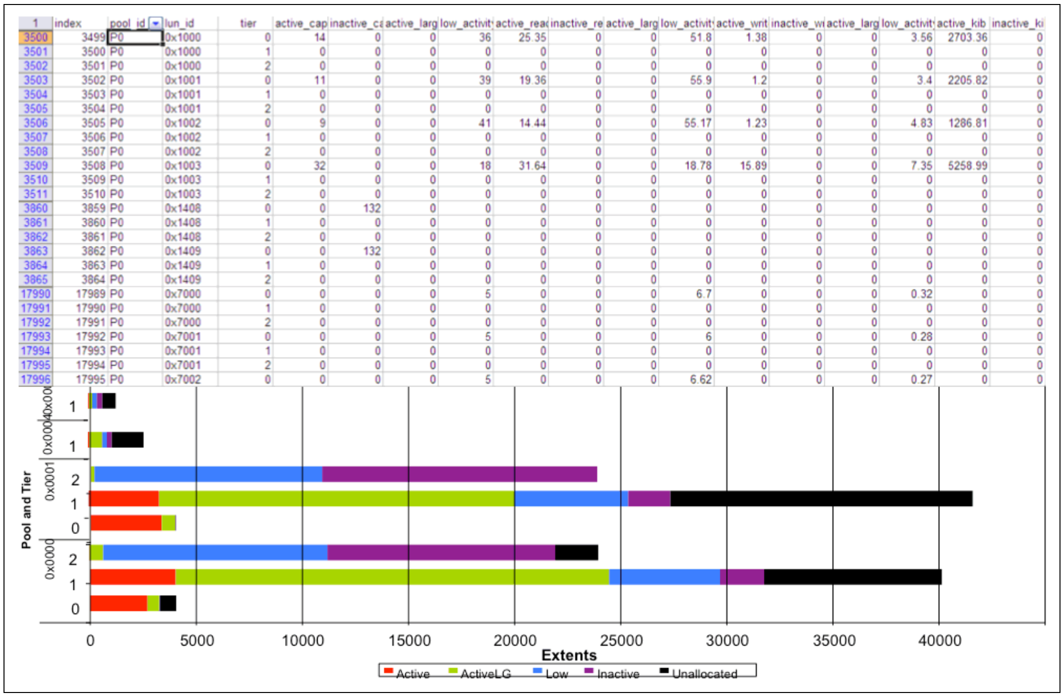
*Figure 10-11 CSV files created by the STAT for Easy Tier*

In addition to the STAT tool, IBM Spectrum Virtualize has another utility, which is a Microsoft SQL file for creating additional graphical reports of the workload that Easy Tier performs. The IBM STAT Charting Utility takes the output of the three CSV files and turns them into graphs for simple reporting.

The new graphs display the following information:

- Workload Categorization

  New workload visuals help you compare activity across tiers within and across pools to help determine the optimal drive mix for the current workloads. The output is illustrated in Figure 10-12.

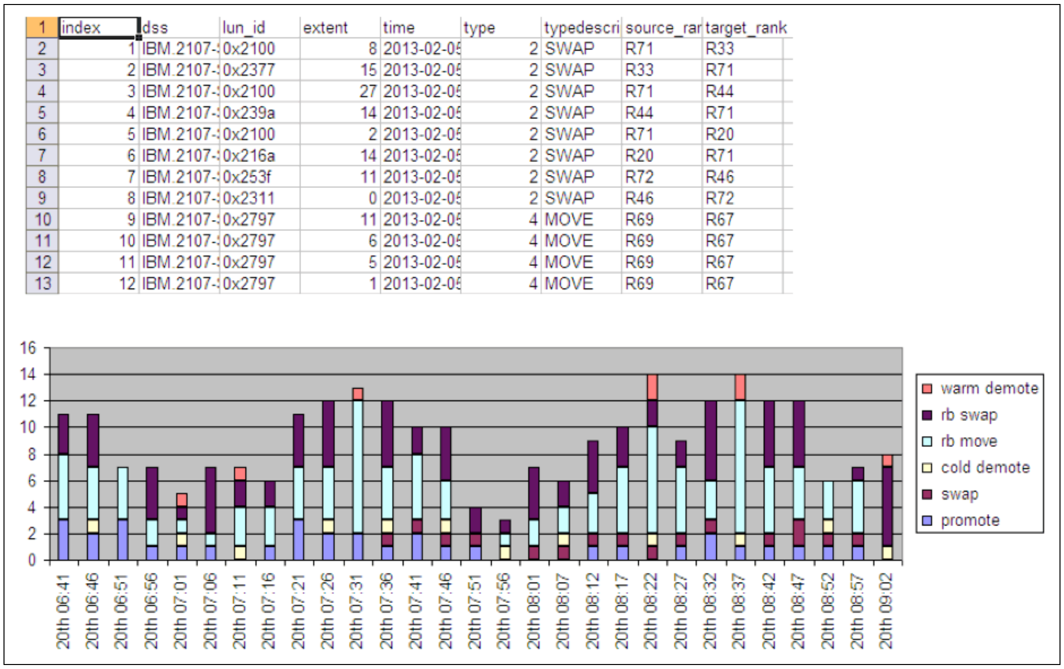


| 1 | index | pool_id | lun_id | tier | active_cap | inactive_ca | active_larg | low_activit | active_rea | inactive_re | active_larg | low_activit | active_writ | inactive_wi | active_larg | low_activit | active_kib | inactive_ki |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 3500 | 3499 | P0 | 0x1000 | 0 | 14 | 0 | 0 | 36 | 25.35 | 0 | 0 | 51.8 | 1.38 | 0 | 0 | 3.56 | 2703.36 | 0 |
| 3501 | 3500 | P0 | 0x1000 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 3502 | 3501 | P0 | 0x1000 | 2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 3503 | 3502 | P0 | 0x1001 | 0 | 11 | 0 | 0 | 39 | 19.36 | 0 | 0 | 55.9 | 1.2 | 0 | 0 | 3.4 | 2205.82 | 0 |
| 3504 | 3503 | P0 | 0x1001 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 3505 | 3504 | P0 | 0x1001 | 2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 3506 | 3505 | P0 | 0x1002 | 0 | 9 | 0 | 0 | 41 | 14.44 | 0 | 0 | 55.17 | 1.23 | 0 | 0 | 4.83 | 1286.81 | 0 |
| 3507 | 3506 | P0 | 0x1002 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 3508 | 3507 | P0 | 0x1002 | 2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 3509 | 3508 | P0 | 0x1003 | 0 | 32 | 0 | 0 | 18 | 31.64 | 0 | 0 | 18.78 | 15.89 | 0 | 0 | 7.35 | 5258.99 | 0 |
| 3510 | 3509 | P0 | 0x1003 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 3511 | 3510 | P0 | 0x1003 | 2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 3860 | 3859 | P0 | 0x1408 | 0 | 0 | 132 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 3861 | 3860 | P0 | 0x1408 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 3862 | 3861 | P0 | 0x1408 | 2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 3863 | 3862 | P0 | 0x1409 | 0 | 0 | 132 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 3864 | 3863 | P0 | 0x1409 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 3865 | 3864 | P0 | 0x1409 | 2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 17990 | 17989 | P0 | 0x7000 | 0 | 0 | 0 | 0 | 5 | 0 | 0 | 0 | 6.7 | 0 | 0 | 0 | 0.32 | 0 | 0 |
| 17991 | 17990 | P0 | 0x7000 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 17992 | 17991 | P0 | 0x7000 | 2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 17993 | 17992 | P0 | 0x7001 | 0 | 0 | 0 | 0 | 5 | 0 | 0 | 0 | 6 | 0 | 0 | 0 | 0.28 | 0 | 0 |
| 17994 | 17993 | P0 | 0x7001 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 17995 | 17994 | P0 | 0x7001 | 2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 17996 | 17995 | P0 | 0x7002 | 0 | 0 | 0 | 0 | 5 | 0 | 0 | 0 | 6.62 | 0 | 0 | 0 | 0.27 | 0 | 0 |

*Figure 10-12 STAT Charting Utility Workload Categorization report*

- Daily Movement report

  A new Easy Tier summary report every 24 hours illustrating data migration activity (5-minute intervals) can help visualize migration types and patterns for current workloads. The output is illustrated in Figure 10-13.

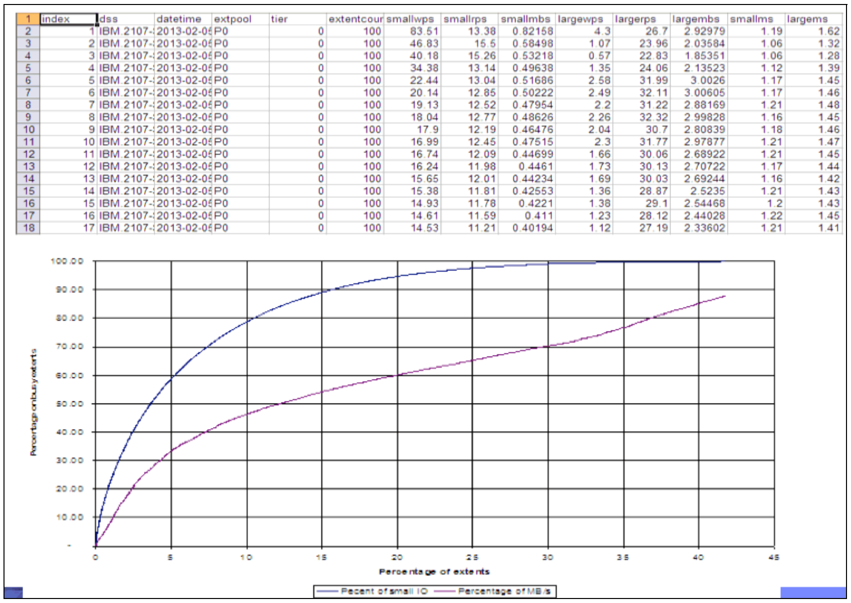


| 1 | index | dss | lun_id | extent | time | type | typedescri | source_rar | target_rank |
|---|---|---|---|---|---|---|---|---|---|
| 2 | 1 | IBM.2107-: | 0x2100 | 8 | 2013-02-05 | 2 | SWAP | R71 | R33 |
| 3 | 2 | IBM.2107-: | 0x2377 | 15 | 2013-02-05 | 2 | SWAP | R33 | R71 |
| 4 | 3 | IBM.2107-: | 0x2100 | 27 | 2013-02-05 | 2 | SWAP | R71 | R44 |
| 5 | 4 | IBM.2107-: | 0x239a | 14 | 2013-02-05 | 2 | SWAP | R44 | R71 |
| 6 | 5 | IBM.2107-: | 0x2100 | 2 | 2013-02-05 | 2 | SWAP | R71 | R20 |
| 7 | 6 | IBM.2107-: | 0x216a | 14 | 2013-02-05 | 2 | SWAP | R20 | R71 |
| 8 | 7 | IBM.2107-: | 0x253f | 11 | 2013-02-05 | 2 | SWAP | R72 | R46 |
| 9 | 8 | IBM.2107-: | 0x2311 | 0 | 2013-02-05 | 2 | SWAP | R46 | R72 |
| 10 | 9 | IBM.2107-: | 0x2797 | 11 | 2013-02-05 | 4 | MOVE | R69 | R67 |
| 11 | 10 | IBM.2107-: | 0x2797 | 6 | 2013-02-05 | 4 | MOVE | R69 | R67 |
| 12 | 11 | IBM.2107-: | 0x2797 | 5 | 2013-02-05 | 4 | MOVE | R69 | R67 |
| 13 | 12 | IBM.2107-: | 0x2797 | 1 | 2013-02-05 | 4 | MOVE | R69 | R67 |

*Figure 10-13 STAT Charting Utility Daily Summary report*

- Workload Skew report

  This report shows the skew of all workloads across the system in a graph to help you visualize and accurately tier configurations when you add capacity or a new system. The output is illustrated in Figure 10-14.



| 1 | index | dss | datetime | extpool | tier | extentcour | smallwps | smallrps | smallmbs | largewps | largerps | largembs | smallms | largems |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 2 | 1 | IBM.2107-: | 2013-02-05 | P0 | 0 | 100 | 83.51 | 13.38 | 0.82158 | 4.3 | 26.7 | 2.92979 | 1.19 | 1.62 |
| 3 | 2 | IBM.2107-: | 2013-02-05 | P0 | 0 | 100 | 46.83 | 15.5 | 0.58498 | 1.07 | 23.96 | 2.03584 | 1.06 | 1.32 |
| 4 | 3 | IBM.2107-: | 2013-02-05 | P0 | 0 | 100 | 40.18 | 15.26 | 0.53218 | 0.57 | 22.83 | 1.85351 | 1.06 | 1.28 |
| 5 | 4 | IBM.2107-: | 2013-02-05 | P0 | 0 | 100 | 34.38 | 13.14 | 0.49638 | 1.35 | 24.06 | 2.13523 | 1.12 | 1.39 |
| 6 | 5 | IBM.2107-: | 2013-02-05 | P0 | 0 | 100 | 22.44 | 13.04 | 0.51686 | 2.58 | 31.99 | 3.0026 | 1.17 | 1.45 |
| 7 | 6 | IBM.2107-: | 2013-02-05 | P0 | 0 | 100 | 20.14 | 12.85 | 0.50222 | 2.49 | 32.11 | 3.00605 | 1.17 | 1.46 |
| 8 | 7 | IBM.2107-: | 2013-02-05 | P0 | 0 | 100 | 19.13 | 12.52 | 0.47954 | 2.2 | 31.22 | 2.88169 | 1.21 | 1.48 |
| 9 | 8 | IBM.2107-: | 2013-02-05 | P0 | 0 | 100 | 18.04 | 12.77 | 0.48626 | 2.26 | 32.32 | 2.99828 | 1.16 | 1.45 |
| 10 | 9 | IBM.2107-: | 2013-02-05 | P0 | 0 | 100 | 17.9 | 12.19 | 0.46476 | 2.04 | 30.7 | 2.80839 | 1.18 | 1.46 |
| 11 | 10 | IBM.2107-: | 2013-02-05 | P0 | 0 | 100 | 16.99 | 12.45 | 0.47515 | 2.3 | 31.77 | 2.97877 | 1.21 | 1.47 |
| 12 | 11 | IBM.2107-: | 2013-02-05 | P0 | 0 | 100 | 16.74 | 12.09 | 0.44699 | 1.66 | 30.06 | 2.68922 | 1.21 | 1.45 |
| 13 | 12 | IBM.2107-: | 2013-02-05 | P0 | 0 | 100 | 16.24 | 11.98 | 0.4461 | 1.73 | 30.13 | 2.70722 | 1.17 | 1.44 |
| 14 | 13 | IBM.2107-: | 2013-02-05 | P0 | 0 | 100 | 15.65 | 12.01 | 0.44234 | 1.69 | 30.03 | 2.69244 | 1.16 | 1.42 |
| 15 | 14 | IBM.2107-: | 2013-02-05 | P0 | 0 | 100 | 15.38 | 11.81 | 0.42553 | 1.36 | 28.87 | 2.5235 | 1.21 | 1.43 |
| 16 | 15 | IBM.2107-: | 2013-02-05 | P0 | 0 | 100 | 14.93 | 11.78 | 0.4221 | 1.38 | 29.1 | 2.54468 | 1.2 | 1.43 |
| 17 | 16 | IBM.2107-: | 2013-02-05 | P0 | 0 | 100 | 14.61 | 11.59 | 0.411 | 1.23 | 28.12 | 2.44028 | 1.22 | 1.45 |
| 18 | 17 | IBM.2107-: | 2013-02-05 | P0 | 0 | 100 | 14.53 | 11.21 | 0.40194 | 1.12 | 27.19 | 2.33602 | 1.21 | 1.41 |

*Figure 10-14 STAT Charting Utility Workload Skew report*

The STAT Charting Utility can be downloaded from the IBM support website:

http://www.ibm.com/support/techdocs/atsmastr.nsf/WebIndex/PRS5251

### 10.2.10 More information

For more information about planning and configuration considerations, best practices, and monitoring and measurement tools, see *IBM System Storage SAN Volume Controller and Storwize V7000 Best Practices and Performance Guidelines*, SG24-7521, and *Implementing IBM Easy Tier with IBM Real-time Compression*, TIPS1072.

## 10.3 Thin provisioning

In a shared storage environment, *thin provisioning* is a method for optimizing the usage of available storage. It relies on allocating blocks of data on demand versus the traditional method of allocating all of the blocks up front. This methodology eliminates almost all white space, which helps avoid the poor usage rates (often as low as 10%) that occur in the traditional storage allocation method. Traditionally, large pools of storage capacity are allocated to individual servers but remain unused (not written to).

Thin provisioning presents more storage space to the hosts or servers that are connected to the storage system than is available on the storage system. The IBM Storwize V7000 supports this capability for Fibre Channel (FC) and Internet Small Computer System Interface (iSCSI) provisioned volumes.

An example of thin provisioning is when a storage system contains 5000 GiB of usable storage capacity, but the storage administrator mapped volumes of 500 GiB each to 15 hosts. In this example, the storage administrator makes 7500 GiB of storage space visible to the hosts, even though the storage system has only 5000 GiB of usable space, as shown in Figure 10-15. In this case, all 15 hosts cannot use immediately all 500 GiB that is provisioned to them. The storage administrator must monitor the system and add storage as needed.

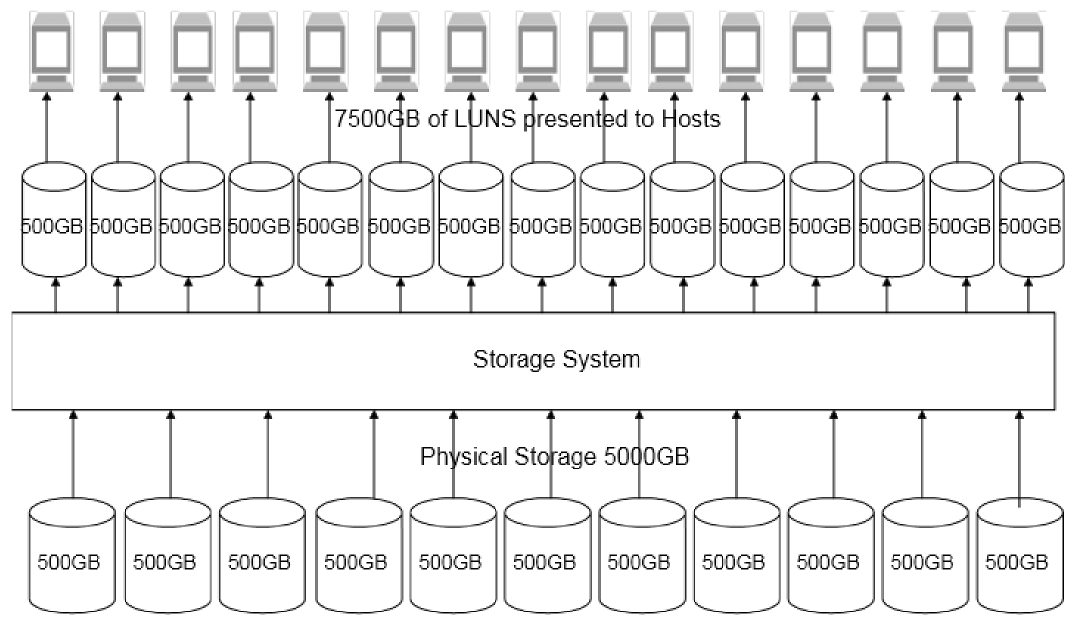


*Figure 10-15 Concept of thin provisioning*

You can imagine thin provisioning as the same process as when airlines sell more tickets on a flight than physical seats are available, assuming that some passengers do not appear at check-in. They do not assign actual seats at the time of sale, which avoids each client having a claim on a specific seat number. The same concept applies to thin provisioning (airline) IBM Storwize V7000 (plane) and its volumes (seats). The storage administrator (airline ticketing system) must closely monitor the allocation process and set proper thresholds.

## 10.3.1 Configuring a thin-provisioned volume

Volumes can be configured as *thin-provisioned* or *fully allocated*. Thin-provisioned volumes are created with real and virtual capacities. You can still create volumes by using a striped, sequential, or image mode virtualization policy, as you can with any other volume.

*Real capacity* defines how much disk space is allocated to a volume. *Virtual capacity* is the capacity of the volume that is reported to other IBM Storwize V7000 components (such as FlashCopy or remote copy) and to the hosts. For example, you can create a volume with real capacity of only 100 GiB but virtual capacity of 1 tebibyte (TiB). The actual space used by the volume on IBM Storwize V7000 is 100 GiB, but hosts see a 1 TiB volume.

A directory maps the virtual address space to the real address space. The directory and the user data share the real capacity.

Thin-provisioned volumes are available in two operating modes:

- Autoexpand
- Non-autoexpand

You can switch the mode at any time. If you select the autoexpand feature, the IBM Storwize V7000 automatically adds a fixed amount of more real capacity to the thin volume as required. Therefore, the autoexpand feature attempts to maintain a fixed amount of unused real capacity for the volume.

This amount is known as the *contingency capacity*. The contingency capacity is initially set to the real capacity that is assigned when the volume is created. If the user modifies the real capacity, the contingency capacity is reset to be the difference between the used capacity and real capacity.

A volume that is created without the autoexpand feature, and therefore has a zero contingency capacity, goes offline when the real capacity is used and the volume must expand.

**Warning threshold:** Enable the warning threshold, by using email or a Simple Network Management Protocol (SNMP) trap, when you work with thin-provisioned volumes. You can enable the warning threshold on the volume, and on the storage pool side, especially when you do not use the autoexpand mode. Otherwise, the thin volume goes offline if it runs out of space.

Autoexpand mode does not cause real capacity to grow much beyond the virtual capacity. The real capacity can be manually expanded to more than the maximum that is required by the current virtual capacity, and the contingency capacity is recalculated.

A thin-provisioned volume can be converted non-disruptively to a fully allocated volume, or vice versa, by using the volume mirroring function. For example, you can add a thin-provisioned copy to a fully allocated primary volume, and then remove the fully allocated copy from the volume after they are synchronized.

The fully allocated to thin-provisioned migration procedure uses a zero-detection algorithm, so that grains that contain all zeros do not cause any real capacity to be used. Usually, if IBM

Storwize V7000 is to detect zeros on the volume, you must use software on the host side to write zeros to all unused space on the disk or file system.

> **Tip:** Consider the use of thin-provisioned volumes as targets in the FlashCopy mappings.

## Space allocation

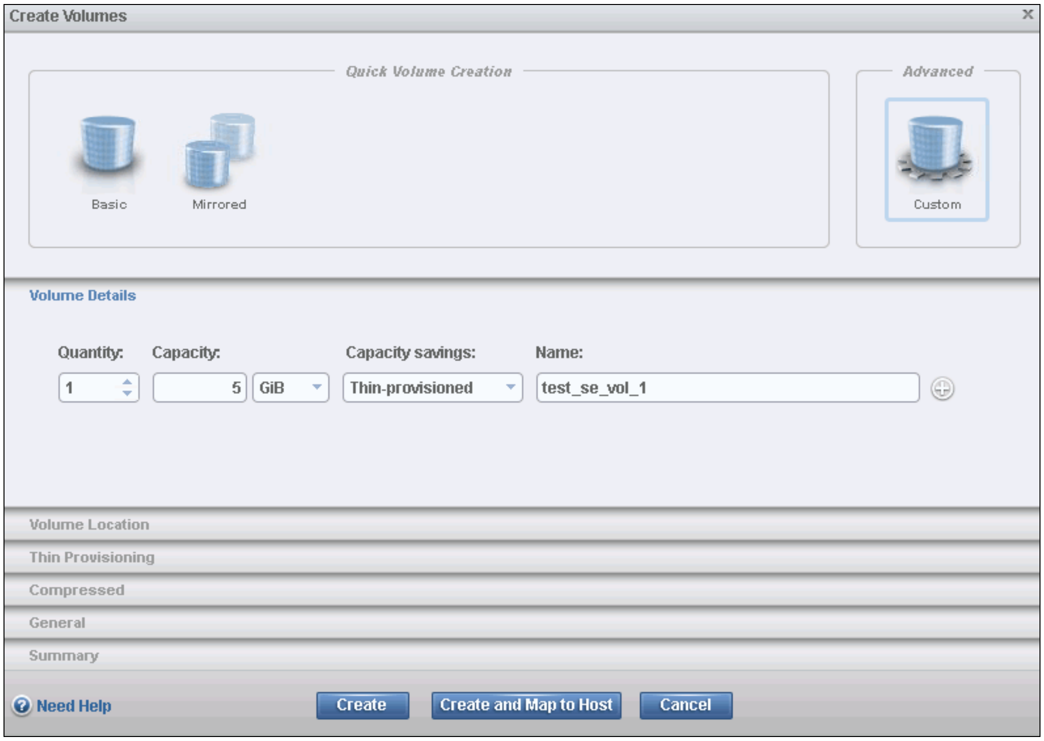
When a thin-provisioned volume is created, a small amount of the real capacity is used for initial metadata. Write I/Os to the grains of the thin volume (that were not previously written to) cause grains of the real capacity to be used to store metadata and user data. Write I/Os to the grains (that were previously written to) update the grain where data was previously written.

> **Grain definition:** The grain is defined when the volume is created, and can be 32 KiB, 64 KiB, 128 KiB, or 256 KiB.

Smaller granularities can save more space, but they have larger directories. When you use thin-provisioning with FlashCopy, specify the same grain size for the thin-provisioned volume and FlashCopy.

To create a thin-provisioned volume from the dynamic menu, complete the following steps:

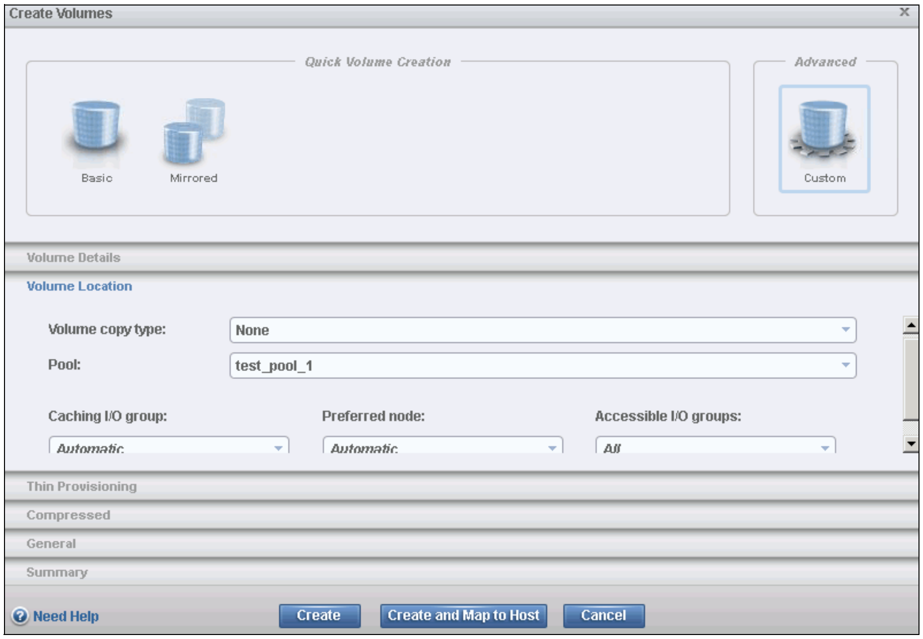
1. Go to **Volumes** → **Volumes** → **Create Volumes** and select **Advanced**, as shown in Figure 10-16. Enter the required capacity and volume name.



Quantity: 1
Capacity: 5 GiB
Capacity savings: Thin-provisioned
Name: test_se_vol_1

*Figure 10-16 Creating thin provisioned volume*

2. Select the volume size and name, and choose **Thin-provisioned** from the Capacity savings menu. If you want to create more volumes, click the Plus sign (+) next to the volume name. Click the **Volume Location** tab and select the storage pool for the volume.
3. If you have more than one I/O group, here you can also select the caching I/O group and preferred node, as shown in Figure 10-17.



Volume copy type: None

Pool: test_pool_1

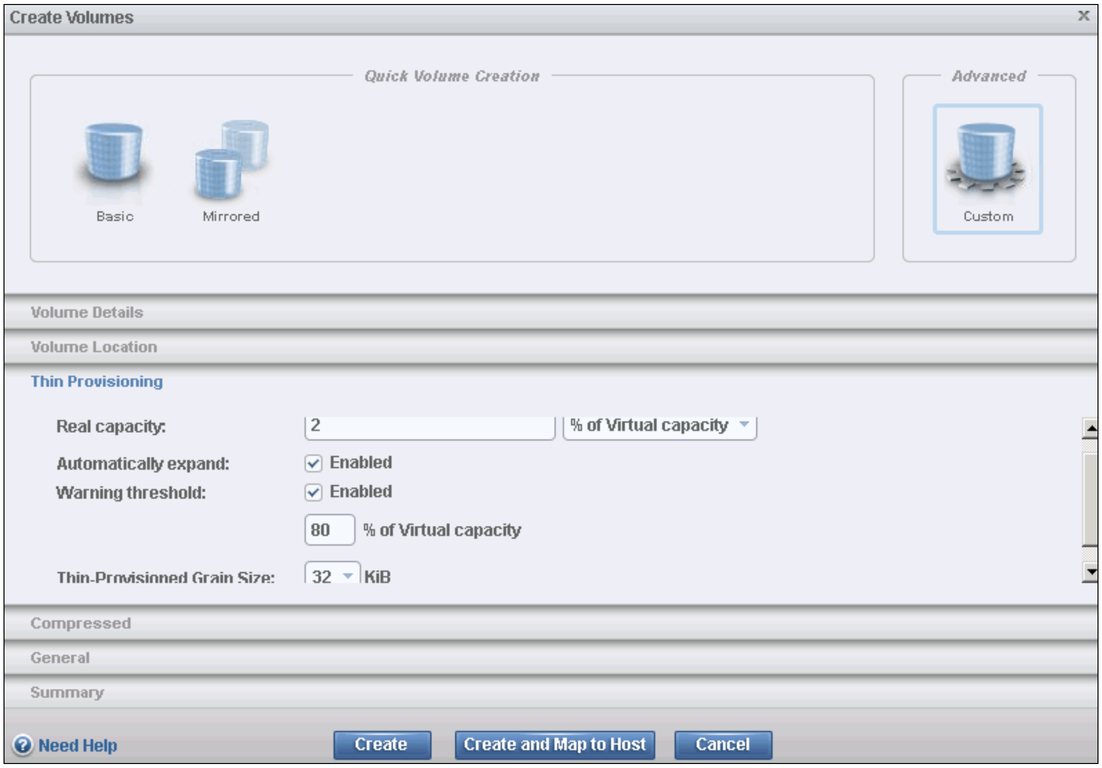
Caching I/O group: Automatic

Preferred node: Automatic

Accessible I/O groups: All

Figure 10-17 Choosing the thin provisioned volume location

4. Next go to *Thin Provisioning* tab and enter the thin provisioned parameters, such as real capacity, warning threshold, or autoexpand enabled or disabled, as seen in Figure 10-18.



Real capacity: 2 % of Virtual capacity

Automatically expand: ☑ Enabled

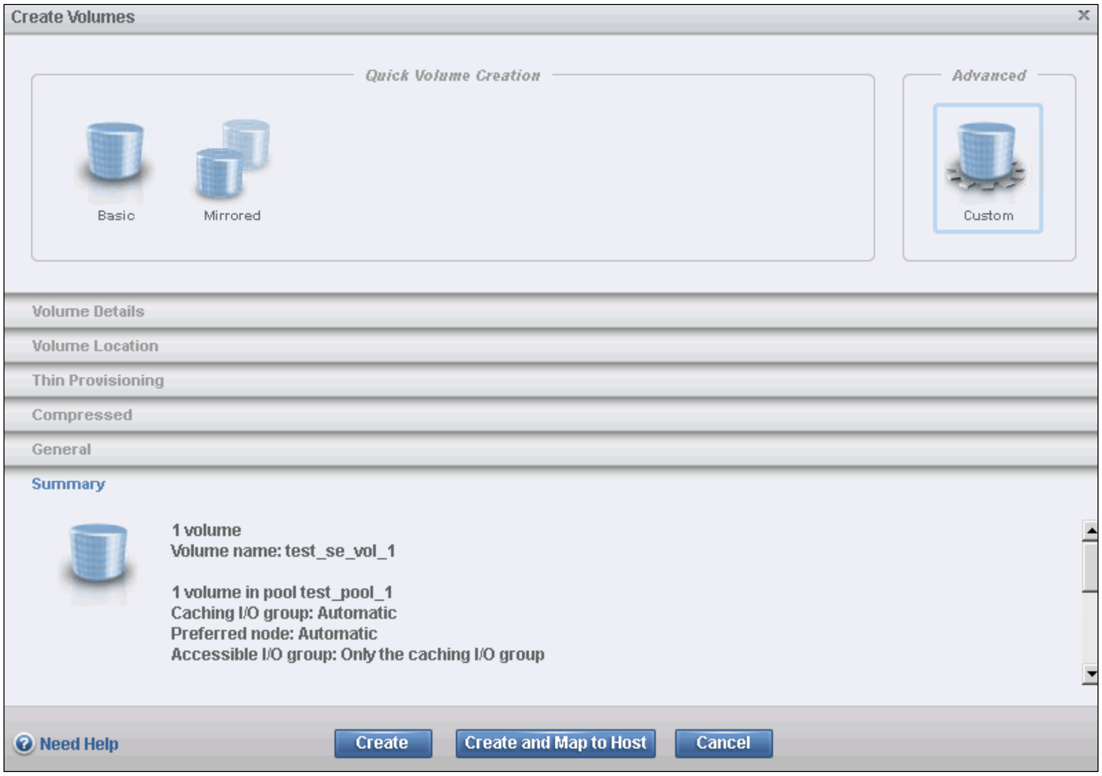
Warning threshold: ☑ Enabled

80 % of Virtual capacity

Thin-Provisioned Grain Size: 32 KiB

Figure 10-18 Choosing the thin provisioned parameters

5. Check your selections in the General tab and click **Create**, as shown in Figure 10-19.



1 volume
Volume name: test_se_vol_1

1 volume in pool test_pool_1
Caching I/O group: Automatic
Preferred node: Automatic
Accessible I/O group: Only the caching I/O group

*Figure 10-19 Creating thin provisioned volume summary*

## 10.3.2 Performance considerations

Thin-provisioned volumes save capacity only if the host server does not write to whole volumes. Whether the thin-provisioned volume works well partly depends on how the file system allocated the space. Some file systems, for example, New Technology File System (NTFS), write to the whole volume before overwriting deleted files. Other file systems reuse space in preference to allocating new space.

File system problems can be moderated by tools, such as `defrag`, or by managing storage by using host Logical Volume Managers (LVMs). The thin-provisioned volume also depends on how applications use the file system. For example, some applications delete log files only when the file system is nearly full.

> **Important:** Do not use defrag on thin-provisioned volumes. The defragmentation process can write data to different areas of a volume, which can cause a thin-provisioned volume to grow up to its virtual size.

There is no recommendation for thin-provisioned volumes. As explained previously, the performance of thin-provisioned volumes depends on what is used in the particular environment. For the best performance, use fully allocated volumes rather than thin-provisioned volumes.

> **Note:** Starting with IBM Spectrum Virtualize V7.3, the cache subsystem architecture was redesigned. Now, thin-provisioned volumes can benefit from lower cache functions (such as coalescing writes or prefetching), which greatly improve performance.

### 10.3.3 Limitations of virtual capacity

A few factors (extent and grain size) limit the virtual capacity of thin-provisioned volumes beyond the factors that limit the capacity of regular volumes. Table 10-2 shows the maximum thin provisioned volume virtual capacities for an extent size.

*Table 10-2 Maximum thin provisioned volume virtual capacities for an extent size*

| Extent size in megabytes (MB) | Maximum volume real capacity in gigabytes (GB) | Maximum thin virtual capacity in GB |
|---|---|---|
| 16 | 2,048 | 2,000 |
| 32 | 4,096 | 4,000 |
| 64 | 8,192 | 8,000 |
| 128 | 16,384 | 16,000 |
| 256 | 32,768 | 32,000 |
| 512 | 65,536 | 65,000 |
| 1,024 | 131,072 | 130,000 |
| 2,048 | 262,144 | 260,000 |
| 4,096 | 262,144 | 262,144 |
| 8,192 | 262,144 | 262,144 |

Table 10-3 shows the maximum thin-provisioned volume virtual capacities for a grain size.

*Table 10-3 Maximum thin volume virtual capacities for a grain size*

| Grain size in KiB | Maximum thin virtual capacity in GiB |
|---|---|
| 32 | 260,000 |
| 64 | 520,000 |
| 128 | 1,040,000 |
| 256 | 2,080,000 |

For more information and detailed performance considerations for configuring thin provisioning, see *IBM System Storage SAN Volume Controller and Storwize V7000 Best Practices and Performance Guidelines*, SG24-7521. You can also go to the V7000 IBM Knowledge Center:

http://www.ibm.com/support/knowledgecenter/ST3FR7

## 10.4 Real-time Compression software

The IBM Real-time Compression software that is embedded in the IBM Spectrum Virtualize addresses the requirements of primary storage data reduction, including performance. It does so by using a purpose-built technology, called *Real-time Compression*, that uses the Random Access Compression Engine (RACE). It offers the following benefits:

- Compression for active primary data

  IBM Real-time Compression can be used with active primary data. Therefore, it supports workloads that are not candidates for compression in other solutions. The solution supports online compression of existing data. Storage administrators can regain free disk space in an existing storage system without requiring administrators and users to clean up or archive data.

  This configuration significantly enhances the value of existing storage assets, and the benefits to the business are immediate. The capital expense of upgrading or expanding the storage system is delayed.

- Compression for replicated or mirrored data

  Remote volume copies can be compressed, in addition to the volumes at the primary storage tier. This process reduces storage requirements in Metro Mirror and Global Mirror destination volumes as well.

- No changes to the existing environment are required

  IBM Real-time Compression is part of the storage system. It was designed with transparency in mind so that it can be implemented without changes to applications, hosts, networks, fabrics, or external storage systems. The solution is not apparent to hosts, so users and applications continue to work as-is. Compression occurs within the IBM Spectrum Virtualize.

- Overall savings in operational expenses

  More data is stored in a rack space, so fewer storage expansion enclosures are required to store a data set. This reduced rack space has the following benefits:

  - Reduced power and cooling requirements. More data is stored in a system, which requires less power and cooling per gigabyte or used capacity.
  - Reduced software licensing for more functions in the system. More data that is stored per enclosure reduces the overall spending on licensing.

  > **Tip:** Implementing compression in IBM Spectrum Virtualize provides the same benefits to internal SSDs and externally virtualized storage systems.

- Disk space savings are immediate

  The space reduction occurs when the host writes the data. This process is unlike other compression solutions in which some or all of the reduction is realized only after a post-process compression batch job is run.

### 10.4.1 Common use cases

This section addresses the most common use cases for implementing compression:

- General-purpose volumes
- Databases
- Virtualized infrastructures
- Log server data stores

### General-purpose volumes

Most general-purpose volumes are used for highly compressible data types, such as home directories, CAD/CAM, oil and gas geo-seismic data, and log data. Storing such types of data in compressed volumes provides immediate capacity reduction to the overall used space. More space can be provided to users without any change to the environment.

There can be many file types stored in general-purpose servers. However, for practical information, the estimated compression ratios are based on actual field experience. Expected compression ratios are 50% to 60%.

File systems that contain audio, video files, and compressed files are not good candidates for compression. The overall capacity savings on these file types are minimal.

### Databases

Database information is stored in table space files. It is common to observe high compression ratios in database volumes. Examples of databases that can greatly benefit from Real-Time Compression are IBM DB2®, Oracle, and Microsoft SQL Server. Expected compression ratios are 50% to 80%.

> **Important:** Some databases offer optional built-in compression. Generally, do not compress already compressed database files.

### Virtualized infrastructures

The proliferation of open systems virtualization in the market has increased the use of storage space, with more virtual server images and backups kept online. The use of compression reduces the storage requirements at the source.

Examples of virtualization solutions that can greatly benefit from Real-time Compression are VMware, Microsoft Hyper-V, and KVM. Expected compression ratios are 45% to 75%.

> **Tip:** Virtual machines with file systems that contain compressed files are not good candidates for compression, as described previously in “Databases”.

### Log server data stores

Logs are a critical part for any information technology (IT) department in any organization. Log aggregates or syslog servers are a central point for the administrators, and immediate access and a smooth work process is necessary. Log server data stores are good candidates for Real-time Compression. Expected compression ratios are up to 90%.

## 10.4.2 Real-time Compression concepts

RACE technology is based on over 50 patents that are not primarily about compression. Instead, they define how to make industry-standard Lempel-Ziv (LZ) compression of primary storage operate in real-time and allow random access. The primary intellectual property behind this is the RACE engine.

At a high level, the IBM RACE component compresses data that is written into the storage system dynamically. This compression occurs transparently, so Fibre Channel and iSCSI connected hosts are not aware of the compression. RACE is an inline compression technology, meaning that each host write is compressed as it passes through the IBM Spectrum Virtualize to the disks. This has a clear benefit over other compression technologies that are post-processing based.

Those technologies do not provide immediate capacity savings. Therefore, they are not a good fit for primary storage workloads, such as databases and active data set applications.

RACE is based on the Lempel-Ziv lossless data compression algorithm and operates in a real-time method. When a host sends a write request, it is acknowledged by the write cache of the system, and then staged to the storage pool. As part of its staging, it passes through the compression engine and is then stored in compressed format onto the storage pool. Therefore, writes are acknowledged immediately after they are received by the write cache, with compression occurring as part of the staging to internal or external physical storage.

Capacity is saved when the data is written by the host because the host writes are smaller when they are written to the storage pool. IBM Real-time Compression is a self-tuning solution. It is adapting to the workload that runs on the system at any particular moment.

## 10.4.3 Random Access Compression Engine

To understand why RACE is unique, you need to review the traditional compression techniques. This description is not about the compression algorithm itself, that is, how the data structure is reduced in size mathematically. Rather, the description is about how the data is laid out within the resulting compressed output.

### Compression utilities

Compression is probably most known to users because of the widespread use of compression utilities, such as the WinZip and gzip utilities. At a high level, these utilities take a file as their input, and parse the data by using a sliding window technique. Repetitions of data are detected within the sliding window history, most often 32 KiB. Repetitions outside of the window cannot be referenced. Therefore, the file cannot be reduced in size unless data is repeated when the window “slides” to the next 32 KiB slot.

Figure 10-20 shows compression that uses a sliding window, where the first two repetitions of the string “ABCD” fall within the same compression window, and can therefore be compressed using the same dictionary. Note that the third repetition of the string falls outside of this window, and therefore cannot be compressed using the same compression dictionary as the first two repetitions, reducing the overall achieved compression ratio.

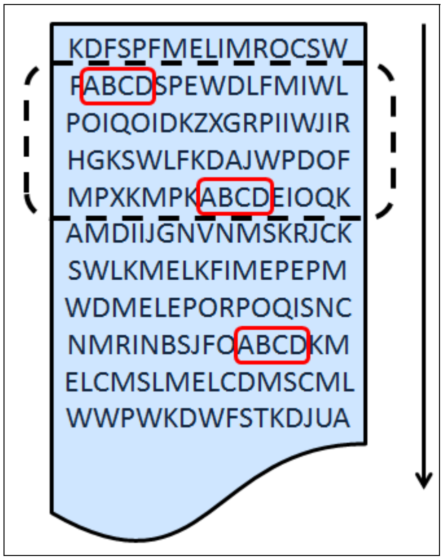


*Figure 10-20 Compression that uses a sliding window*

## Traditional data compression in storage systems

The traditional approach taken to implement data compression in storage systems is an extension of how compression works in the compression utilities previously mentioned. Similar to compression utilities, the incoming data is broken into fixed chunks, and then each chunk is compressed and extracted independently.

However, there are drawbacks to this approach. An update to a chunk requires a read of the chunk followed by a recompression of the chunk to include the update. The larger the chunk size chosen, the heavier the I/O penalty to recompress the chunk. If a small chunk size is chosen, the compression ratio is reduced because the repetition detection potential is reduced.

Figure 10-21 shows an example of how the data is broken into fixed-size chunks (in the upper-left side of the figure). It also shows how each chunk gets compressed independently into variable length compressed chunks (in the upper-right side of the figure). The resulting compressed chunks are stored sequentially in the compressed output.

Although this approach is an evolution from compression utilities, it is limited to low-performance use cases. This limitation is mainly because it does not provide real random access to the data.

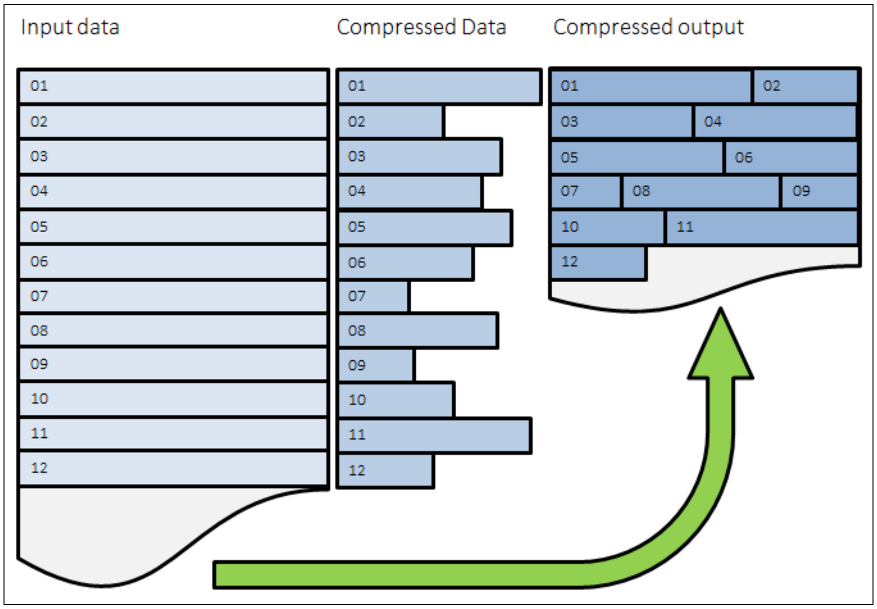


*Figure 10-21 Traditional data compression in storage systems*

## Random Access Compression Engine

The IBM patented Random Access Compression Engine implements an inverted approach when compared to traditional approaches to compression. RACE uses variable-size chunks for the input, and produces fixed-size chunks for the output.

This method enables an efficient and consistent way to index the compressed data because it is stored in fixed-size containers (Figure 10-22).

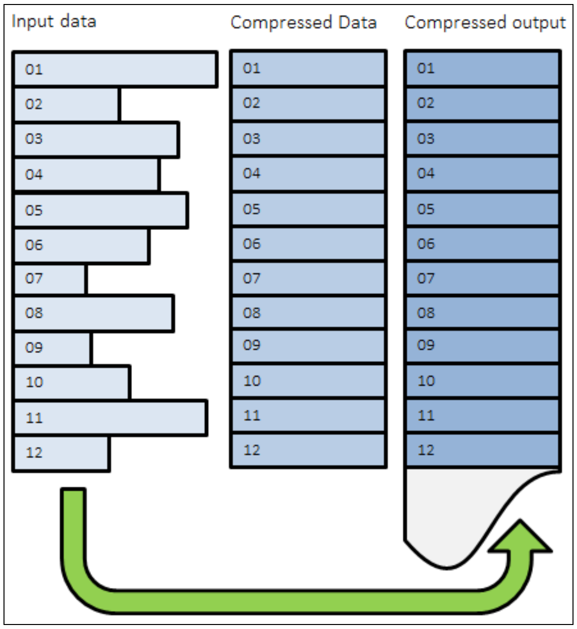


*Figure 10-22 Random Access Compression*

## Location-based compression

Both compression utilities and traditional storage systems compression compress data by finding repetitions of bytes within the chunk that is being compressed. The compression ratio of this chunk depends on how many repetitions can be detected within the chunk. The number of repetitions is affected by how much the bytes stored in the chunk are related to each other. The relation between bytes is driven by the format of the object. For example, an office document might contain textual information, and an embedded drawing (like this page).

Because the chunking of the file is arbitrary, it has no concept of how the data is laid out within the document. Therefore, a compressed chunk can be a mixture of the textual information and part of the drawing. This process yields a lower compression ratio, because the different data types mixed together cause a suboptimal dictionary of repetitions. That is, fewer repetitions can be detected because a repetition of bytes in a text object is unlikely to be found in a drawing.

This traditional approach to data compression is also called *location-based compression*. The data repetition detection is based on the location of data within the same chunk.

This challenge was also addressed with the *predecide* mechanism that was introduced in V7.1.

## Predecide mechanism

Some data chunks have a higher compression ratio than others. Compressing some of the chunks saves very little space, but still requires resources, such as processor (CPU) and memory. To avoid spending resources on uncompressible data, and to provide the ability to use a different, more effective (in this particular case) compression algorithm, IBM has invented a *predecide* mechanism that was first introduced in V7.1.

The chunks that are below a given compression ratio are skipped by the compression engine, therefore saving CPU time and memory processing. Chunks that are decided not to be compressed with the main compression algorithm, but that still can be compressed well with the other, are marked and processed accordingly. The result might vary because predecide does not check the entire block, only a sample of it.

Figure 10-23 shows how the detection mechanism works.

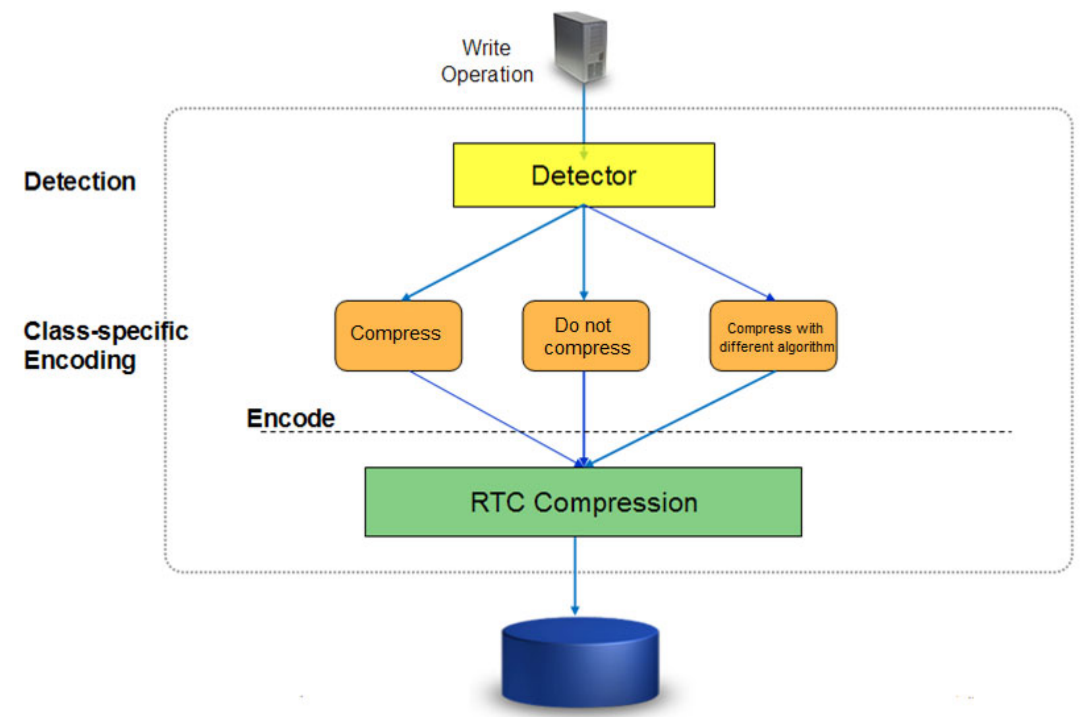


*Figure 10-23 Detection mechanism*

## Temporal compression

RACE offers a technology leap beyond location-based compression, called *temporal compression*. When host writes arrive to RACE, they are compressed and fill up fixed size chunks, also called *compressed blocks*. Multiple compressed writes can be aggregated into a single compressed block. A dictionary of the detected repetitions is stored within the compressed block.

When applications write new data or update existing data, it is typically sent from the host to the storage system as a series of writes. Because these writes are likely to originate from the same application and be from the same data type, more repetitions are usually detected by the compression algorithm. This type of data compression is called *temporal compression* because the data repetition detection is based on the time the data was written into the same compressed block.

Temporal compression adds the time dimension that is not available to other compression algorithms. It offers a higher compression ratio because the compressed data in a block represents a more homogeneous set of input data.

Figure 10-24 shows how three writes sent one after the other by a host end up in different chunks. They get compressed in different chunks because their location in the volume is not adjacent. This yields a lower compression ratio because the same data must be compressed non-natively by using three separate dictionaries.

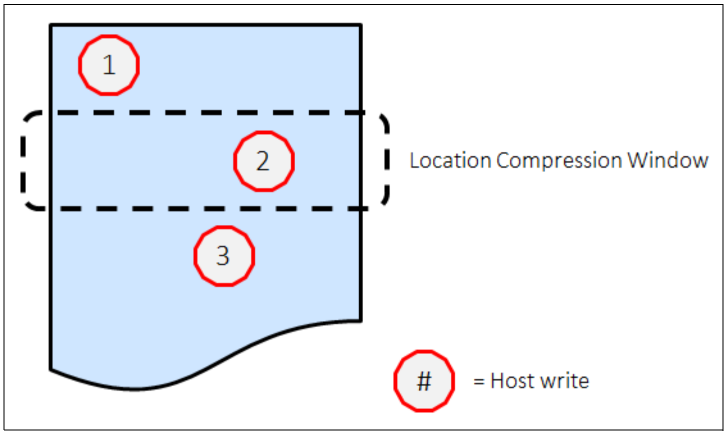


*Figure 10-24 Location-based compression*

When the same three writes are sent through RACE, as shown on Figure 10-25, the writes are compressed together by using a single dictionary. This yields a higher compression ratio than location-based compression (Figure 10-25).

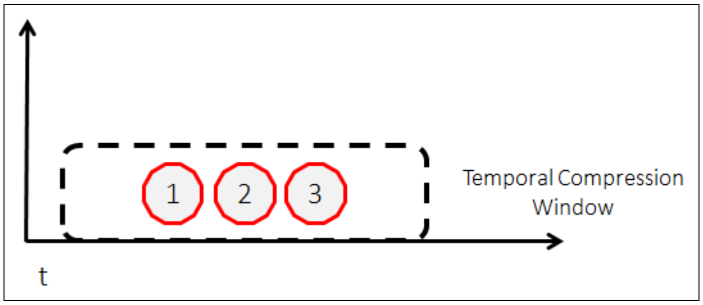


*Figure 10-25 Temporal compression*

## 10.4.4 Dual RACE component

In V7.4 compression code was enhanced by the addition of a second RACE component per node. This feature takes advantage of multi-core controller architecture, and uses the compression accelerator cards more effectively. The dual RACE component adds the second RACE instance, which works in parallel with the first instance, as shown in Figure 10-26.

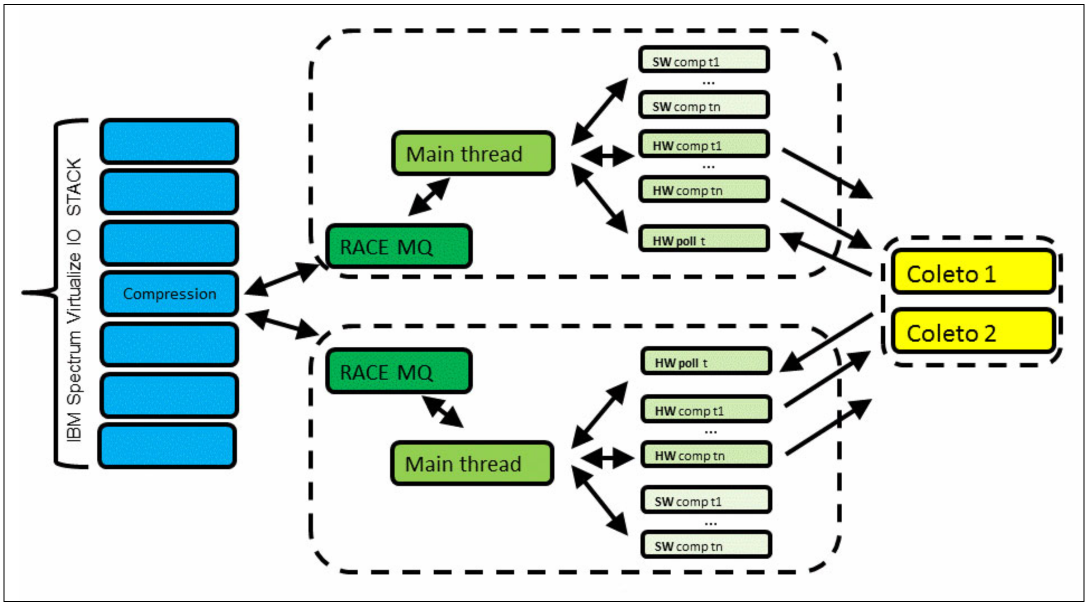


*Figure 10-26 Dual RACE architecture*

With dual RACE enhancement, the compression performance can be boosted up to two times for compressed workloads when compared to previous versions.

To take advantage of dual RACE, several software and hardware requirements must be met:

- The software must be at V7.4.
- Only Storwize V7000 Gen2 is supported.

**Note:** We advise using two acceleration cards for the best performance.

When using the dual RACE feature, the acceleration cards are shared between RACE components, which means that the acceleration cards are used simultaneously by both RACE components. The rest of resources, such as processor (CPU) cores and random access memory (RAM), are evenly divided between the RACE components.

You do not need to manually enable dual RACE; dual RACE triggers automatically when all minimal software and hardware requirements are met. If the Storwize V7000 Gen2 is compression capable but the minimal requirements for dual RACE are not met, only one RACE instance is used (as in the previous versions of the code).

## 10.4.5 Random Access Compression Engine in IBM Spectrum Virtualize stack

It is important to understand where the RACE technology is implemented in the IBM Spectrum Virtualize stack. This location determines how it applies to other Storwize components.

RACE technology is implemented into the Storwize thin provisioning layer, and is an organic part of the stack. The IBM Spectrum Virtualize stack is shown in Figure 10-27. Compression is transparently integrated with existing system management design. All of the IBM Spectrum Virtualize advanced features are supported on compressed volumes. You can create, delete, migrate, map (assign), and unmap (unassign) a compressed volume as though it were a fully allocated volume.

In addition, you can use Real-time Compression along with Easy Tier on the same volumes. This compression method provides nondisruptive conversion between compressed and decompressed volumes. This conversion provides a uniform user-experience and eliminates the need for special procedures when dealing with compressed volumes.

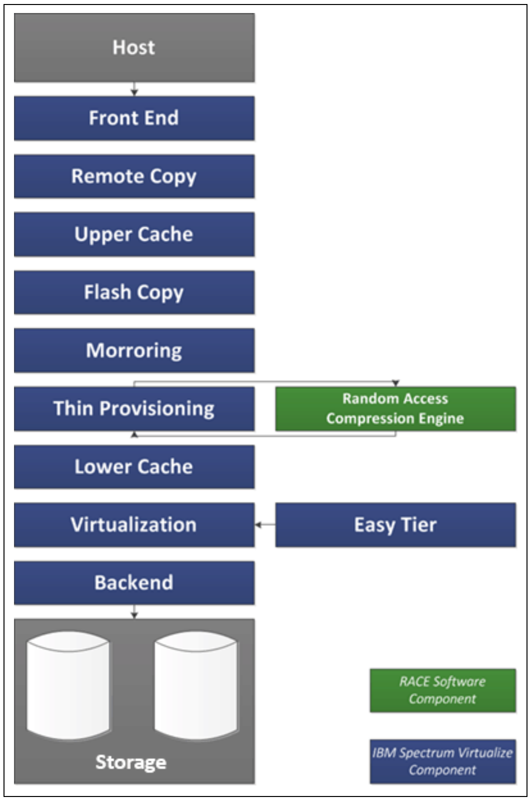


*Figure 10-27 RACE integration within IBM Spectrum Virtualize stack*

## 10.4.6 Data write flow

When a host sends a write request to Storwize V7000, it reaches the upper cache layer. The host is immediately sent an acknowledgment of its I/Os.

When the upper cache layer destages to the RACE, the I/Os are sent to the thin-provisioning layer. They are then sent to RACE, and if necessary, to the original host write or writes. The metadata that holds the index of the compressed volume is updated if needed, and is compressed as well.

## 10.4.7 Data read flow

When a host sends a read request to the Storwize V7000 for compressed data, it is forwarded directly to the Real-time Compression (RtC) component:

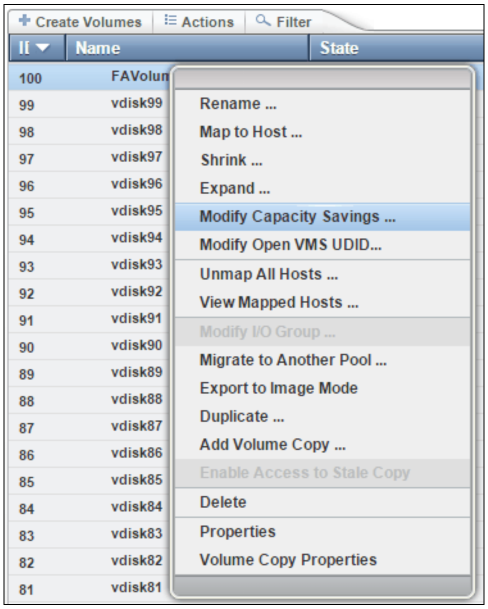
- If the RtC component contains the requested data, IBM Spectrum Virtualize cache replies to the host with the requested data without having to read the data from the lower-level cache or disk.
- If the RtC component does not contain the requested data, the request is forwarded to the Storwize IBM Spectrum Virtualize lower-level cache.
- If the lower-level cache contains the requested data, it is sent up the stack and returned to the host without accessing the storage.

- If the lower-level cache does not contain the requested data, it sends a read request to the storage for the requested data.

## 10.4.8 Compression of existing data

In addition to compressing data in real time, you can also compress existing data sets (convert volume to compressed). To do so, you have to change the capacity savings settings of the volume:

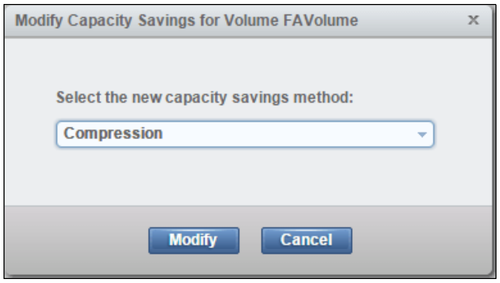
1. Right-click a particular volume and select **Modify Capacity Settings**, as shown in Figure 10-28.



| Il | Name | State |
|---|---|---|
| 100 | FAVolum | |
| 99 | vdisk99 | |
| 98 | vdisk98 | |
| 97 | vdisk97 | |
| 96 | vdisk96 | |
| 95 | vdisk95 | |
| 94 | vdisk94 | |
| 93 | vdisk93 | |
| 92 | vdisk92 | |
| 91 | vdisk91 | |
| 90 | vdisk90 | |
| 89 | vdisk89 | |
| 88 | vdisk88 | |
| 87 | vdisk87 | |
| 86 | vdisk86 | |
| 85 | vdisk85 | |
| 84 | vdisk84 | |
| 83 | vdisk83 | |
| 82 | vdisk82 | |
| 81 | vdisk81 | |

*Figure 10-28 Modifying Capacity Settings*

2. In the menu select **Compression** as the Capacity Savings option, as shown in Figure 10-29.



Select the new capacity savings method: Compression

*Figure 10-29 Selecting Capacity Setting*

3. After the copies are fully synchronized, the original volume copy is deleted automatically.

As a result, you have compressed data on the existing volume. This process is nondisruptive, so the data remains online and accessible by applications and users.

With virtualization of external storage systems, the ability to compress already stored data significantly enhances and accelerates the benefit to users. It enables them to see a tremendous return on their Storwize V7000 investment. On initial purchase of a Storwize

V7000 with Real-time Compression, customers can defer their purchase of new storage. As new storage needs to be acquired, IT purchases a lower amount of the required storage before compression.

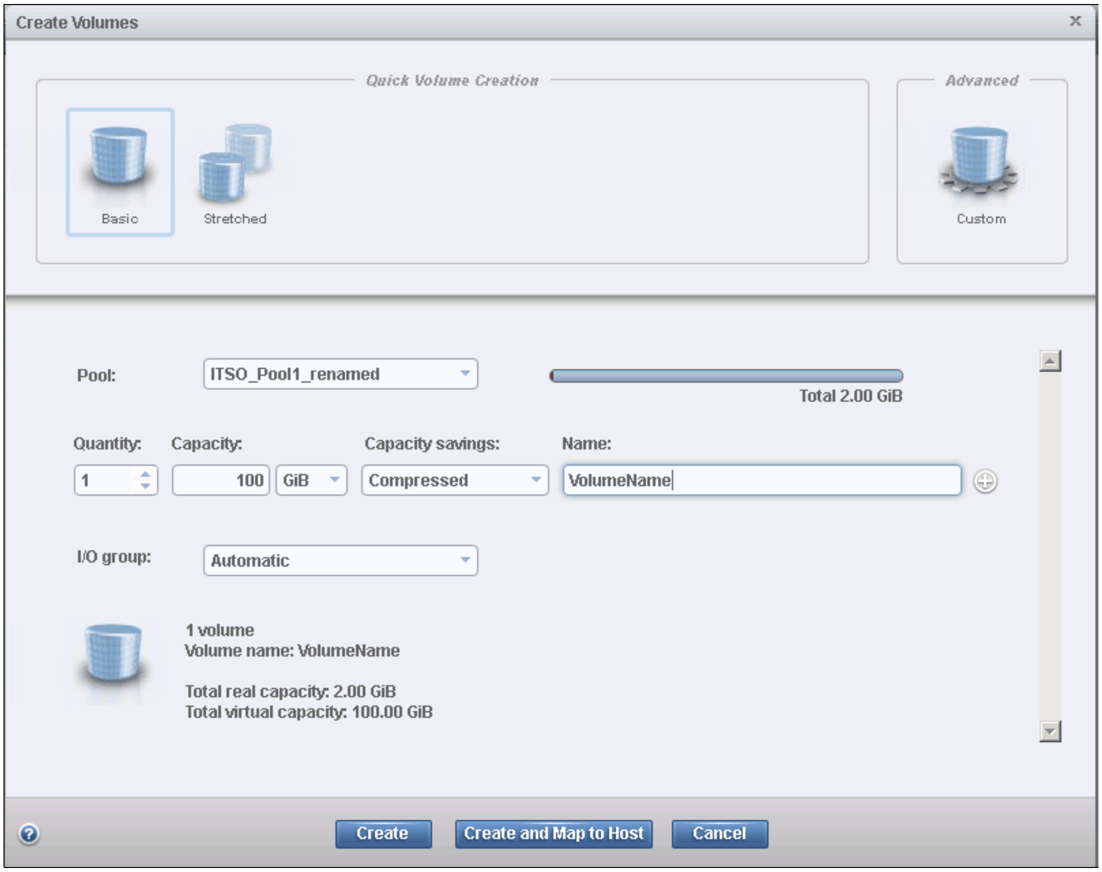
> **Important:** Remember that Storwize will reserve some of its resources like CPU cores and RAM after you create just one compressed volume or volume copy. This can affect your system performance if you do not plan accordingly in advance.

## 10.4.9 Configuring compressed volumes

To use compression on the Storwize V7000, licensing is required. With the Storwize V7000, Real-time Compression is licensed by capacity, per terabyte of virtual data.

There are two ways of creating a compressed volume: Basic and Advanced.

To create a compressed volume using Basic option, from the top bar under Volumes menu chose **Create Volumes** and select **Basic** in the **Quick Volume Creation** section, as shown in Figure 10-30.



Pool: ITSO_Pool1_renamed

Total 2.00 GiB

Quantity: 1

Capacity: 100 GiB

Capacity savings: Compressed

Name: VolumeName

I/O group: Automatic

1 volume
Volume name: VolumeName

Total real capacity: 2.00 GiB
Total virtual capacity: 100.00 GiB

*Figure 10-30 Creating Basic compressed volume*

To create a compressed volume using the Advanced option, complete the following steps:

1. From the top bar under the Volumes menu, choose **Create Volumes** and select **Custom** in the **Advanced** section, as shown in Figure 10-31.

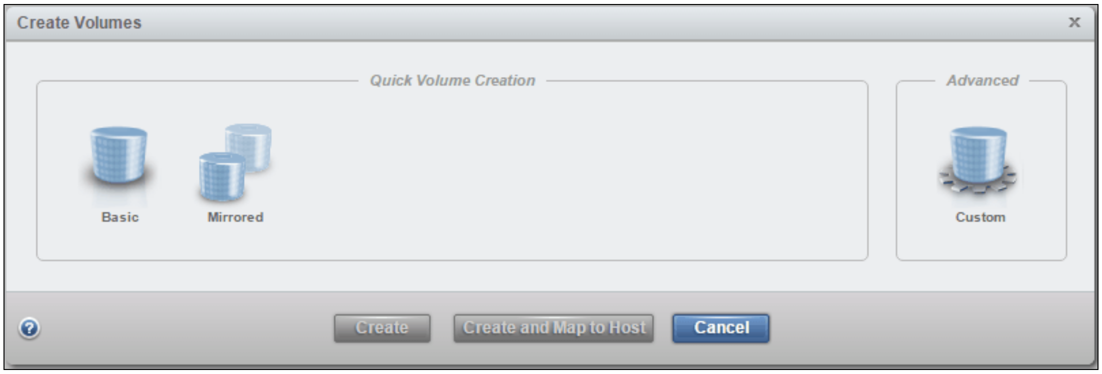
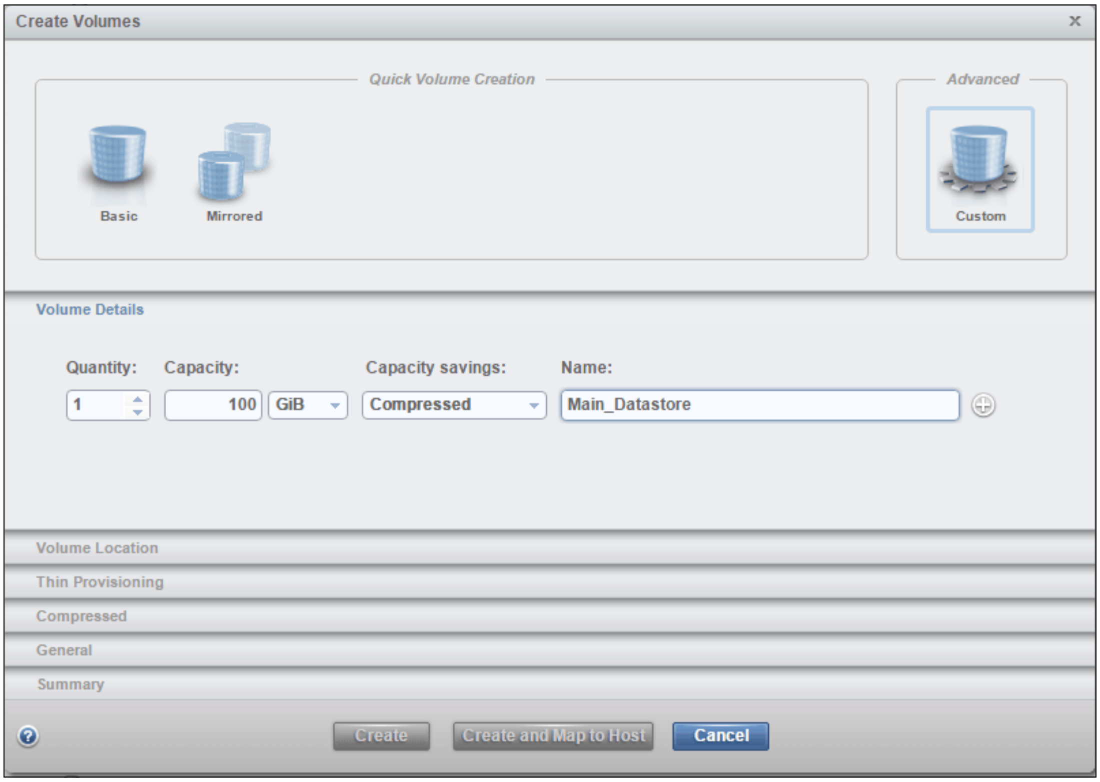


*Figure 10-31 Creating Advanced compressed volume*

2. In the **Volume Details** section, set up Capacity and Saving Type (Compression), and give the volume a Name, as shown in Figure 10-32.

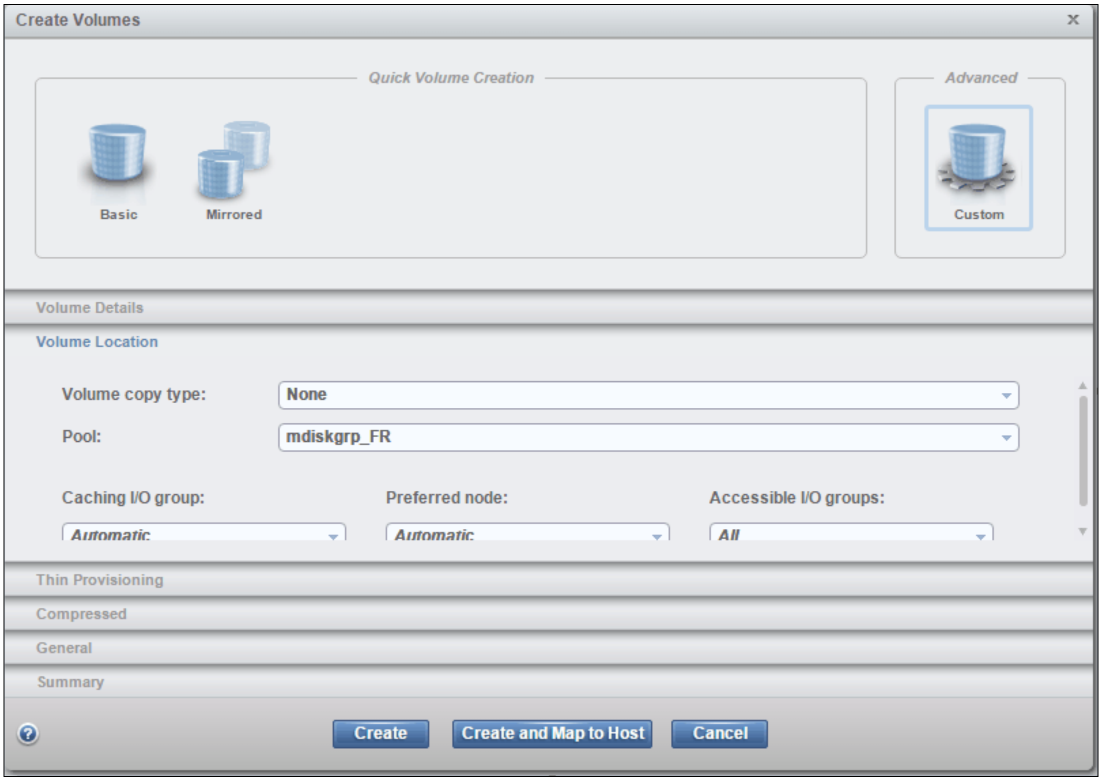


Volume Details

Quantity: 1
Capacity: 100 GiB
Capacity savings: Compressed
Name: Main_Datastore

*Figure 10-32 Setting up basic properties*

3. Setting location properties in the **Volume Location** section while setting **Pool** is required, as shown in Figure 10-33.



Volume copy type: None

Pool: mdiskgrp_FR

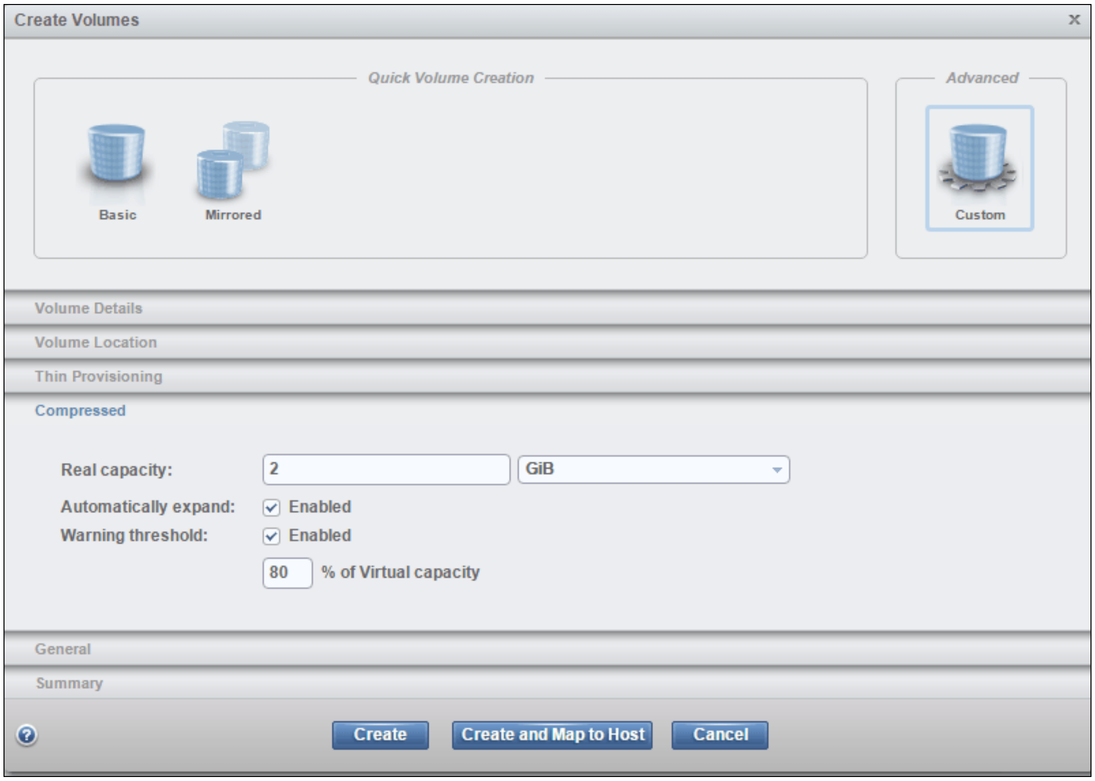
Caching I/O group: Automatic

Preferred node: Automatic

Accessible I/O groups: All

*Figure 10-33 Setting up location properties*

4. The **Compressed** section provides the ability to set or change allocated (virtual) capacity, the real capacity that data uses on this volume, autoexpand, and warning thresholds (Figure 10-34).



Real capacity: 2 GiB

Automatically expand: ☑ Enabled

Warning threshold: ☑ Enabled

80 % of Virtual capacity

*Figure 10-34 Setting up capacity properties of the compressed volume*

## 10.4.10 Comprestimator

IBM Spectrum Virtualize V7.6 introduced a utility to estimate expected compression ratios on existing volumes. V7.5 and V7.6 also includes a line of reliability, availability, and serviceability (RAS) improvements and features that helps IBM services to troubleshoot and monitor the client environment in a much better way.

The built-in Comprestimator is a command-line utility that can be used to estimate an expected compression rate for a given volume.

Comprestimator uses advanced mathematical and statistical algorithms to perform the sampling and analysis process in a very short and efficient way. The utility also displays its accuracy level by showing the maximum error range of the results achieved based on the formulas that it uses.

The following commands are available:

- The **`analyzevdisk`** command provides an option to analyze a single volume.

  Usage: `analyzevdisk <volume ID>`
  Example: **`analyzevdisk 0`**

  This command can be canceled by running the **`analyzevdisk <volume ID> -cancel`** command.

- The **`lsvdiskanalysis`** command provides a list and the status of the volumes. Some of them can be analyzed already, some of them not yet. The command can either be used for all volumes on the system or it can be used per volume, similar to **`lsvdisk`**. See Example 10-5.

*Example 10-5 Example of the command run over one volume with ID 0*

```
IBM_2076:ITSO Gen2:superuser>lsvdiskanalysis 0
id 0
name SQL_Data0
state estimated
started_time 151012104343
analysis_time 151012104353
capacity 300.00GB
thin_size 290.85GB
thin_savings 9.15GB
thin_savings_ratio 3.05
compressed_size 141.58GB
compression_savings 149.26GB
compression_savings_ratio 51.32
total_savings 158.42GB
total_savings_ratio 52.80
accuracy 4.97
```

The **`state`** parameter can have the following values:

  - `idle`. Was never estimated and not currently scheduled.
  - `scheduled`. Volume is queued for estimation, and will be processed based on lowest volume ID first.
  - `active`. Volume is being analyzed.
  - `canceling`. Volume was requested to cancel an active analysis, analysis was not yet canceled.
  - `estimated`. Volume was analyzed and results show the expected savings of thin provisioning and compression.
  - `sparse`. Volume was analyzed but comprestimator could not find enough non-zero samples to establish a good estimation.
  - `compression_savings_ratio`. Compression saving ratio is the estimated amount of space that can be saved on the storage in the frame of this specific volume expressed as a percentage.

- The **`analyzevdiskbysystem`** command provides an option to run Comprestimator on all volumes within the system. The analyzing process is nondisruptive and should not affect the system significantly. Analysis speed might vary due to the fullness of the volume, but should not take more than a few minutes per volume.

  This command can be canceled by running the **`analyzevdiskbysystem -cancel`** command.

- The **`lsvdiskanalysisprogress`** command shows the progress of the Comprestimator analysis as shown in Example 10-6.

*Example 10-6 Comprestimator progress*

```
id vdisk_count pending_analysis estimated_completion_time
0  45          12               151012154400
```

# 11

# Advanced Copy Services

This chapter describes the Advanced Copy Services and the storage software capabilities to support the interaction of your IBM Spectrum Virtualize with clouds. Both functions are enabled by IBM Spectrum Virtualize software running inside of IBM SAN Volume Controller and Storwize family products.

This chapter includes the following topics:

- FlashCopy
- Reverse FlashCopy
- FlashCopy functional overview
- Implementing FlashCopy
- Managing FlashCopy using GUI
- Transparent Cloud Tiering
- Implementing Transparent Cloud Tiering
- Volume mirroring and migration options
- Native IP replication
- Remote Copy
- Remote Copy commands
- Managing Remote Copy using the GUI
- Troubleshooting remote copy

# 11.1 FlashCopy

By using the IBM FlashCopy function of the IBM Spectrum Virtualize, you can perform a *point-in-time copy* of one or more volumes. In this section, we describe the inner workings of FlashCopy, and provide details of its configuration and use.

You can use FlashCopy to help you solve critical and challenging business needs that require duplication of data of your source volume. Volumes can remain online and active while you create consistent copies of the data sets. Because the copy is performed at the block level, it operates below the host operating system and its cache. Therefore, the copy is not apparent to the host.

> **Important:** Because FlashCopy operates at the block level below the host operating system and cache, those levels do need to be flushed for consistent FlashCopies.

While the FlashCopy operation is performed, the source volume is frozen briefly to initialize the FlashCopy bitmap, and then input/output (I/O) can resume. Although several FlashCopy options require the data to be copied from the source to the target in the background, which can take time to complete, the resulting data on the target volume is presented so that the copy appears to complete immediately.

This process is performed by using a bitmap (or bit array), which tracks changes to the data after the FlashCopy is started, and an indirection layer, which enables data to be read from the source volume transparently.

## 11.1.1 Business requirements for FlashCopy

When you are deciding whether FlashCopy addresses your needs, you must adopt a combined business and technical view of the problems that you want to solve. First, determine the needs from a business perspective. Then, determine whether FlashCopy can address the technical needs of those business requirements.

The business applications for FlashCopy are wide-ranging. Common use cases for FlashCopy include, but are not limited to, the following examples:

- Rapidly creating consistent backups of dynamically changing data
- Rapidly creating consistent copies of production data to facilitate data movement or migration between hosts
- Rapidly creating copies of production data sets for application development and testing
- Rapidly creating copies of production data sets for auditing purposes and data mining
- Rapidly creating copies of production data sets for quality assurance

Regardless of your business needs, FlashCopy within the IBM Spectrum Virtualize is flexible and offers a broad feature set, which makes it applicable to many scenarios.

## 11.1.2 Backup improvements with FlashCopy

FlashCopy does not reduce the time that it takes to perform a backup to traditional backup infrastructure. However, it can be used to minimize and, under certain conditions, eliminate application downtime that is associated with performing backups. FlashCopy can also transfer the resource usage of performing intensive backups from production systems.

After the FlashCopy is performed, the resulting image of the data can be backed up to tape, as though it were the source system. After the copy to tape is complete, the image data is redundant and the target volumes can be discarded. For time-limited applications, such as these examples, “no copy” or incremental FlashCopy is used most often. The use of these methods puts less load on your infrastructure.

When FlashCopy is used for backup purposes, the target data usually is managed as read-only at the operating system level. This approach provides extra security by ensuring that your target data was not modified and remains true to the source.

### 11.1.3 Restore with FlashCopy

FlashCopy can perform a restore from any existing FlashCopy mapping. Therefore, you can restore (or copy) from the target to the source of your regular FlashCopy relationships. When restoring data from FlashCopy, this method can be qualified as reversing the direction of the FlashCopy mappings.

This capability has the following benefits:

- There is no need to worry about pairing mistakes; you trigger a restore.
- The process appears instantaneous.
- You can maintain a pristine image of your data while you are restoring what was the primary data.

This approach can be used for various applications, such as recovering your production database application after an errant batch process that caused extensive damage.

> **Preferred practices:** Although restoring from a FlashCopy is quicker than a traditional tape media restore, you must not use restoring from a FlashCopy as a substitute for good archiving practices. Instead, keep one to several iterations of your FlashCopies so that you can near-instantly recover your data from the most recent history, and keep your long-term archive as appropriate for your business.

In addition to the restore option, which copies the original blocks from the target volume to modified blocks on the source volume, the target can be used to perform a restore of individual files. To do that you need to make the target available on a host. We suggest that you do not make the target available to the source host, because seeing duplicates of disks causes problems for most host operating systems. Copy the files to the source using normal host data copy methods for your environment.

### 11.1.4 Moving and migrating data with FlashCopy

FlashCopy can be used to facilitate the movement or migration of data between hosts while minimizing downtime for applications. By using FlashCopy, application data can be copied from source volumes to new target volumes while applications remain online. After the volumes are fully copied and synchronized, the application can be brought down and then immediately brought back up on the new server that is accessing the new FlashCopy target volumes.

This method differs from the other migration methods, which are described later in this chapter. Common uses for this capability are host and back-end storage hardware refreshes.

### 11.1.5 Application testing with FlashCopy

It is often important to test a new version of an application or operating system that is using actual production data. This testing ensures the highest quality possible for your environment. FlashCopy makes this type of testing easy to accomplish without putting the production data at risk or requiring downtime to create a constant copy.

You create a FlashCopy of your source and use that for your testing. This copy is a duplicate of your production data down to the block level so that even physical disk identifiers are copied. Therefore, it is impossible for your applications to tell the difference.

### 11.1.6 Host and application considerations to ensure FlashCopy integrity

Because FlashCopy is at the block level, it is necessary to understand the interaction between your application and the host operating system. From a logical standpoint, it is easiest to think of these objects as “layers” that sit on top of one another. The application is the topmost layer, and beneath it is the operating system layer.

Both of these layers have various levels and methods of caching data to provide better speed. Because the FlashCopy sit below these layers, they are unaware of the cache at the application or operating system layers.

To ensure the integrity of the copy that is made, it is necessary to flush the host operating system and application cache for any outstanding reads or writes before the FlashCopy operation is performed. Failing to flush the host operating system and application cache produces what is referred to as a *crash consistent* copy.

The resulting copy requires the same type of recovery procedure, such as log replay and file system checks, that is required following a host crash. FlashCopies that are crash consistent often can be used following file system and application recovery procedures.

Various operating systems and applications provide facilities to stop I/O operations and ensure that all data is flushed from host cache. If these facilities are available, they can be used to prepare for a FlashCopy operation. When this type of facility is unavailable, the host cache must be flushed manually by quiescing the application and unmounting the file system or drives.

**Preferred practice:** From a practical standpoint, when you have an application that is backed by a database and you want to make a FlashCopy of that application’s data, it is sufficient in most cases to use the write-suspend method that is available in most modern databases, because the database maintains strict control over I/O.

This method is as opposed to flushing data from both the application and the backing database, which is always the suggested method because it is safer. However, this method can be used when facilities do not exist or your environment includes time sensitivity.

### 11.1.7 FlashCopy attributes

The FlashCopy function in IBM Spectrum Virtualize features the following attributes:

- The target is the time-zero copy of the source, which is known as *FlashCopy mapping targets*.
- FlashCopy produces an exact copy of the source volume, including any metadata that was written by the host operating system, logical volume manager, and applications.

- The source volume and target volume are available (almost) immediately following the FlashCopy operation.
- The source and target volumes must be the same “virtual” size.
- The source and target volumes must be on the same IBM Storwize system.
- The source and target volumes do not need to be in the same I/O Group or storage pool.
- The storage pool extent sizes can differ between the source and target.
- The source volumes can have up to 256 target volumes (Multiple Target FlashCopy).
- The target volumes can be the source volumes for other FlashCopy relationships (*cascaded FlashCopy*).
- Consistency groups are supported to enable FlashCopy across multiple volumes at the same time.
- Up to 255 FlashCopy consistency groups are supported per system.
- Up to 512 FlashCopy mappings can be placed in one consistency group.
- The target volume can be updated independently of the source volume.
- Bitmaps that are governing I/O redirection (I/O indirection layer) are maintained in both node canisters of the IBM Storwize I/O Group to prevent a single point of failure.
- FlashCopy mapping and Consistency Groups can be automatically withdrawn after the completion of the background copy.
- Thin-provisioned FlashCopy (or Snapshot in the graphical user interface (GUI)) use disk space only when updates are made to the source or target data, and not for the entire capacity of a volume copy.
- FlashCopy licensing is based on the virtual capacity of the source volumes.
- Incremental FlashCopy copies all of the data when you first start FlashCopy and then only the changes when you stop and start FlashCopy mapping again. Incremental FlashCopy can substantially reduce the time that is required to re-create an independent image.
- Reverse FlashCopy enables FlashCopy targets to become restore points for the source without breaking the FlashCopy relationship, and without having to wait for the original copy operation to complete.
- The maximum number of supported FlashCopy mappings is 4096 per clustered system.
- The size of the source and target volumes cannot be altered (increased or decreased) while a FlashCopy mapping is defined.

## 11.1.8 Reverse FlashCopy

Reverse FlashCopy enables FlashCopy targets to become restore points for the source without breaking the FlashCopy relationship, and without having to wait for the original copy operation to complete. It supports multiple targets (up to 256) and therefore multiple rollback points.

A key advantage of the IBM Spectrum Virtualize Multiple Target Reverse FlashCopy function is that the reverse FlashCopy does not destroy the original target, which enables processes that are using the target, such as a tape backup, to continue uninterrupted.

IBM Spectrum Virtualize also provides the ability to create an optional copy of the source volume to be made before the reverse copy operation starts. This ability to restore back to the original source data can be useful for diagnostic purposes.

Complete the following steps to restore from an on-disk backup:

1. (Optional) Create a target volume (volume Z) and use FlashCopy to copy the production volume (volume X) onto the new target for later problem analysis.
2. Create a FlashCopy map with the backup to be restored (volume Y) or (volume W) as the source volume and volume X as the target volume, if this map does not exist.
3. Start the FlashCopy map (volume Y → volume X) with the `-restore` option to copy the backup data onto the production disk. If the `-restore` option is specified and no FlashCopy mapping exists, the command is ignored, which preserves your data integrity.

The production disk is instantly available with the backup data. Figure 11-1 shows an example of Reverse FlashCopy.

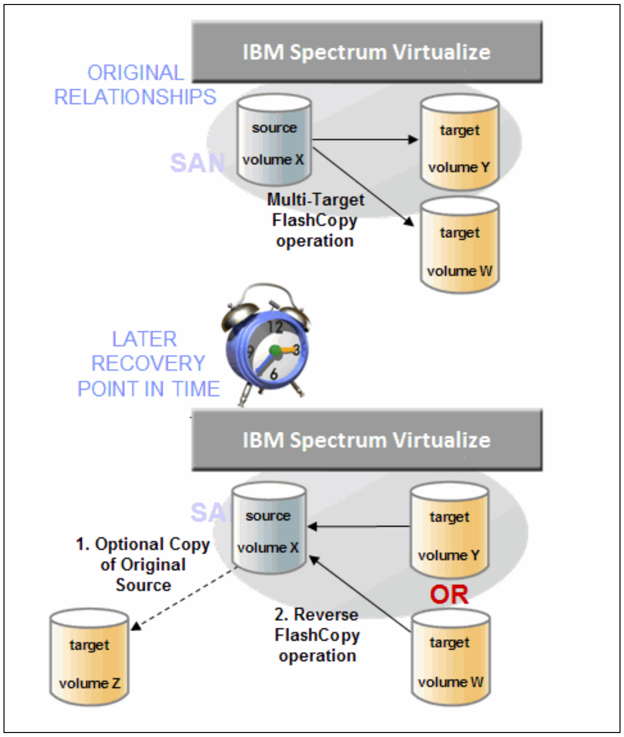


*Figure 11-1 Reverse FlashCopy*

Regardless of whether the initial FlashCopy map (volume X → volume Y) is incremental, the Reverse FlashCopy operation copies the modified data only.

Consistency Groups are reversed by creating a set of new reverse FlashCopy maps and adding them to a new reverse Consistency Group. Consistency Groups cannot contain more than one FlashCopy map with the same target volume.

### 11.1.9 IBM Spectrum Protect Snapshot

The management of many large FlashCopy relationships and Consistency Groups is a complex task without a form of automation for assistance.

IBM Spectrum Protect Snapshot (formerly Tivoli Flash Copy Manager) provides fast application-aware backups and restores using advanced point-in-time image technologies in the IBM Spectrum Virtualize. In addition, it allows you to manage frequent, near-instant, non-disruptive, application-aware backups and restores using integrated application and

VMware snapshot technologies. IBM Spectrum Protect Snapshot can be widely used in both IBM and non-IBM storage systems.

For more information about IBM Spectrum Protect Snapshot, see the following website:

https://ibm.biz/Bdsjv7

## 11.2 FlashCopy functional overview

FlashCopy works by defining a FlashCopy mapping that consists of one source volume with one target volume. Multiple FlashCopy mappings (source-to-target relationships) can be defined, and point-in-time consistency can be maintained across multiple individual mappings by using Consistency Groups. For more information, see “Consistency Group with Multiple Target FlashCopy” on page 423.

Before you start a FlashCopy (regardless of the type and options specified), you must issue a **prestartfcmap** or **prestartfcconsistgrp** command, which puts the cache into write-through mode and provides a flushing of the I/O currently bound for your volume. After FlashCopy is started, an effective copy of a source volume to a target volume is created.

The content of the source volume is presented immediately on the target volume and the original content of the target volume is lost. This FlashCopy operation is also referred to as a *time-zero copy* (T0).

> **Tip:** Rather than using **prestartfcmap** or **prestartfcconsistgrp**, you can also use the **-prep** parameter in the **startfcmap** or **startfcconsistgrp** command to prepare and start FlashCopy in one step.

The source and target volumes are available for use immediately after the FlashCopy operation. The FlashCopy operation creates a bitmap that is referenced and maintained to direct I/O requests within the source and target relationship. This bitmap is updated to reflect the active block locations as data is copied in the background from the source to the target, and updates are made to the source.

For more information about background copy, see Grains and the FlashCopy bitmap. Figure 11-2 on page 420 shows the redirection of the host I/O toward the source volume and the target volume.

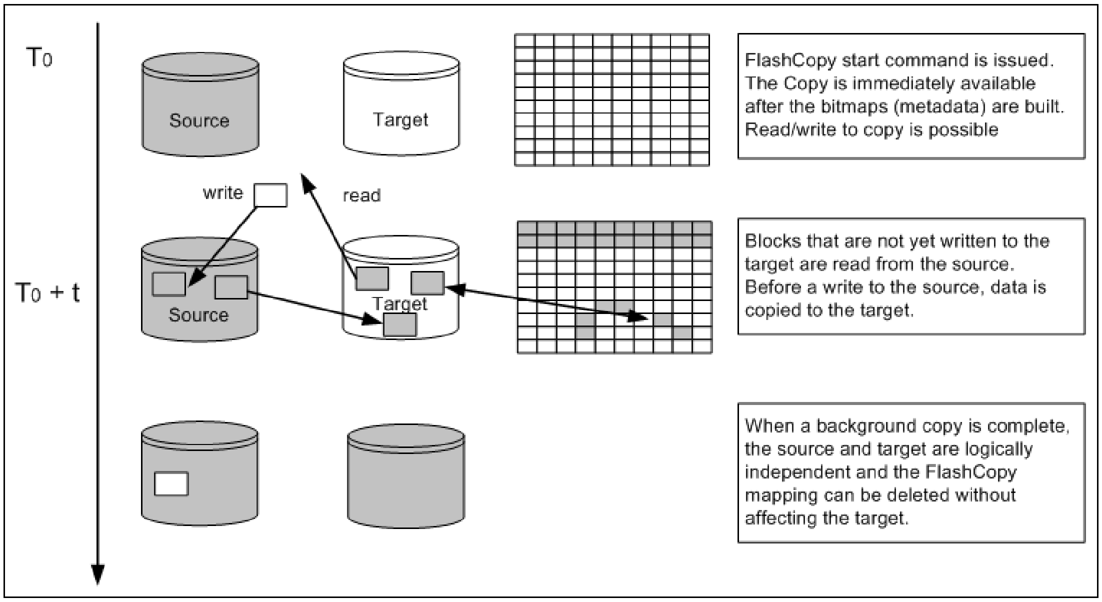

T0
Source
Target
write
read
T0 + t
Source
Target
FlashCopy start command is issued. The Copy is immediately available after the bitmaps (metadata) are built. Read/write to copy is possible
Blocks that are not yet written to the target are read from the source. Before a write to the source, data is copied to the target.
When a background copy is complete, the source and target are logically independent and the FlashCopy mapping can be deleted without affecting the target.

*Figure 11-2 Redirection of host I/O*

# 11.3 Implementing FlashCopy

In this section, we describe how FlashCopy is implemented in the IBM Spectrum Virtualize running on IBM Storwize.

## 11.3.1 FlashCopy mappings

FlashCopy occurs between a source volume and a target volume. The source and target volumes must be the same size. The minimum granularity that IBM Spectrum Virtualize supports for FlashCopy is an entire volume. It is not possible to use FlashCopy to copy only part of a volume.

**Important:** As with any point-in-time copy technology, you are bound by operating system and application requirements for interdependent data and the restriction to an entire volume.

The source and target volumes must belong to the same IBM Storwize V7000 system, but they do not have to be in the same I/O Group or storage pool. FlashCopy associates a source volume to a target volume through FlashCopy mapping.

To become members of a FlashCopy mapping, source and target volumes must be the same size. Volumes that are members of a FlashCopy mapping cannot have their size increased or decreased while they are members of the FlashCopy mapping.

A *FlashCopy mapping* is the act of creating a relationship between a source volume and a target volume. FlashCopy mappings can be stand-alone or a member of a Consistency Group. You can perform the actions of preparing, starting, or stopping FlashCopy on either a stand-alone mapping or a Consistency Group.

Figure 11-3 on page 421 shows the concept of FlashCopy mapping.

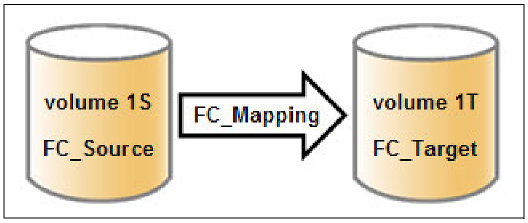


*Figure 11-3 FlashCopy mapping*

## 11.3.2 Multiple Target FlashCopy

The IBM Storwize V7000 supports up to 256 target volumes from a single source volume. Each copy is managed by a unique mapping. Figure 11-4 shows the Multiple Target FlashCopy implementation.

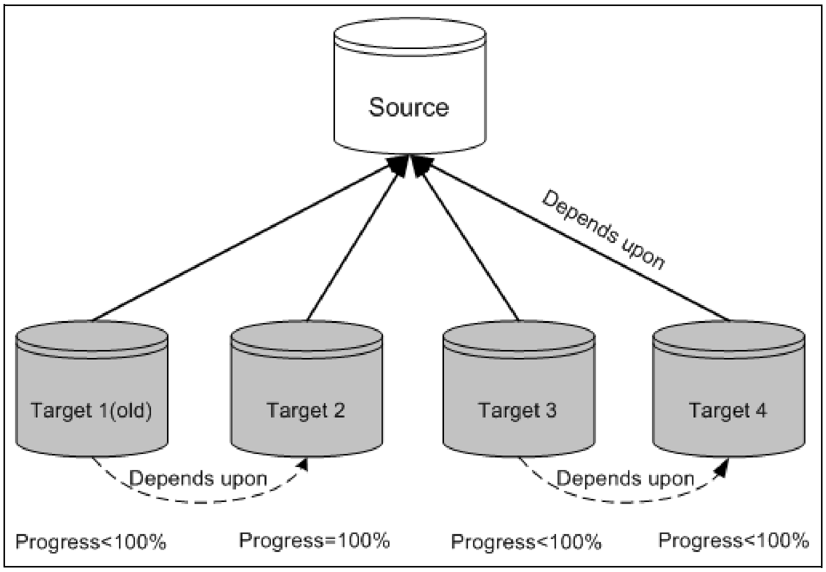


*Figure 11-4 Multiple Target FlashCopy implementation*

Figure 11-4 also shows four targets and mappings that are taken from a single source, along with their interdependencies. In this example, Target 1 is the oldest (as measured from the time that it was started) through to Target 4, which is the newest. The ordering is important because of how the data is copied when multiple target volumes are defined and because of the dependency chain that results.

A write to the source volume does not cause its data to be copied to all of the targets. Instead, it is copied to the newest target volume only (Target 4 in Figure 11-4). The older targets refer to new targets first before referring to the source.

From the point of view of an intermediate target disk (neither the oldest nor the newest), it treats the set of newer target volumes and the true source volume as a type of composite source. It treats all older volumes as a kind of target (and behaves like a source to them).

If the mapping for an intermediate target volume shows 100% progress, its target volume contains a complete set of data. In this case, mappings treat the set of newer target volumes (up to and including the 100% progress target) as a form of composite source. A dependency relationship exists between a particular target and all newer targets (up to and including a target that shows 100% progress) that share the source until all data is copied to this target and all older targets.

For more information about Multiple Target FlashCopy, see Interaction and dependency between multiple target FlashCopy mappings.

## 11.3.3 Consistency Groups

*Consistency Groups* address the requirement to preserve point-in-time data consistency across multiple volumes for applications that include related data that spans multiple volumes. For these volumes, Consistency Groups maintain the integrity of the FlashCopy by ensuring that “dependent writes” are run in the application’s intended sequence.

When Consistency Groups are used, the FlashCopy commands are issued to the FlashCopy Consistency Group, which performs the operation on all FlashCopy mappings that are contained within the Consistency Group at the same time.

Figure 11-5 shows a Consistency Group that includes two FlashCopy mappings.

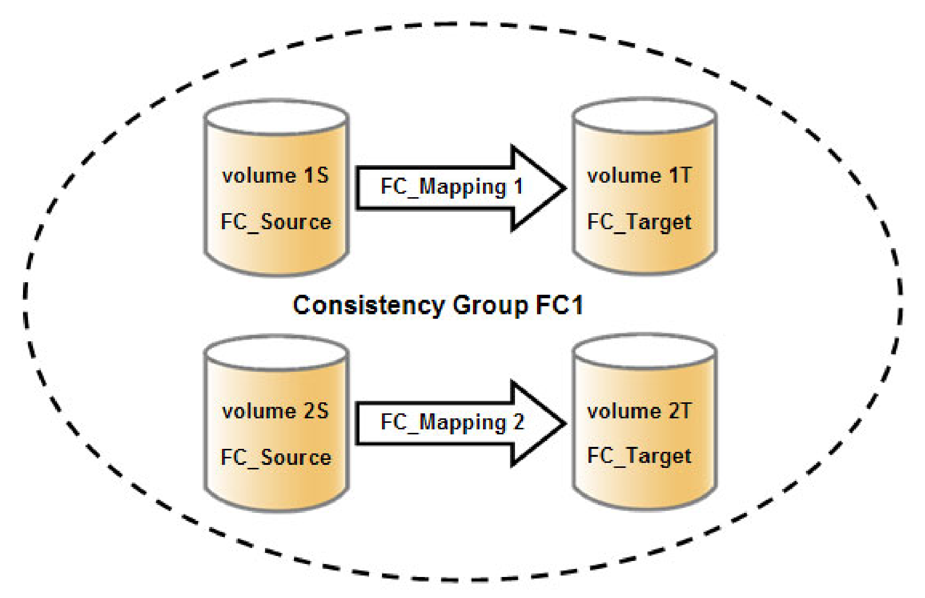


*Figure 11-5 FlashCopy Consistency Group*

> **Important:** After an individual FlashCopy mapping is added to a Consistency Group, it can be managed as part of the group only. Operations, such as `prepare`, `start`, and `stop`, are no longer allowed on the individual mapping.

### Dependent writes

To show why it is crucial to use Consistency Groups when a data set spans multiple volumes, consider the following typical sequence of writes for a database update transaction:

1. A write is run to update the database log, which indicates that a database update is about to be performed.
2. A second write is run to perform the actual update to the database.
3. A third write is run to update the database log, which indicates that the database update completed successfully.

The database ensures the correct ordering of these writes by waiting for each step to complete before the next step is started. However, if the database log (updates 1 and 3) and the database (update 2) are on separate volumes, it is possible for the FlashCopy of the database volume to occur before the FlashCopy of the database log. This sequence can result in the target volumes seeing writes 1 and 3 but not 2 because the FlashCopy of the database volume occurred before the write was completed.

In this case, if the database was restarted by using the backup that was made from the FlashCopy target volumes, the database log indicates that the transaction completed successfully. In fact, it did not complete successfully because the FlashCopy of the volume with the database file was started (the bitmap was created) before the write completed to the volume. Therefore, the transaction is lost and the integrity of the database is in question.

To overcome the issue of dependent writes across volumes and to create a consistent image of the client data, a FlashCopy operation must be performed on multiple volumes as an atomic operation. To accomplish this method, the IBM Spectrum Virtualize supports the concept of *Consistency Groups*.

A FlashCopy Consistency Group can contain up to 512 FlashCopy mappings. The maximum number of FlashCopy mappings that is supported by the IBM Storwize V7000 V7.8 is 4096. FlashCopy commands can then be issued to the FlashCopy Consistency Group and therefore, simultaneously for all of the FlashCopy mappings that are defined in the Consistency Group.

For example, when a FlashCopy **`start`** command is issued to the Consistency Group, all of the FlashCopy mappings in the Consistency Group are started at the same time. This simultaneous start results in a point-in-time copy that is consistent across all of the FlashCopy mappings that are contained in the Consistency Group.

## Consistency Group with Multiple Target FlashCopy

A Consistency Group aggregates FlashCopy mappings, not volumes. Therefore, where a source volume has multiple FlashCopy mappings, they can be in the same or separate Consistency Groups.

If a particular volume is the source volume for multiple FlashCopy mappings, you might want to create separate Consistency Groups to separate each mapping of the same source volume. Regardless of whether the source volume with multiple target volumes is in the same consistency group or in separate consistency groups, the resulting FlashCopy produces multiple identical copies of the source data.

## Maximum configurations

Table 11-1 lists the FlashCopy properties and maximum configurations.

*Table 11-1 FlashCopy properties and maximum configurations*

| FlashCopy property | Maximum | Comment |
|---|---:|---|
| FlashCopy targets per source | 256 | This maximum is the number of FlashCopy mappings that can exist with the same source volume. |
| FlashCopy mappings per system | 4096 | The number of mappings is no longer limited by the number of volumes in the system, so the FlashCopy component limit applies. |
| FlashCopy Consistency Groups per system | 255 | This maximum is an arbitrary limit that is policed by the software. |

| FlashCopy property | Maximum | Comment |
|---|---|---|
| FlashCopy volume capacity per I/O Group | 4 pebibytes (PiB) | This maximum is a limit on the quantity of FlashCopy mappings that are using bitmap space from this I/O Group. This maximum configuration uses all 4 gibibytes (GiB) of bitmap space for the I/O Group, and allows no Metro or Global Mirror bitmap space. The default is 40 tebibytes (TiB). |
| FlashCopy mappings per Consistency Group | 512 | This limit is because of the time that is taken to prepare a Consistency Group with many mappings. |

## 11.3.4 FlashCopy indirection layer

The *FlashCopy indirection layer* governs the I/O to the source and target volumes when a FlashCopy mapping is started, which is done by using a FlashCopy bitmap. The purpose of the FlashCopy indirection layer is to enable the source and target volumes for read and write I/O immediately after the FlashCopy is started.

To show how the FlashCopy indirection layer works, we examine what happens when a FlashCopy mapping is prepared and then started.

When a FlashCopy mapping is prepared and started, the following sequence is applied:

1. Flush the write cache to the source volume or volumes that are part of a Consistency Group.
2. Put cache into write-through mode on the source volumes.
3. Discard cache for the target volumes.
4. Establish a sync point on all of the source volumes in the Consistency Group (which creates the FlashCopy bitmap).
5. Ensure that the indirection layer governs all of the I/O to the source volumes and target volumes.
6. Enable cache on the source volumes and target volumes.

FlashCopy provides the semantics of a point-in-time copy by using the indirection layer, which intercepts I/O that is directed at the source or target volumes. The act of starting a FlashCopy mapping causes this indirection layer to become active in the I/O path, which occurs automatically across all FlashCopy mappings in the Consistency Group. The indirection layer then determines how each I/O is to be routed, based on the following factors:

- The volume and the logical block address (LBA) to which the I/O is addressed
- Its direction (read or write)
- The state of an internal data structure, the FlashCopy bitmap

The indirection layer allows the I/O to go through to the underlying volume, redirects the I/O from the target volume to the source volume, or queues the I/O while it arranges for data to be copied from the source volume to the target volume. To explain in more detail which action is applied for each I/O, we first look at the FlashCopy bitmap.

## 11.3.5 Grains and the FlashCopy bitmap

When data is copied between volumes, it is copied in units of address space that are known as *grains*. Grains are units of data that are grouped to optimize the use of the bitmap that tracks changes to the data between the source and target volume. You can use 64 kibibytes (KiB) or 256 KiB grain sizes (256 KiB is the default). The FlashCopy bitmap contains 1 bit for each grain, and is used to show whether the source grain was copied to the target. The 64 KiB grain size uses bitmap space at a rate of four times the default 256 KiB size.

The FlashCopy bitmap dictates read and write behavior for the source and target volumes.

### Source reads

Reads are performed from the source volume, which is the same as for non-FlashCopy volumes.

### Source writes

Writes to the source cause one of the following actions:

- If the grain was not copied to the target yet, the grain is copied before the actual write is performed to the source. The bitmap is updated to indicate that this grain is already copied to the target.
- If the grain was already copied, the write is performed to the source as usual.

### Target reads

Reads are performed from the target if the grain was copied. Otherwise, the read is performed from the source and no copy is performed.

### Target writes

Writes to the target cause one of the following actions:

- If the grain was not copied from the source to the target, the grain is copied from the source to the target before the actual write is performed to the source. The bitmap is updated to indicate that this grain is already copied to the target.
- If the entire grain is being updated on the target, the target is marked as split with the source (if there is no I/O error during the write) and the write goes directly to the target.
- If the grain in question was already copied from the source to the target, the write goes directly to the target.

### The FlashCopy indirection layer algorithm

Imagine the FlashCopy indirection layer as the I/O traffic director when a FlashCopy mapping is active. The I/O is intercepted and handled according to whether it is directed at the source volume or at the target volume, depending on the nature of the I/O (read or write) and the state of the grain (whether it was copied).

Figure 11-6 shows how the background copy runs while I/Os are handled according to the indirection layer algorithm.

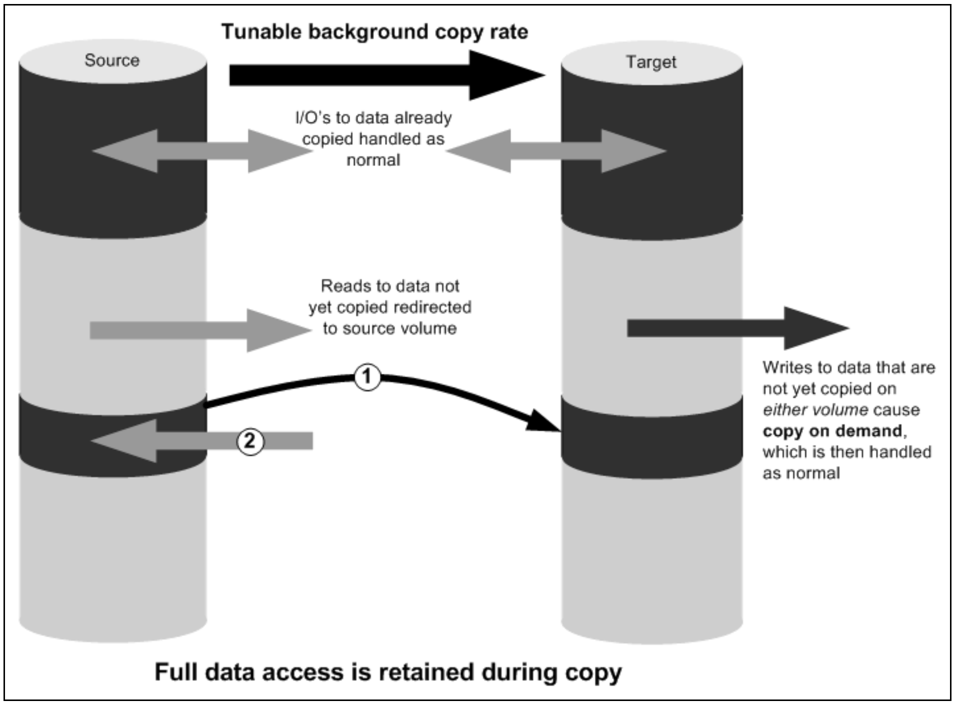

Tunable background copy rate
Source
Target
I/O's to data already copied handled as normal
Reads to data not yet copied redirected to source volume
1
2
Writes to data that are not yet copied on *either volume* cause **copy on demand**, which is then handled as normal
Full data access is retained during copy

*Figure 11-6 I/O processing with FlashCopy*

## 11.3.6 Interaction and dependency between multiple target FlashCopy mappings

Figure 11-7 shows a set of four FlashCopy mappings that share a common source. The FlashCopy mappings target volumes Target 0, Target 1, Target 2, and Target 3.

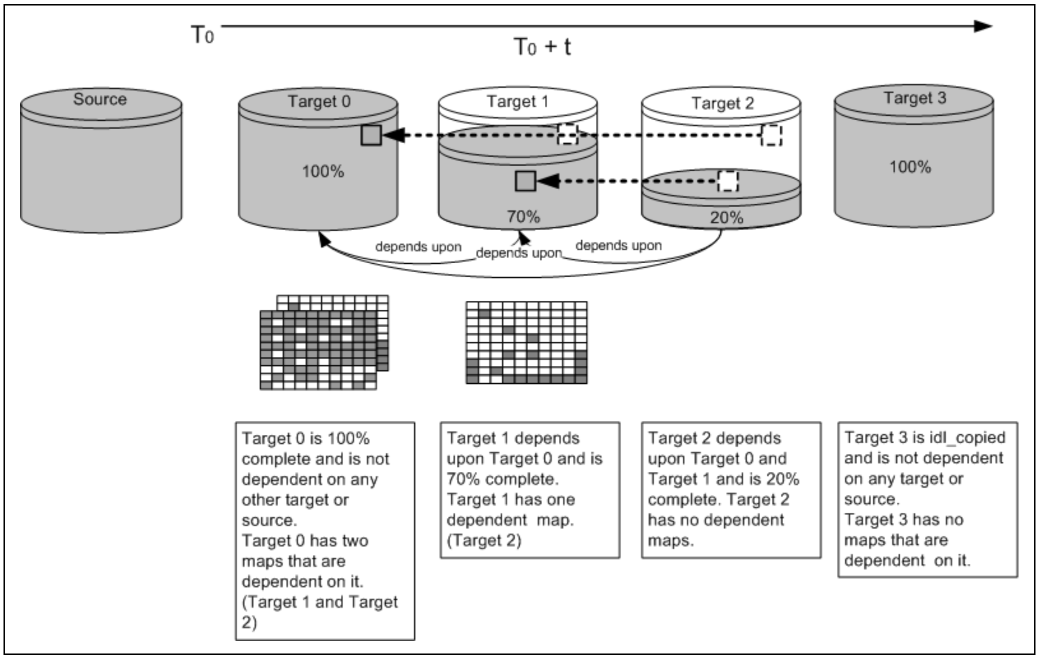

To
To + t
Source
Target 0
Target 1
Target 2
Target 3
100%
70%
20%
100%
depends upon
depends upon
depends upon
Target 0 is 100% complete and is not dependent on any other target or source. Target 0 has two maps that are dependent on it. (Target 1 and Target 2)
Target 1 depends upon Target 0 and is 70% complete. Target 1 has one dependent map. (Target 2)
Target 2 depends upon Target 0 and Target 1 and is 20% complete. Target 2 has no dependent maps.
Target 3 is idl_copied and is not dependent on any target or source. Target 3 has no maps that are dependent on it.

*Figure 11-7 Interactions among multiple target FlashCopy mappings*

In Figure 11-7 on page 426, Target 0 is not dependent on a source because it completed copying. Target 0 has two dependent mappings (Target 1 and Target 2).

Target 1 depends on Target 0. It remains dependent until all of Target 1 is copied. Target 2 depends on it because Target 2 is 20% copy complete. After all of Target 1 is copied, it can then move to the idle_copied state.

Target 2 is dependent upon Target 0 and Target 1 and remains dependent until all of Target 2 is copied. No target depends on Target 2; therefore, when all of the data is copied to Target 2, it can move to the `idle_copied` state.

Target 3 completed copying, so it is not dependent on any other maps.

### Target writes with Multiple Target FlashCopy

A write to an intermediate or the newest target volume must consider the state of the grain within its own mapping, and the state of the grain of the next oldest mapping.

If the grain of the next oldest mapping is not yet copied, it must be copied before the write can proceed, to preserve the contents of the next oldest mapping. The data that is written to the next oldest mapping comes from a target or source.

If the grain in the target that is being written is not yet copied, the grain is copied from the oldest copied grain in the mappings that are newer than the target, or from the source if none is copied. After this copy is done, the write can be applied to the target.

### Target reads with Multiple Target FlashCopy

If the grain being read is copied from the source to the target, the read returns data from the target that is being read. If the grain is not yet copied, each of the newer mappings is examined in turn, and the read is performed from the first copy that is found. If none is found, the read is performed from the source.

### Stopping the copy process

When a **`stop`** command is issued to a mapping that contains a target that has dependent mappings, the mapping enters the `stopping` state and begins copying all grains that are uniquely held on the target volume of the mapping that is being stopped to the next oldest mapping that is in the Copying state. The mapping remains in the `stopping` state until all grains are copied, and then enters the `stopped` state.

> **Note:** The stopping copy process can be ongoing for several mappings that share the source at the same time. At the completion of this process, the mapping automatically makes an asynchronous state transition to the `stopped` state, or the `idle_copied` state if the mapping was in the `copying` state with progress = 100%.

For example, if the mapping that is associated with Target 0 was issued a **`stopfcmap`** or **`stopfcconsistgrp`** command, Target 0 enters the `stopping` state while a process copies the data of Target 0 to Target 1. After all of the data is copied, Target 0 enters the `stopped` state, and Target 1 is no longer dependent upon Target 0; however, Target 1 remains dependent on Target 2.

## 11.3.7 Summary of the FlashCopy indirection layer algorithm

Table 11-2 summarizes the indirection layer algorithm.

*Table 11-2 Summary table of the FlashCopy indirection layer algorithm*

| Accessed volume | Was the grain copied? | Host I/O operation | |
|---|---|---|---|
| | | **Read** | **Write** |
| Source | No | Read from the source volume. | Copy grain to most recently started target for this source, then write to the source. |
| | Yes | Read from the source volume. | Write to the source volume. |
| Target | No | If any newer targets exist for this source in which this grain was copied, read from the oldest of these targets. Otherwise, read from the source. | Hold the write. Check the dependency target volumes to see whether the grain was copied. If the grain is not copied to the next oldest target for this source, copy the grain to the next oldest target. Then, write to the target. |
| | Yes | Read from the target volume. | Write to the target volume. |

## 11.3.8 Interaction with the cache

Starting with V7.3, the entire cache subsystem was redesigned and changed accordingly. Cache has been divided into upper and lower cache. Upper cache serves mostly as write cache and hides the write latency from the hosts and application. Lower cache is a read/write cache and optimizes I/O to and from disks. Figure 11-8 shows the new IBM Spectrum Virtualize cache architecture.

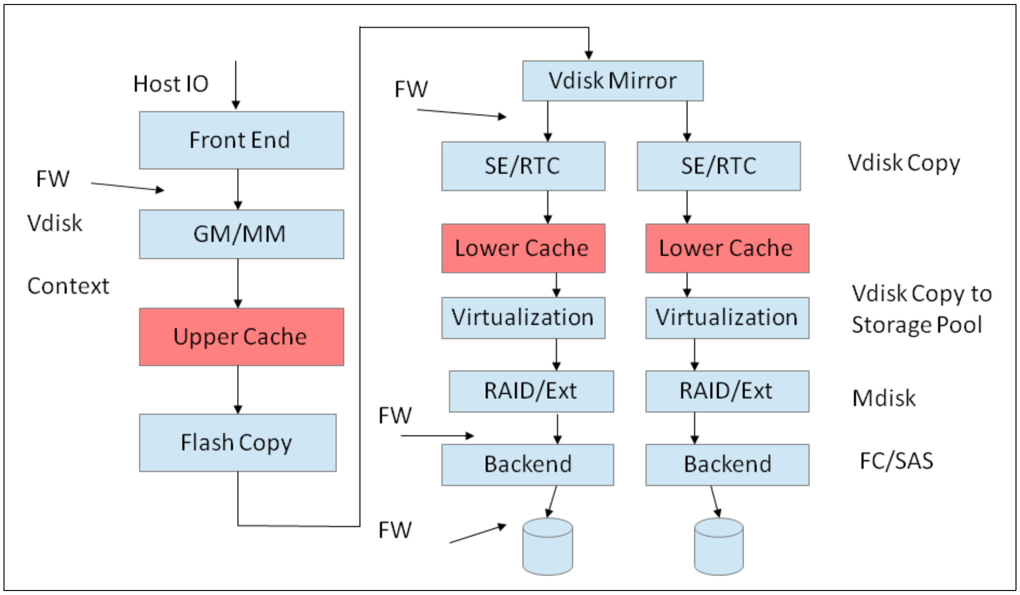


*Figure 11-8 New cache architecture*

This copy-on-write process introduces significant latency into write operations. To isolate the active application from this additional latency, the FlashCopy indirection layer is placed logically between upper and lower cache. Therefore, the additional latency that is introduced by the copy-on-write process is encountered only by the internal cache operations, and not by the application.

The logical placement of the FlashCopy indirection layer is shown in Figure 11-9.

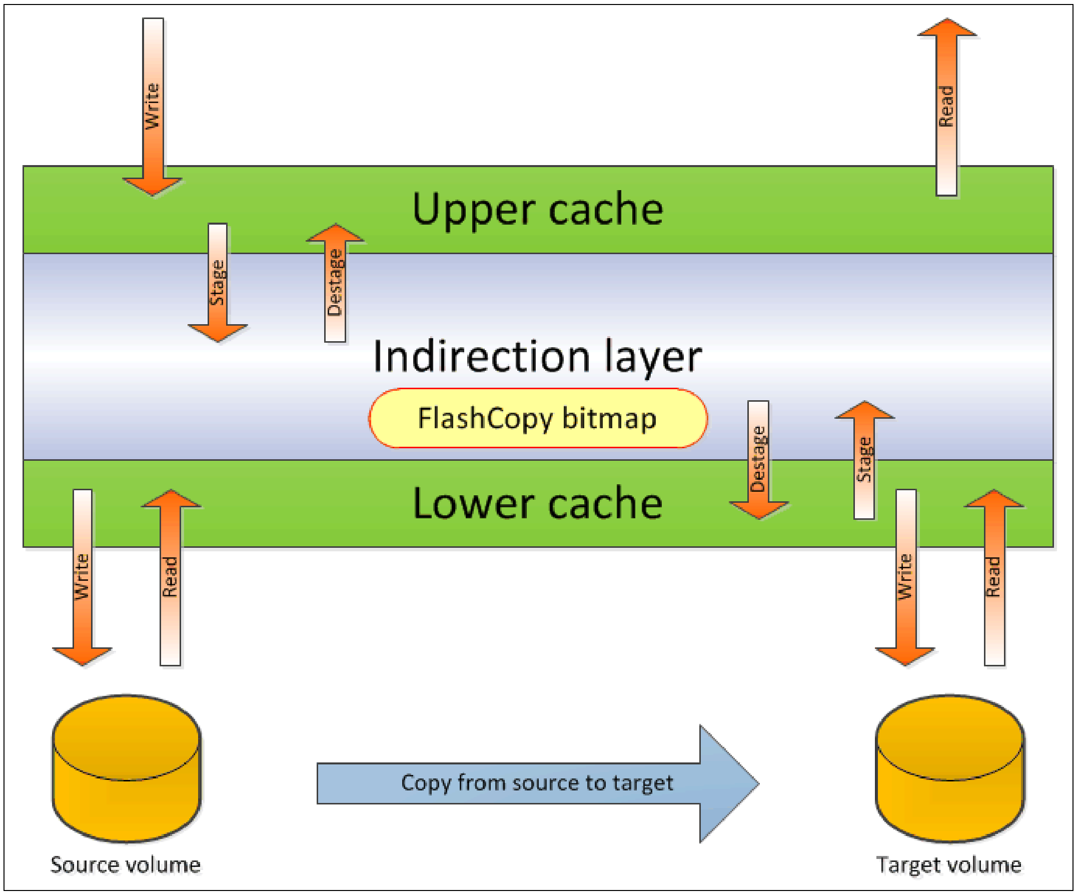


*Figure 11-9 Logical placement of the FlashCopy indirection layer*

Introduction of the two-level cache provides additional performance improvements to the FlashCopy mechanism. Because now the FlashCopy layer is above lower cache in the IBM Spectrum Virtualize software stack, it can benefit from read prefetching and coalescing writes to backend storage. Also, preparing FlashCopy is much faster, because upper cache write data does not have to go directly to backend storage but to lower cache layer.

Additionally, in the multitarget FlashCopy the target volumes of the same image share cache data. This design is opposite to previous IBM Spectrum Virtualize code versions, where each volume had its own copy of cached data.

## 11.3.9 FlashCopy and image mode volumes

FlashCopy can be used with image mode volumes. Because the source and target volumes must be the same size, you must create a target volume with the same size as the image mode volume when you are creating a FlashCopy mapping. To accomplish this task, use the **`svcinfo lsvdisk -bytes volumeName`** command. The size in bytes is then used to create the volume that is used in the FlashCopy mapping.

This method provides an exact number of bytes because image mode volumes might not line up one-to-one on other measurement unit boundaries. In Example 11-1, we list the size of the `test_image_vol_1` volume.

The `test_image_vol_copy_1` volume is then created, which specifies the same size.

*Example 11-1 Listing the size of a volume in bytes and creating a volume of equal size*

```
IBM_V7000_Gen2:superuser>lsvdisk -bytes test_image_vol_1
id 12
name test_image_vol_1
IO_group_id 0
IO_group_name io_grp0
status online
mdisk_grp_id 3
mdisk_grp_name temp_migration_pool
capacity 21474836480
type image
formatted no
formatting no
mdisk_id 5
mdisk_name mdisk3
FC_id
FC_name
RC_id
RC_name
vdisk_UID 600507680283818B300000000000000E
throttling 0
preferred_node_id 2
fast_write_state empty
cache readwrite
udid
fc_map_count 0
sync_rate 50
copy_count 1
se_copy_count 0
filesystem
mirror_write_priority latency
RC_change no
compressed_copy_count 0
access_IO_group_count 1
last_access_time
parent_mdisk_grp_id 3
parent_mdisk_grp_name temp_migration_pool
owner_type none
owner_id
owner_name
encrypt no
volume_id 12
volume_name test_image_vol_1
function

copy_id 0
status online
sync yes
auto_delete no
```

```
primary yes
mdisk_grp_id 3
mdisk_grp_name temp_migration_pool
type image
mdisk_id 5
mdisk_name mdisk3
fast_write_state empty
used_capacity 21474836480
real_capacity 21474836480
free_capacity 0
overallocation 100
autoexpand
warning
grainsize
se_copy no
easy_tier on
easy_tier_status measured
tier ssd
tier_capacity 0
tier enterprise
tier_capacity 21474836480
tier nearline
tier_capacity 0
compressed_copy no
uncompressed_used_capacity 21474836480
parent_mdisk_grp_id 3
parent_mdisk_grp_name temp_migration_pool
encrypt no


IBM_V7000_Gen2:superuser>mkvdisk -mdiskgrp test_pool_1 -iogrp 0 -size 21474836480
-unit b -name test_image_vol_copy_1
Virtual Disk, id [13], successfully created


IBM_V7000_Gen2:superuser>lsvdisk -delim " "
12 test_image_vol_1 0 io_grp0 online 3 temp_migration_pool 20.00GB image
600507680283818B300000000000000E 0 1 empty 0 no 0 3 temp_migration_pool no no 12
test_image_vol_1
13 test_image_vol_copy_1 0 io_grp0 online 0 test_pool_1 20.00GB striped
600507680283818B300000000000000F 0 1 not_empty 0 no 0 0 test_pool_1 yes no 13
test_image_vol_copy_1
```

**Tip:** Alternatively, you can use the **`expandvolumesize`** and **`shrinkvolumesize`** volume commands to modify the size of the volume.

These actions must be performed before a mapping is created.

You can use an image mode volume as a FlashCopy source volume or target volume.

## 11.3.10 FlashCopy mapping events

In this section, we describe the events that modify the states of a FlashCopy. We also describe the mapping events that are listed in Table 11-3.

**Overview of a FlashCopy sequence of events:** The following tasks show the FlashCopy sequence:

1. Associate the source data set with a target location (one or more source and target volumes).
2. Create a FlashCopy mapping for each source volume to the corresponding target volume. The target volume must be equal in size to the source volume.
3. Discontinue access to the target (application dependent).
4. Prepare (pre-trigger) the FlashCopy:
   a. Flush the cache for the source.
   b. Discard the cache for the target.
5. Start (trigger) the FlashCopy:
   a. Pause I/O (briefly) on the source.
   b. Resume I/O on the source.
   c. Start I/O on the target.

*Table 11-3 Mapping events*

| Mapping event | Description |
|---|---|
| Create | A FlashCopy mapping is created between the specified source volume and the specified target volume. The operation fails if any one of the following conditions is true:<br>▸ The source volume is a member of 256 FlashCopy mappings.<br>▸ The node has insufficient bitmap memory.<br>▸ The source and target volumes are different sizes. |
| Prepare | The **prestartfcmap** or **prestartfcconsistgrp** command is directed to a Consistency Group for FlashCopy mappings that are members of a normal Consistency Group or to the mapping name for FlashCopy mappings that are stand-alone mappings. The **prestartfcmap** or **prestartfcconsistgrp** command places the FlashCopy mapping into the Preparing state.<br><br>The **prestartfcmap** or **prestartfcconsistgrp** command can corrupt any data that was on the target volume because cached writes are discarded. Even if the FlashCopy mapping is never started, the data from the target might be changed logically during the act of preparing to start the FlashCopy mapping. |
| Flush done | The FlashCopy mapping automatically moves from the `preparing` state to the `prepared` state after all cached data for the source is flushed and all cached data for the target is no longer valid. |

| Mapping event | Description |
|---|---|
| Start | When all of the FlashCopy mappings in a Consistency Group are in the `prepared` state, the FlashCopy mappings can be started. To preserve the cross-volume Consistency Group, the start of all of the FlashCopy mappings in the Consistency Group must be synchronized correctly concerning I/Os that are directed at the volumes by using the **`startfcmap`** or **`startfcconsistgrp`** command.<br><br>The following actions occur during the running of the **`startfcmap`** command or the **`startfcconsistgrp`** command:<br>► New reads and writes to all source volumes in the Consistency Group are paused in the cache layer until all ongoing reads and writes beneath the cache layer are completed.<br>► After all FlashCopy mappings in the Consistency Group are paused, the internal cluster state is set to enable FlashCopy operations.<br>► After the cluster state is set for all FlashCopy mappings in the Consistency Group, read and write operations continue on the source volumes.<br>► The target volumes are brought online.<br>As part of the **`startfcmap`** or **`startfcconsistgrp`** command, read and write caching is enabled for the source and target volumes. |
| Modify | The following FlashCopy mapping properties can be modified:<br>► FlashCopy mapping name<br>► Clean rate<br>► Consistency group<br>► Copy rate (for background copy or stopping copy priority)<br>► Automatic deletion of the mapping when the background copy is complete |
| Stop | The following separate mechanisms can be used to stop a FlashCopy mapping:<br>► Issue a command<br>► An I/O error occurred |
| Delete | This command requests that the specified FlashCopy mapping is deleted. If the FlashCopy mapping is in the `copying` state, the **`force`** flag must be used. |
| Flush failed | If the flush of data from the cache cannot be completed, the FlashCopy mapping enters the `stopped` state. |
| Copy complete | After all of the source data is copied to the target and there are no dependent mappings, the state is set to `copied`. If the option to automatically delete the mapping after the background copy completes is specified, the FlashCopy mapping is deleted automatically. If this option is not specified, the FlashCopy mapping is not deleted automatically and can be reactivated by preparing and starting again. |
| Bitmap online/offline | The node failed. |

## 11.3.11 FlashCopy mapping states

In this section, we describe the states of a FlashCopy mapping.

### Idle_or_copied

The source and target volumes act as independent volumes even if a mapping exists between the two. Read and write caching is enabled for the source and the target volumes.

If the mapping is incremental and the background copy is complete, the mapping records the differences between the source and target volumes only. If the connection to both nodes in the I/O group that the mapping is assigned to is lost, the source and target volumes are offline.

### Copying

The copy is in progress. Read and write caching is enabled on the source and the target volumes.

### Prepared

The mapping is ready to start. The target volume is online, but is not accessible. The target volume cannot perform read or write caching. Read and write caching is failed by the Small Computer System Interface (SCSI) front end as a hardware error. If the mapping is incremental and a previous mapping is completed, the mapping records the differences between the source and target volumes only. If the connection to both nodes in the I/O group that the mapping is assigned to is lost, the source and target volumes go offline.

### Preparing

The target volume is online, but not accessible. The target volume cannot perform read or write caching. Read and write caching is failed by the SCSI front end as a hardware error. Any changed write data for the source volume is flushed from the cache. Any read or write data for the target volume is discarded from the cache.

If the mapping is incremental and a previous mapping is completed, the mapping records the differences between the source and target volumes only. If the connection to both nodes in the I/O group that the mapping is assigned to is lost, the source and target volumes go offline.

Performing the cache flush that is required as part of the **startfcmap** or **startfcconsistgrp** command causes I/Os to be delayed while they are waiting for the cache flush to complete. To overcome this problem, FlashCopy supports the **prestartfcmap** or **prestartfcconsistgrp** commands, which prepare for a FlashCopy start while still allowing I/Os to continue to the source volume.

In the Preparing state, the FlashCopy mapping is prepared by completing the following steps:

1. Flushing any modified write data that is associated with the source volume from the cache. Read data for the source is left in the cache.
2. Placing the cache for the source volume into write-through mode so that subsequent writes wait until data is written to disk before the **write** command that is received from the host is complete.
3. Discarding any read or write data that is associated with the target volume from the cache.

## Stopped

The mapping is stopped because you issued a `stop` command or an I/O error occurred. The target volume is offline and its data is lost. To access the target volume, you must restart or delete the mapping. The source volume is accessible and the read and write cache is enabled. If the mapping is incremental, the mapping is recording write operations to the source volume. If the connection to both nodes in the I/O group that the mapping is assigned to is lost, the source and target volumes go offline.

## Stopping

The mapping is copying data to another mapping.

If the background copy process is complete, the target volume is online while the stopping copy process completes.

If the background copy process is not complete, data is discarded from the target volume cache. The target volume is offline while the stopping copy process runs.

The source volume is accessible for I/O operations.

## Suspended

The mapping started, but it did not complete. Access to the metadata is lost, which causes the source and target volume to go offline. When access to the metadata is restored, the mapping returns to the copying or stopping state and the source and target volumes return online. The background copy process resumes. Any data that was not flushed and was written to the source or target volume before the suspension is in cache until the mapping leaves the suspended state.

## Summary of FlashCopy mapping states

Table 11-4 lists the various FlashCopy mapping states, and the corresponding states of the source and target volumes.

*Table 11-4 FlashCopy mapping state summary*

| State | Source | | Target | |
|---|---|---|---|---|
| | **Online/Offline** | **Cache state** | **Online/Offline** | **Cache state** |
| Idling/Copied | Online | Write-back | Online | Write-back |
| Copying | Online | Write-back | Online | Write-back |
| Stopped | Online | Write-back | Offline | N/A |
| Stopping | Online | Write-back | ► Online if copy complete<br>► Offline if copy incomplete | N/A |
| Suspended | Offline | Write-back | Offline | N/A |
| Preparing | Online | Write-through | Online but not accessible | N/A |
| Prepared | Online | Write-through | Online but not accessible | N/A |

## 11.3.12 Thin provisioned FlashCopy

FlashCopy source and target volumes can be thin-provisioned.

### Source or target thin-provisioned

The most common configuration is a fully allocated source and a thin-provisioned target. By using this configuration, the target uses a smaller amount of real storage than the source. With this configuration, use the `NOCOPY` (background copy rate = 0%) option only. Although the `COPY` option is supported, this option creates a fully allocated target, which defeats the purpose of thin provisioning.

### Source and target thin-provisioned

When the source and target volumes are thin-provisioned, only the data that is allocated to the source is copied to the target. In this configuration, the background copy option has no effect.

> **Performance:** The best performance is obtained when the grain size of the thin-provisioned volume is the same as the grain size of the FlashCopy mapping.

### Thin-provisioned incremental FlashCopy

The implementation of thin-provisioned volumes does not preclude the use of incremental FlashCopy on the same volumes. It does not make sense to have a fully allocated source volume and then use incremental FlashCopy (which is always a full copy the first time) to copy this fully allocated source volume to a thin-provisioned target volume. However, this action is not prohibited.

Consider the following optional configurations:

- A thin-provisioned source volume can be copied incrementally by using FlashCopy to a thin-provisioned target volume. Whenever the FlashCopy is performed, only data that was modified is recopied to the target. If space is allocated on the target because of I/O to the target volume, this space is not reclaimed with subsequent FlashCopy operations.
- A fully allocated source volume can be copied incrementally by using FlashCopy to another fully allocated volume at the same time as it is being copied to multiple thin-provisioned targets (taken at separate points in time). By using this combination, a single full backup can be kept for recovery purposes, and the backup workload is separated from the production workload. At the same time, older thin-provisioned backups can be retained.

## 11.3.13 Background copy

With FlashCopy background copy enabled, the source volume data is copied to the corresponding target volume. With the FlashCopy background copy disabled, only data that changed on the source volume is copied to the target volume.

The benefit of the use of a FlashCopy mapping with background copy enabled is that the target volume becomes a real clone (independent from the source volume) of the FlashCopy mapping source volume after the copy is complete. When the background copy function is not performed, the target volume remains a valid copy of the source data only while the FlashCopy mapping remains in place.

The *background copy rate* is a property of a FlashCopy mapping that is defined as a value 0 - 100. The background copy rate can be defined and changed dynamically for individual FlashCopy mappings. A value of 0 disables the background copy.

Table 11-5 shows the relationship of the background copy rate value to the attempted number of grains to be copied per second.

*Table 11-5 Background copy rate*

| Value | Data copied per second | Grains per second (256 kilobyte (KB) grain) | Grains per second (64 KB grain) |
|---|---|---|---|
| 01 - 10 | 128 KiB | 0.5 | 2 |
| 11 - 20 | 256 KiB | 1 | 4 |
| 21 - 30 | 512 KiB | 2 | 8 |
| 31 - 40 | 1 mebibyte (MiB) | 4 | 16 |
| 41 - 50 | 2 MiB | 8 | 32 |
| 51 - 60 | 4 MiB | 16 | 64 |
| 61 - 70 | 8 MiB | 32 | 128 |
| 71 - 80 | 16 MiB | 64 | 256 |
| 81 - 90 | 32 MiB | 128 | 512 |
| 91 - 100 | 64 MiB | 256 | 1024 |

The grains per second numbers represent the maximum number of grains that the IBM Storwize V7000 copies per second, assuming that the bandwidth to the managed disks (MDisks) can accommodate this rate.

If the IBM Storwize V7000 cannot achieve these copy rates because of insufficient width from the nodes to the MDisks, the background copy I/O contends for resources on an equal basis with the I/O that is arriving from the hosts. Background copy I/O and I/O that is arriving from the hosts tend to see an increase in latency and a consequential reduction in throughput.

Background copy and foreground I/O continue to make progress, and do not stop, hang, or cause the node to fail. The background copy is performed by both nodes of the I/O Group in which the source volume is found.

## 11.3.14 Serialization of I/O by FlashCopy

In general, the FlashCopy function in the IBM Spectrum Virtualize introduces no explicit serialization into the I/O path. Therefore, many concurrent I/Os are allowed to the source and target volumes.

However, there is a lock for each grain. The lock can be in shared or exclusive mode. For multiple targets, a common lock is shared, and the mappings are derived from a particular source volume. The lock is used in the following modes under the following conditions:

- The lock is held in shared mode during a read from the target volume, which touches a grain that was not copied from the source.
- The lock is held in exclusive mode while a grain is being copied from the source to the target.

If the lock is held in shared mode and another process wants to use the lock in shared mode, this request is granted unless a process is already waiting to use the lock in exclusive mode.

If the lock is held in shared mode and it is requested to be exclusive, the requesting process must wait until all holders of the shared lock free it.

Similarly, if the lock is held in exclusive mode, a process that is wanting to use the lock in shared or exclusive mode must wait for it to be freed.

## 11.3.15 Event handling

When a FlashCopy mapping is not copying or stopping, the FlashCopy function does not affect the handling or reporting of events for error conditions that are encountered in the I/O path. Event handling and reporting are affected only by FlashCopy when a FlashCopy mapping is copying or stopping; that is, actively moving data.

We describe these scenarios next.

### Node failure

Normally, two copies of the FlashCopy bitmap are maintained. One copy of the FlashCopy bitmap is on each of the two nodes that make up the I/O Group of the source volume. When a node fails, one copy of the bitmap for all FlashCopy mappings whose source volume is a member of the failing node's I/O Group becomes inaccessible.

FlashCopy continues with a single copy of the FlashCopy bitmap that is stored as non-volatile in the remaining node in the source I/O Group. The system metadata is updated to indicate that the missing node no longer holds a current bitmap. When the failing node recovers or a replacement node is added to the I/O Group, the bitmap redundancy is restored.

### Path failure (Path Offline state)

In a fully functioning system, all of the nodes have a software representation of every volume in the system within their application hierarchy.

Because the storage area network (SAN) that links IBM Storwize V7000 nodes to each other and to the MDisks is made up of many independent links, it is possible for a subset of the nodes to be temporarily isolated from several of the MDisks. When this situation happens, the managed disks are said to be *path offline* on certain nodes.

**Other nodes:** Other nodes might see the managed disks as Online because their connection to the managed disks is still functioning.

#### *Path Offline for the source volume*

If a FlashCopy mapping is in the `copying` state and the source volume goes path offline, this path offline state is propagated to all target volumes up to, but not including, the target volume for the newest mapping that is 100% copied but remains in the `copying` state. If no mappings are 100% copied, all of the target volumes are taken offline. `Path offline` is a state that exists on a per-node basis. Other nodes might not be affected. If the source volume comes online, the target and source volumes are brought back online.

#### *Path Offline for the target volume*

If a target volume goes path offline but the source volume is still online, and if there are any dependent mappings, those target volumes also go path offline. The source volume remains online.

### 11.3.16 Asynchronous notifications

FlashCopy raises informational event log entries for certain mapping and Consistency Group state transitions.

These state transitions occur as a result of configuration events that complete asynchronously. The informational events can be used to generate Simple Network Management Protocol (SNMP) traps to notify the user. Other configuration events complete synchronously, and no informational events are logged as a result of the following events:

- `PREPARE_COMPLETED`. This state transition is logged when the FlashCopy mapping or Consistency Group enters the `prepared` state as a result of a user request to prepare. The user can now start (or stop) the mapping or Consistency Group.
- `COPY_COMPLETED`. This state transition is logged when the FlashCopy mapping or Consistency Group enters the `idle_or_copied` state when it was in the `copying` or `stopping` state. This state transition indicates that the target disk now contains a complete copy and no longer depends on the source.
- `STOP_COMPLETED`. This state transition is logged when the FlashCopy mapping or Consistency Group enters the `stopped` state as a result of a user request to stop. It is logged after the automatic copy process completes. This state transition includes mappings where no copying needed to be performed. This state transition differs from the event that is logged when a mapping or group enters the `stopped` state as a result of an I/O error.

### 11.3.17 Interoperation with Metro Mirror and Global Mirror

A volume can be part of any copy relationship (FlashCopy, Metro Mirror or Remote Mirror). Therefore, FlashCopy can work with Metro Mirror and Global Mirror to provide better protection of the data.

For example, we can perform a Metro Mirror copy to duplicate data from Site_A to Site_B and then perform a daily FlashCopy to back up the data to another location.

> **Note:** A volume cannot be part of FlashCopy, Metro Mirror or Remote Mirror, if it is set to Transparent Cloud Tiering function.

Table 11-6 lists the supported combinations of FlashCopy and remote copy. In the table, *remote copy* refers to Metro Mirror and Global Mirror.

*Table 11-6 FlashCopy and remote copy interaction*

| Component | Remote copy primary site | Remote copy secondary site |
|---|---|---|
| FlashCopy Source | Supported | Supported latency: When the FlashCopy relationship is in the `preparing` and `prepared` states, the cache at the remote copy secondary site operates in write-through mode.<br>This process adds latency to the latent remote copy relationship. |

| Component | Remote copy primary site | Remote copy secondary site |
|---|---|---|
| FlashCopy<br>Target | This is a supported combination and has the following restrictions:<br>► Issuing a **`stop -force`** might cause the remote copy relationship to be fully resynchronized.<br>► Code level must be 6.2.*x* or later.<br>► I/O Group must be the same. | This is a supported combination with the major restriction that the FlashCopy mapping cannot be `copying`, `stopping`, or `suspended`. Otherwise, the restrictions are the same as at the remote copy primary site. |

## 11.3.18 FlashCopy presets

The IBM Spectrum Virtualize GUI interface provides three FlashCopy presets (Snapshot, Clone, and Backup) to simplify the more common FlashCopy operations.

Although these presets meet most FlashCopy requirements, they do not provide support for all possible FlashCopy options. If more specialized options are required that are not supported by the presets, the options must be performed by using CLI commands.

In this section, we describe the three preset options and their use cases.

### Snapshot

This preset creates a copy-on-write point-in-time copy. The snapshot is not intended to be an independent copy. Instead, the copy is used to maintain a view of the production data at the time that the snapshot is created. Therefore, the snapshot holds only the data from regions of the production volume that changed since the snapshot was created. Because the snapshot preset uses thin provisioning, only the capacity that is required for the changes is used.

Snapshot uses the following preset parameters:

- Background copy: None
- Incremental: No
- Delete after completion: No
- Cleaning rate: No
- Primary copy source pool: Target pool

#### *Use case*

The user wants to produce a copy of a volume without affecting the availability of the volume. The user does not anticipate many changes to be made to the source or target volume; a significant proportion of the volumes remains unchanged.

By ensuring that only changes require a copy of data to be made, the total amount of disk space that is required for the copy is reduced. Therefore, many Snapshot copies can be used in the environment.

Snapshots are useful for providing protection against corruption or similar issues with the validity of the data, but they do not provide protection from physical controller failures. Snapshots can also provide a vehicle for performing repeatable testing (including “what-if” modeling that is based on production data) without requiring a full copy of the data to be provisioned.

### Clone

The clone preset creates a replica of the volume, which can be changed without affecting the original volume. After the copy completes, the mapping that was created by the preset is automatically deleted.

Clone uses the following preset parameters:

- Background copy rate: 50
- Incremental: No
- Delete after completion: Yes
- Cleaning rate: 50
- Primary copy source pool: Target pool

#### *Use case*

Users want a copy of the volume that they can modify without affecting the original volume. After the clone is established, there is no expectation that it is refreshed or that there is any further need to reference the original production data again. If the source is thin-provisioned, the target is thin-provisioned for the auto-create target.

### Backup

The backup preset creates a point-in-time replica of the production data. After the copy completes, the backup view can be refreshed from the production data, with minimal copying of data from the production volume to the backup volume.

Backup uses the following preset parameters:

- Background Copy rate: 50
- Incremental: Yes
- Delete after completion: No
- Cleaning rate: 50
- Primary copy source pool: Target pool

#### *Use case*

The user wants to create a copy of the volume that can be used as a backup if the source becomes unavailable, as in the case of loss of the underlying physical controller. The user plans to periodically update the secondary copy, and does not want to suffer from the resource demands of creating a new copy each time (and incremental FlashCopy times are faster than full copy, which helps to reduce the window where the new backup is not yet fully effective). If the source is thin-provisioned, the target is also thin-provisioned in this option for the auto-create target.

Another use case, which is not supported by the name, is to create and maintain (periodically refresh) an independent image that can be subjected to intensive I/O (for example, data mining) without affecting the source volume’s performance.

## 11.4 Managing FlashCopy using GUI

It is often easier to work with the FlashCopy function from the GUI if you have a reasonable number of host mappings. However, in enterprise data centers with many host mappings, we suggest that you use the CLI to run your FlashCopy commands.

In this section, we describe the tasks that you can perform at a FlashCopy level using the IBM Spectrum Virtualize GUI.

The following methods can be used to visualize and manage your FlashCopy:

- Use the Overview pane. Move the mouse pointer over Copy Services in the dynamic menu and click **FlashCopy**, as shown in Figure 11-10.

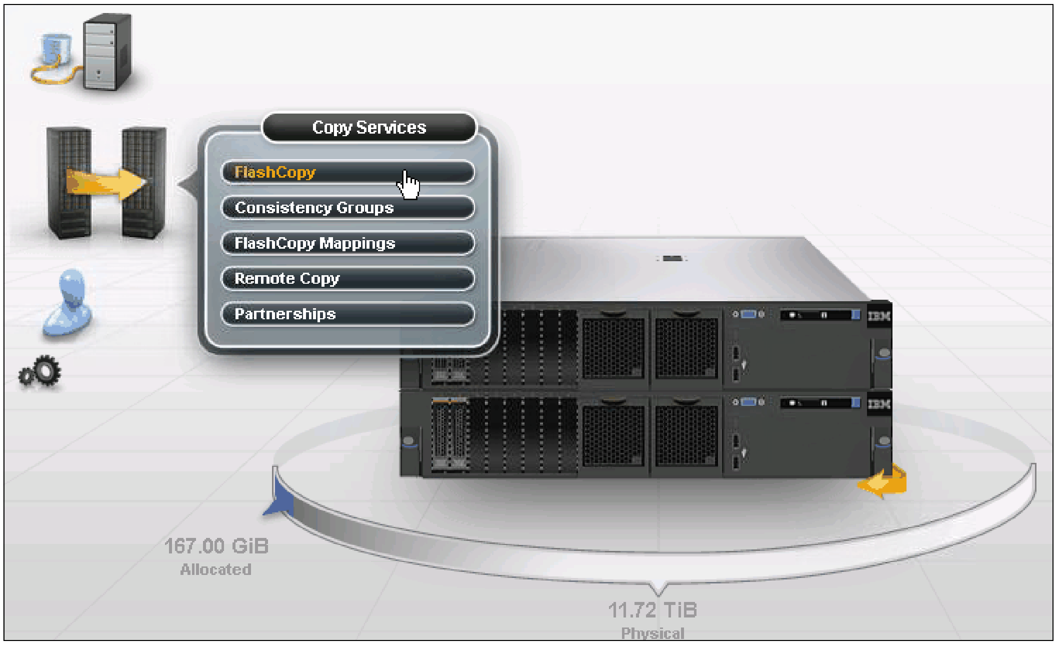
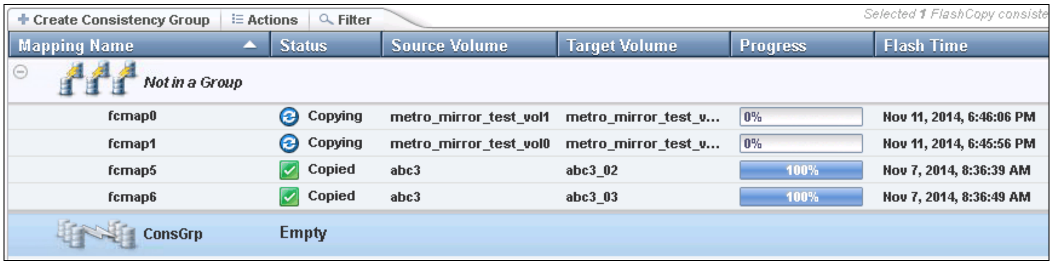


*Figure 11-10 FlashCopy pane*

In its basic mode, the IBM FlashCopy function copies the contents of a source volume to a target volume. Any data that existed on the target volume is lost, and that data is replaced by the copied data.

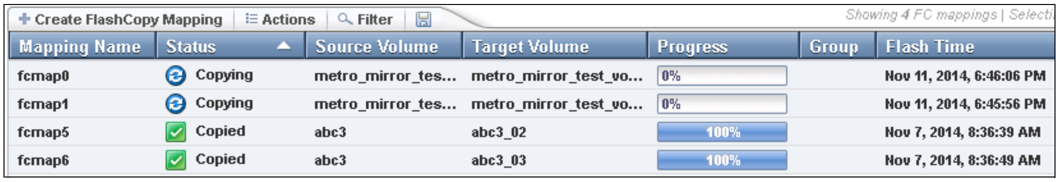
- Use the Consistency Groups pane, as shown in Figure 11-11. A *Consistency Group* is a container for mappings. You can add many mappings to a Consistency Group.



| Mapping Name | Status | Source Volume | Target Volume | Progress | Flash Time |
|---|---|---|---|---|---|
| Not in a Group | | | | | |
| fcmap0 | Copying | metro_mirror_test_vol1 | metro_mirror_test_v... | 0% | Nov 11, 2014, 6:46:06 PM |
| fcmap1 | Copying | metro_mirror_test_vol0 | metro_mirror_test_v... | 0% | Nov 11, 2014, 6:45:56 PM |
| fcmap5 | ✓ Copied | abc3 | abc3_02 | 100% | Nov 7, 2014, 8:36:39 AM |
| fcmap6 | ✓ Copied | abc3 | abc3_03 | 100% | Nov 7, 2014, 8:36:49 AM |
| ConsGrp | Empty | | | | |

*Figure 11-11 Consistency Groups pane*

- Use the FlashCopy Mappings pane, as shown in Figure 11-12. A *FlashCopy mapping* defines the relationship between a source volume and a target volume.



| Mapping Name | Status | Source Volume | Target Volume | Progress | Group | Flash Time |
|---|---|---|---|---|---|---|
| fcmap0 | Copying | metro_mirror_tes... | metro_mirror_test_vo... | 0% | | Nov 11, 2014, 6:46:06 PM |
| fcmap1 | Copying | metro_mirror_tes... | metro_mirror_test_vo... | 0% | | Nov 11, 2014, 6:45:56 PM |
| fcmap5 | ✓ Copied | abc3 | abc3_02 | 100% | | Nov 7, 2014, 8:36:39 AM |
| fcmap6 | ✓ Copied | abc3 | abc3_03 | 100% | | Nov 7, 2014, 8:36:49 AM |

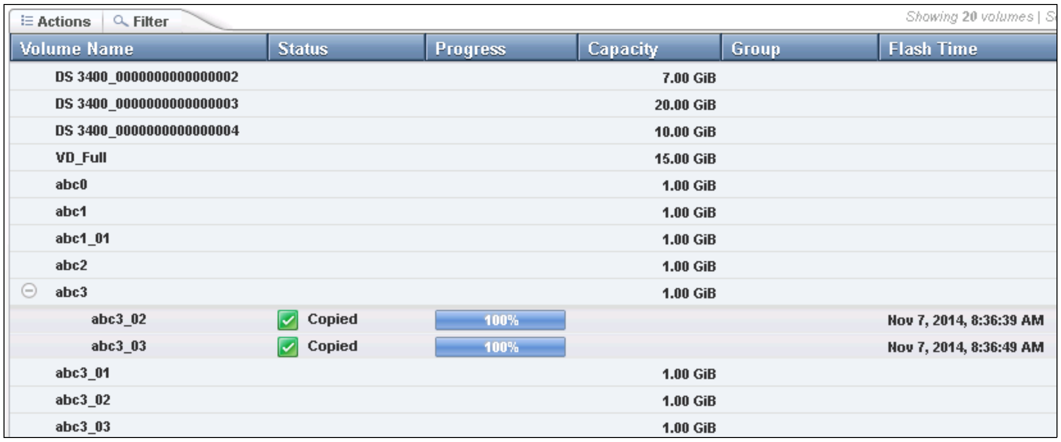
*Figure 11-12 FlashCopy Mappings pane*

## 11.4.1 Creating a FlashCopy mapping

In this section, we create FlashCopy mappings for volumes and their targets.

Complete the following steps:

1. From the Overview pane, move the mouse pointer over Copy Services in the dynamic menu and click **FlashCopy**. The FlashCopy pane opens, as shown in Figure 11-13.

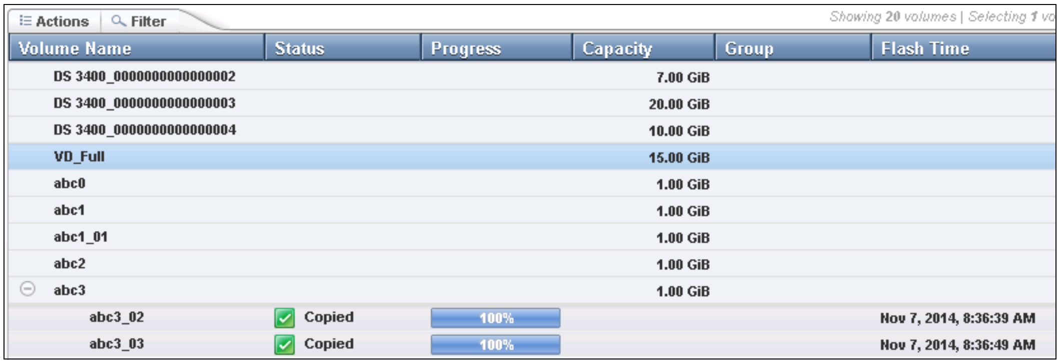


| Volume Name | Status | Progress | Capacity | Group | Flash Time |
|---|---|---|---|---|---|
| DS 3400_0000000000000002 | | | 7.00 GiB | | |
| DS 3400_0000000000000003 | | | 20.00 GiB | | |
| DS 3400_0000000000000004 | | | 10.00 GiB | | |
| VD_Full | | | 15.00 GiB | | |
| abc0 | | | 1.00 GiB | | |
| abc1 | | | 1.00 GiB | | |
| abc1_01 | | | 1.00 GiB | | |
| abc2 | | | 1.00 GiB | | |
| ⊖ abc3 | | | 1.00 GiB | | |
| abc3_02 | ✓ Copied | 100% | | | Nov 7, 2014, 8:36:39 AM |
| abc3_03 | ✓ Copied | 100% | | | Nov 7, 2014, 8:36:49 AM |
| abc3_01 | | | 1.00 GiB | | |
| abc3_02 | | | 1.00 GiB | | |
| abc3_03 | | | 1.00 GiB | | |

*Figure 11-13 FlashCopy pane*

2. Select the volume for which you want to create the FlashCopy relationship, as shown in Figure 11-14.

> **Multiple FlashCopy mappings:** To create multiple FlashCopy mappings at one time, select multiple volumes by holding down Ctrl and clicking the entries that you want.



| Volume Name | Status | Progress | Capacity | Group | Flash Time |
|---|---|---|---|---|---|
| DS 3400_0000000000000002 | | | 7.00 GiB | | |
| DS 3400_0000000000000003 | | | 20.00 GiB | | |
| DS 3400_0000000000000004 | | | 10.00 GiB | | |
| VD_Full | | | 15.00 GiB | | |
| abc0 | | | 1.00 GiB | | |
| abc1 | | | 1.00 GiB | | |
| abc1_01 | | | 1.00 GiB | | |
| abc2 | | | 1.00 GiB | | |
| ⊖ abc3 | | | 1.00 GiB | | |
| abc3_02 | ✓ Copied | 100% | | | Nov 7, 2014, 8:36:39 AM |
| abc3_03 | ✓ Copied | 100% | | | Nov 7, 2014, 8:36:49 AM |

*Figure 11-14 FlashCopy mapping: Select the volume (or volumes)*

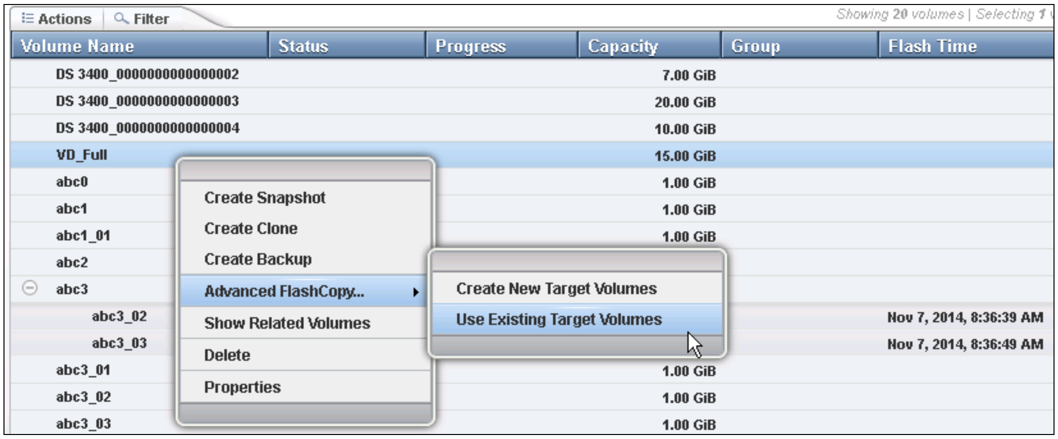
Depending on whether you created the target volumes for your FlashCopy mappings or you want the system to create the target volumes for you, the following options are available:

- If you created the target volumes, see “Using existing target volumes” on page 443.
- If you want the system to create the target volumes for you, see “Creating target volumes” on page 447.

### Using existing target volumes

Complete the following steps to use existing target volumes for the FlashCopy mappings:

1. Select the target volume that you want to use. Then, click **Actions** → **Advanced FlashCopy** → **Use Existing Target Volumes**, as shown in Figure 11-15.

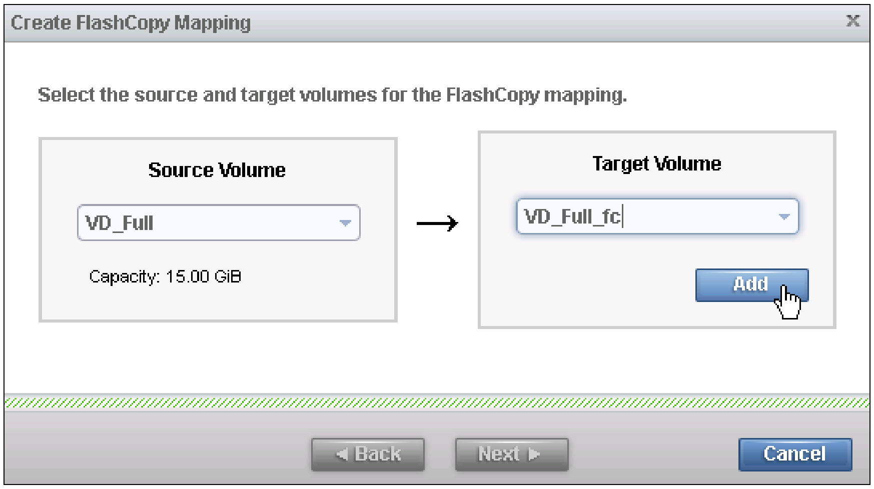


| Volume Name | Status | Progress | Capacity | Group | Flash Time |
|---|---|---|---|---|---|
| DS 3400_0000000000000002 | | | 7.00 GiB | | |
| DS 3400_0000000000000003 | | | 20.00 GiB | | |
| DS 3400_0000000000000004 | | | 10.00 GiB | | |
| VD_Full | | | 15.00 GiB | | |
| abc0 | | | 1.00 GiB | | |
| abc1 | | | 1.00 GiB | | |
| abc1_01 | | | 1.00 GiB | | |
| abc2 | | | | | |
| abc3 | | | | | |
| abc3_02 | | | | | Nov 7, 2014, 8:36:39 AM |
| abc3_03 | | | | | Nov 7, 2014, 8:36:49 AM |
| abc3_01 | | | 1.00 GiB | | |
| abc3_02 | | | 1.00 GiB | | |
| abc3_03 | | | 1.00 GiB | | |

*Figure 11-15 Using existing target volumes*

2. The Create FlashCopy Mapping window opens (Figure 11-16). In this window, you must create the relationship between the source volume (the disk that is copied) and the target volume (the disk that receives the copy). A mapping can be created between any two volumes managed by the same clustered system. Select a source volume and a target volume for your FlashCopy mapping, and then click **Add**. If you must create other copies, repeat this step.

> **Important:** The source volume and the target volume must be of equal size. Therefore, only targets of the same size are shown in the list for a source volume.



Select the source and target volumes for the FlashCopy mapping.

Source Volume: VD_Full

Capacity: 15.00 GiB

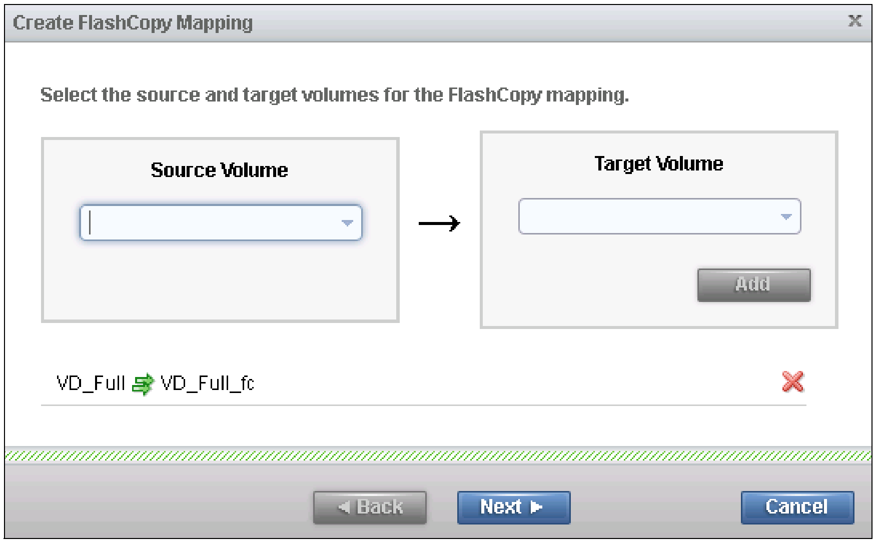
Target Volume: VD_Full_fc

*Figure 11-16 Create a FlashCopy Mapping by using an existing target volume*

To remove a relationship that was created, click ✖, as shown in Figure 11-17 on page 445.

> **Volumes:** The volumes do not have to be in the same I/O Group or storage pool.

3. Click **Next** after you create all of the relationships that you need, as shown in Figure 11-17.

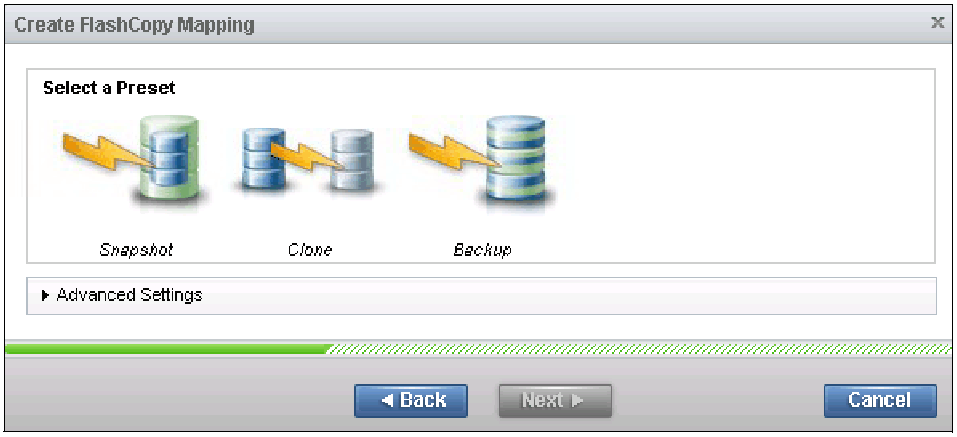


Create FlashCopy Mapping

Select the source and target volumes for the FlashCopy mapping.

Figure 11-17 Create FlashCopy Mapping window

4. In the next window, select one FlashCopy preset. The GUI provides the following presets to simplify common FlashCopy operations, as shown in Figure 11-18:
   – Snapshot: Creates a copy-on-write point-in-time copy.
   – Clone: Creates a replica of the source volume on a target volume. The copy can be changed without affecting the original volume.
   – Backup: Creates a FlashCopy mapping that can be used to recover data or objects if the system experiences data loss. These backups can be copied multiple times from source and target volumes.

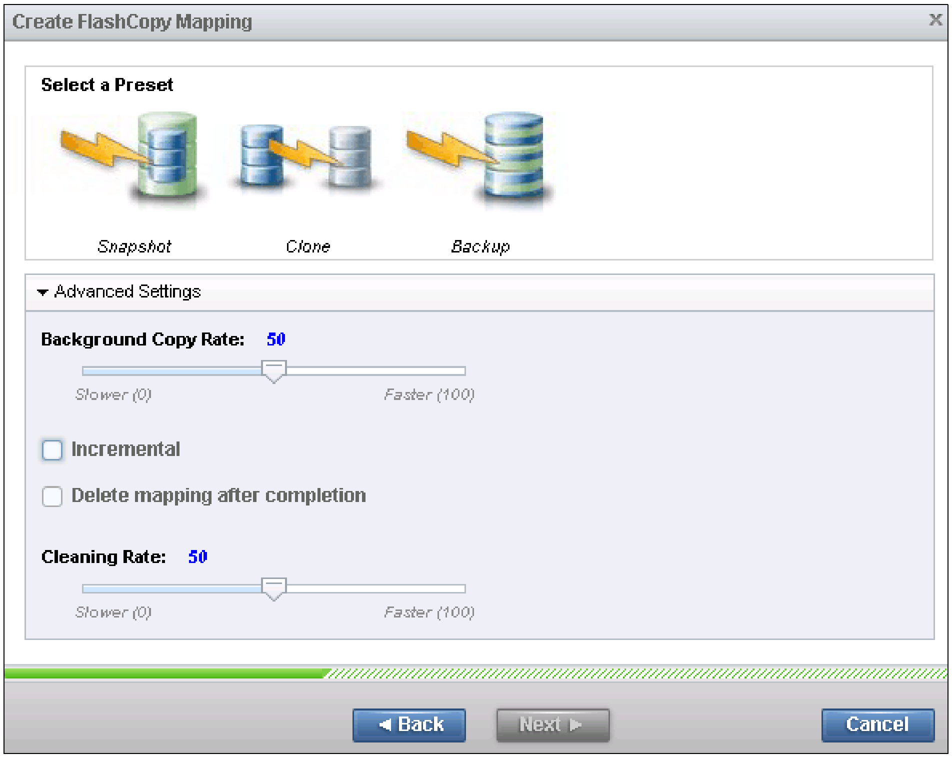


Create FlashCopy Mapping

Figure 11-18 Create FlashCopy Mapping window

For each preset, you can customize various advanced options. You can access these settings by clicking **Advanced Settings**.

5. The advanced setting options are shown in Figure 11-19.



Background Copy Rate: 50

Slower (0) Faster (100)

□ Incremental

□ Delete mapping after completion

Cleaning Rate: 50

Slower (0) Faster (100)

*Figure 11-19 Create FlashCopy Mapping Advanced Settings*

If you prefer not to customize these settings, go directly to step 6.

You can customize the following advanced setting options, as shown in Figure 11-19:

– Background Copy Rate

– Incremental

> **Incremental FlashCopy mapping:** Even if the type of the FlashCopy mapping is incremental, the first copy process copies all of the data from the source volume to the target volume.

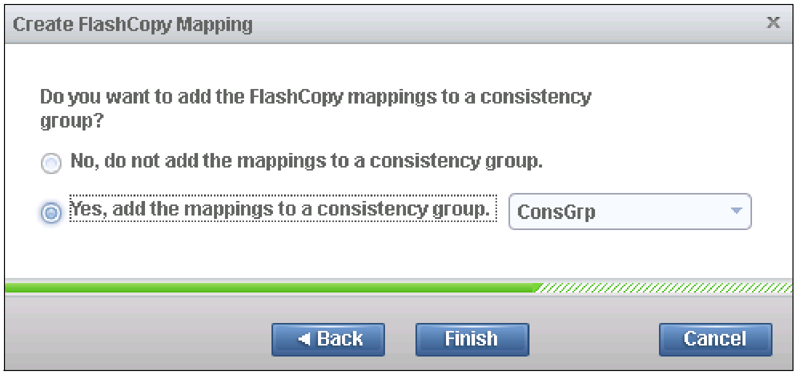
– Delete mapping after completion

– Cleaning Rate

After you complete your modifications, click **Next**.

6. You can choose whether to add the mappings to a Consistency Group or not.

If you want to include this FlashCopy mapping in a Consistency Group, select **Yes, add the mappings to a consistency group** in the window that is shown in Figure 11-20. You also can select the Consistency Group from the drop-down list.



Create FlashCopy Mapping

Do you want to add the FlashCopy mappings to a consistency group?

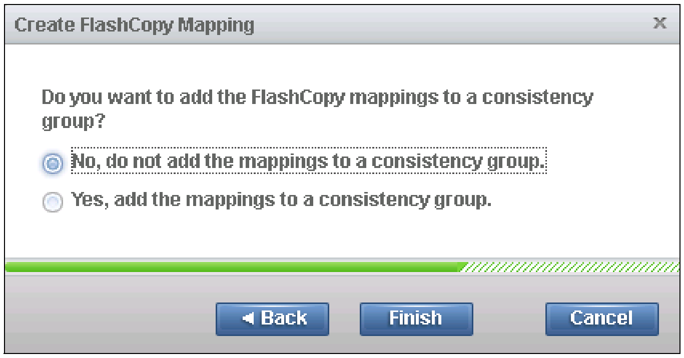
○ No, do not add the mappings to a consistency group.

● Yes, add the mappings to a consistency group. ConsGrp

*Figure 11-20 Add the mappings to a Consistency Group*

Or, if you do not want to include this FlashCopy mapping in a Consistency Group, select **No, do not add the mappings to a consistency group**.

Click **Finish**, as shown in Figure 11-21.



Create FlashCopy Mapping

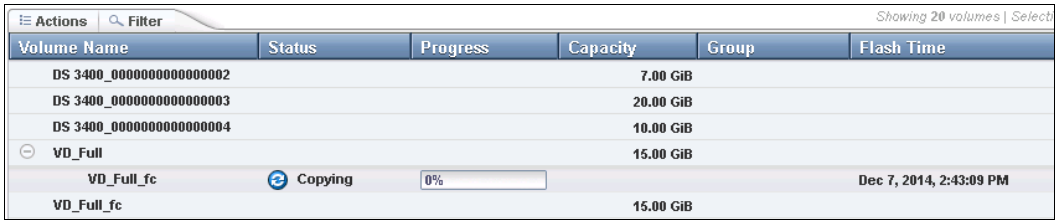
Do you want to add the FlashCopy mappings to a consistency group?

● No, do not add the mappings to a consistency group.

○ Yes, add the mappings to a consistency group.

*Figure 11-21 Do not add the mappings to a Consistency Group*

7. Check the result of this FlashCopy mapping. For each FlashCopy mapping relationship that was created, a mapping name is automatically generated that starts with `fcmapX`, where *X* is the next available number. If needed, you can rename these mappings, as shown in Figure 11-22. For more information, see Renaming FlashCopy mapping.



| Volume Name | Status | Progress | Capacity | Group | Flash Time |
|---|---|---|---|---|---|
| DS 3400_0000000000000002 | | | 7.00 GiB | | |
| DS 3400_0000000000000003 | | | 20.00 GiB | | |
| DS 3400_0000000000000004 | | | 10.00 GiB | | |
| ⊖ VD_Full | | | 15.00 GiB | | |
| VD_Full_fc | Copying | 0% | | | Dec 7, 2014, 2:43:09 PM |
| VD_Full_fc | | | 15.00 GiB | | |

*Figure 11-22 FlashCopy Mapping*

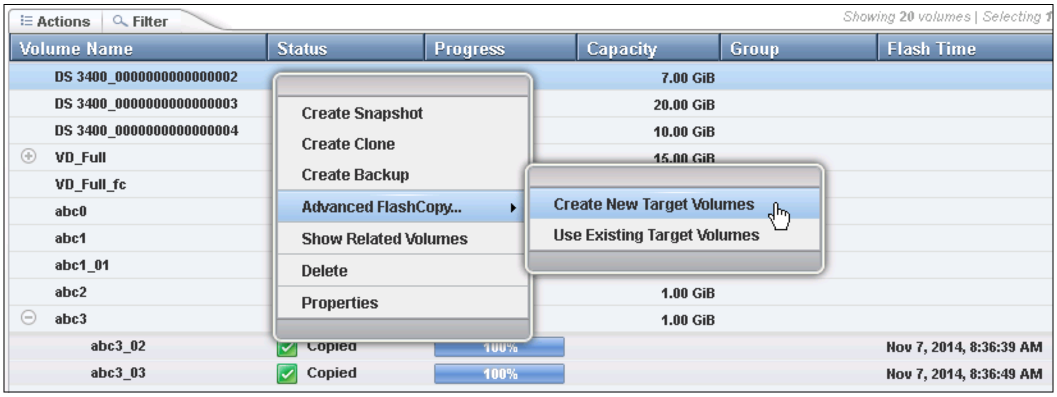
The FlashCopy mapping is now ready for use.

## Creating target volumes

Complete the following steps to create target volumes for FlashCopy mapping:

1. If you did not create a target volume for this source volume, click **Actions** → **Advanced FlashCopy** → **Create New Target Volumes**, as shown in Figure 11-23 on page 448.

> **Target volume naming:** If the target volume does not exist, the target volume is created. The target volume name is based on its source volume and a generated number at the end, for example, `source_volume_name_XX`, where *XX* is a number that was generated dynamically.



| Volume Name | Status | Progress | Capacity | Group | Flash Time |
|---|---|---|---|---|---|
| DS 3400_0000000000000002 | | | 7.00 GiB | | |
| DS 3400_0000000000000003 | | | 20.00 GiB | | |
| DS 3400_0000000000000004 | | | 10.00 GiB | | |
| VD_Full | | | 15.00 GiB | | |
| VD_Full_fc | | | | | |
| abc0 | | | | | |
| abc1 | | | | | |
| abc1_01 | | | | | |
| abc2 | | | 1.00 GiB | | |
| abc3 | | | 1.00 GiB | | |
| abc3_02 | ✓ Copied | 100% | | | Nov 7, 2014, 8:36:39 AM |
| abc3_03 | ✓ Copied | 100% | | | Nov 7, 2014, 8:36:49 AM |

*Figure 11-23 Selecting Create New Target Volumes*

2. In the Create FlashCopy Mapping window (Figure 11-24), you must select one FlashCopy preset. The GUI provides the following presets to simplify common FlashCopy operations:
   – Snapshot
   – Clone
   – Backup

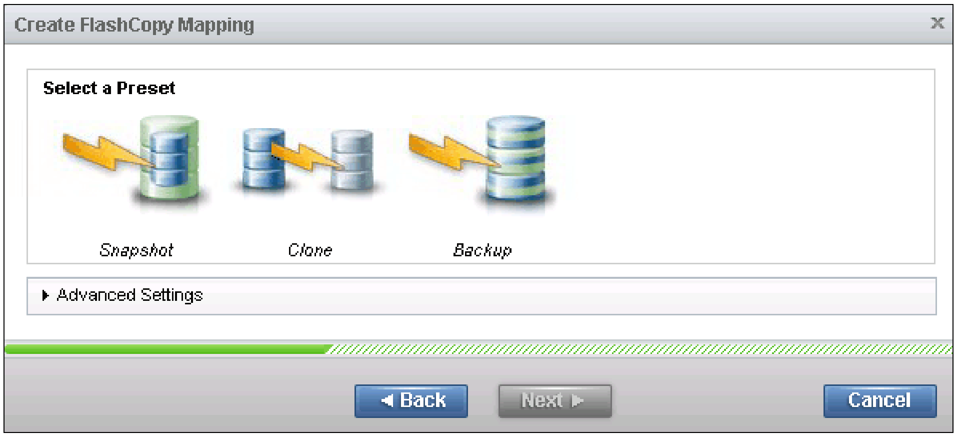
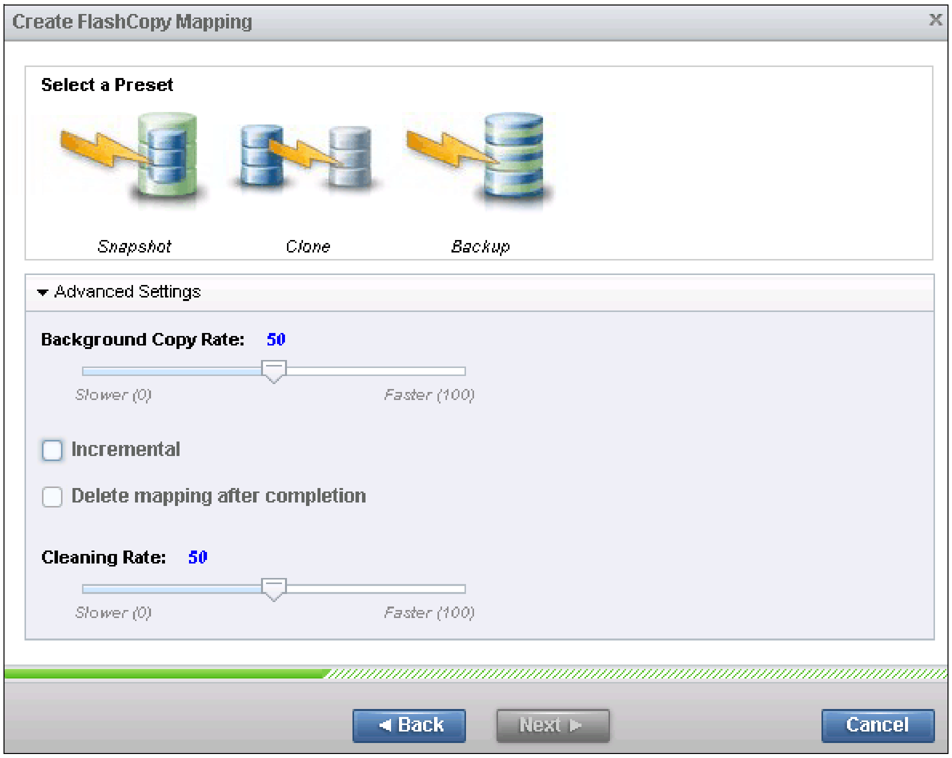


*Figure 11-24 Create FlashCopy Mapping window*

For each preset, you can customize various advanced options. To access these settings, click **Advanced Settings**. The Advanced Setting options show in Figure 11-25.



Background Copy Rate: 50

□ Incremental

□ Delete mapping after completion

Cleaning Rate: 50

*Figure 11-25 Create FlashCopy Mapping Advanced Settings*

If you prefer not to customize these advanced settings, go directly to step 3.

You can customize the advanced setting options that are shown in Figure 11-25:

– Background Copy Rate

– Incremental

> **Incremental FlashCopy mapping:** Even if the type of the FlashCopy mapping is incremental, the first copy process copies all of the data from the source volume to the target volume.

– Delete mapping after completion (This option automatically deletes a FlashCopy mapping after the background copy is completed. Do not use this option when the background copy rate is set to zero).

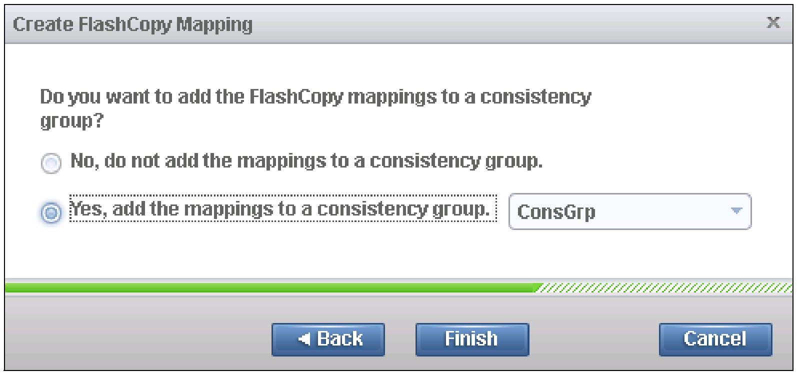
– Cleaning Rate

3. You can choose whether to add this FlashCopy mapping to a Consistency Group or not.

   If you want to include this FlashCopy mapping in a Consistency Group, select **Yes, add the mappings to a consistency group** in the next window (Figure 11-26 on page 450). Select the Consistency Group from the drop-down list.

   If you do not want to include this FlashCopy mapping in a Consistency Group, select **No, do not add the mappings to a consistency group**.

   Click **Finish**.



Create FlashCopy Mapping

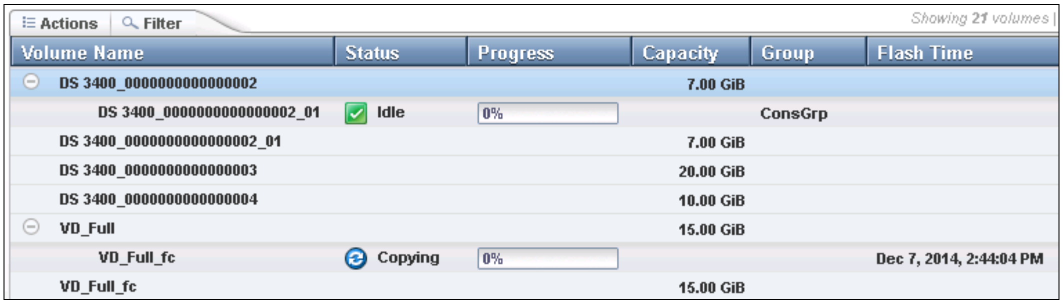
Do you want to add the FlashCopy mappings to a consistency group?

○ No, do not add the mappings to a consistency group.

● Yes, add the mappings to a consistency group. ConsGrp

*Figure 11-26 Selecting the option to add the mappings to a Consistency Group*

4. Check the result of this FlashCopy mapping, as shown in Figure 11-27. For each FlashCopy mapping relationship that is created, a mapping name is automatically generated that starts with `fcmapX` where *X* is the next available number. If necessary, you can rename these mappings, as shown in Figure 11-27. For more information, see Renaming FlashCopy mapping.



| Volume Name | Status | Progress | Capacity | Group | Flash Time |
|---|---|---|---|---|---|
| DS 3400_0000000000000002 | | | 7.00 GiB | | |
| DS 3400_0000000000000002_01 | ✓ Idle | 0% | | ConsGrp | |
| DS 3400_0000000000000002_01 | | | 7.00 GiB | | |
| DS 3400_0000000000000003 | | | 20.00 GiB | | |
| DS 3400_0000000000000004 | | | 10.00 GiB | | |
| VD_Full | | | 15.00 GiB | | |
| VD_Full_fc | Copying | 0% | | | Dec 7, 2014, 2:44:04 PM |
| VD_Full_fc | | | 15.00 GiB | | |

*Figure 11-27 FlashCopy mapping*

The FlashCopy mapping is ready for use.

> **Tip:** You can start FlashCopy from the GUI. However, the use of the GUI might be impractical if you plan to handle many FlashCopy mappings or Consistency Groups periodically, or at varying times. In these cases, creating a script by using the CLI might be more convenient.

## 11.4.2 Single-click snapshot

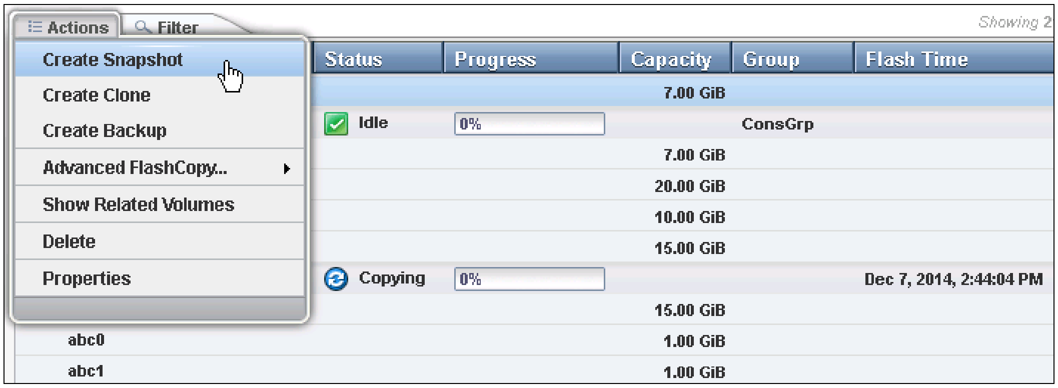
The *snapshot* creates a point-in-time backup of production data. The snapshot is not intended to be an independent copy. Instead, it is used to maintain a view of the production data at the time that the snapshot is created. Therefore, the snapshot holds only the data from regions of the production volume that changed since the snapshot was created. Because the snapshot preset uses thin provisioning, only the capacity that is required for the changes is used.

Snapshot uses the following preset parameters:

- Background copy: No
- Incremental: No
- Delete after completion: No
- Cleaning rate: No
- Primary copy source pool: Target pool

To create and start a snapshot, complete the following steps:

1. From the System pane, move the mouse pointer over Copy Services in the dynamic menu and click **FlashCopy**.
2. Select the volume that you want to create a snapshot of and click **Actions** → **Create Snapshot**, as shown in Figure 11-28.

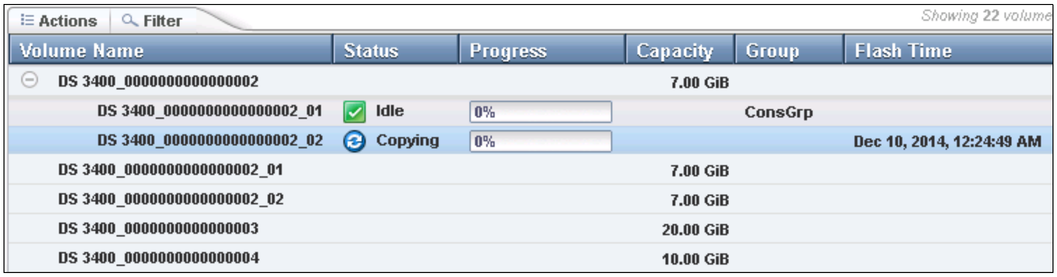


| | Status | Progress | Capacity | Group | Flash Time |
|---|---|---|---|---|---|
| | | | 7.00 GiB | | |
| | Idle | 0% | | ConsGrp | |
| | | | 7.00 GiB | | |
| | | | 20.00 GiB | | |
| | | | 10.00 GiB | | |
| | | | 15.00 GiB | | |
| | Copying | 0% | | | Dec 7, 2014, 2:44:04 PM |
| | | | 15.00 GiB | | |
| abc0 | | | 1.00 GiB | | |
| abc1 | | | 1.00 GiB | | |

*Figure 11-28 Create Snapshot option*

3. A volume is created as a target volume for this snapshot in the same pool as the source volume. The FlashCopy mapping is created and started.

   You can check the FlashCopy progress in the Progress column Status area, as shown in Figure 11-29.



| Volume Name | Status | Progress | Capacity | Group | Flash Time |
|---|---|---|---|---|---|
| DS 3400_0000000000000002 | | | 7.00 GiB | | |
| DS 3400_0000000000000002_01 | Idle | 0% | | ConsGrp | |
| DS 3400_0000000000000002_02 | Copying | 0% | | | Dec 10, 2014, 12:24:49 AM |
| DS 3400_0000000000000002_01 | | | 7.00 GiB | | |
| DS 3400_0000000000000002_02 | | | 7.00 GiB | | |
| DS 3400_0000000000000003 | | | 20.00 GiB | | |
| DS 3400_0000000000000004 | | | 10.00 GiB | | |

*Figure 11-29 Snapshot created and started*

## 11.4.3 Single-click clone

The *clone preset* creates an exact replica of the volume, which can be changed without affecting the original volume. After the copy completes, the mapping that was created by the preset is automatically deleted.

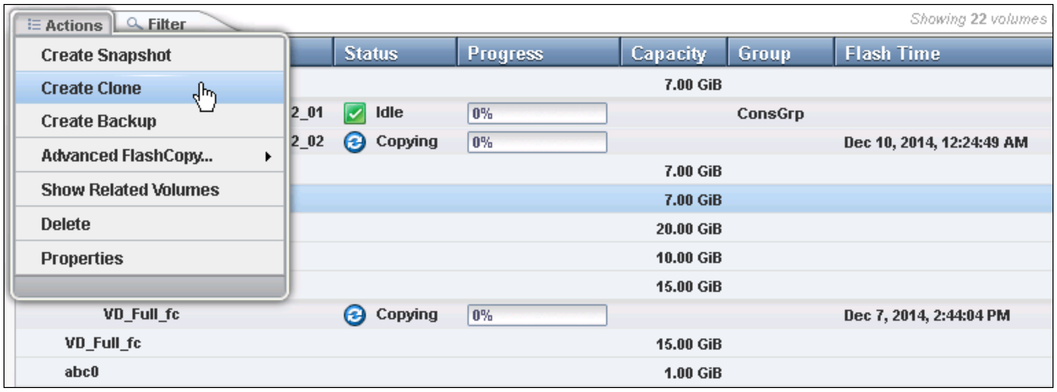
The clone preset uses the following parameters:

- Background copy rate: 50
- Incremental: No
- Delete after completion: Yes
- Cleaning rate: 50
- Primary copy source pool: Target pool

To create and start a clone, complete the following steps:

1. From the System pane, move the mouse pointer over Copy Services in the dynamic menu and click **FlashCopy**.

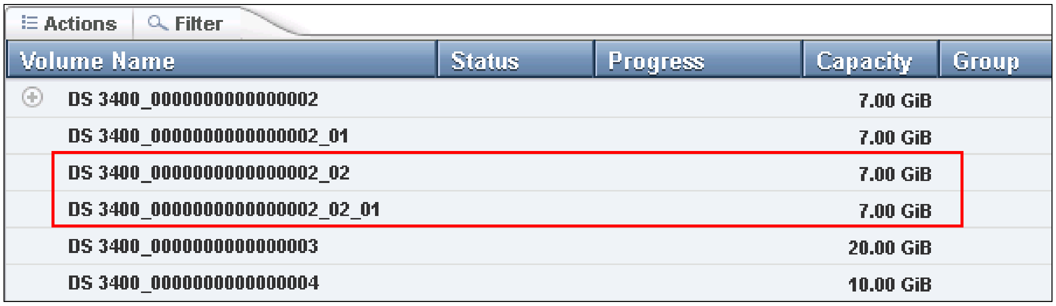
2. Select the volume that you want to clone.
3. Click **Actions** → **Create Clone**, as shown in Figure 11-30.



| | Status | Progress | Capacity | Group | Flash Time |
|---|---|---|---|---|---|
| | | | 7.00 GiB | | |
| 2_01 | Idle | 0% | | ConsGrp | |
| 2_02 | Copying | 0% | | | Dec 10, 2014, 12:24:49 AM |
| | | | 7.00 GiB | | |
| | | | 7.00 GiB | | |
| | | | 20.00 GiB | | |
| | | | 10.00 GiB | | |
| | | | 15.00 GiB | | |
| VD_Full_fc | Copying | 0% | | | Dec 7, 2014, 2:44:04 PM |
| VD_Full_fc | | | 15.00 GiB | | |
| abc0 | | | 1.00 GiB | | |

*Figure 11-30 Create Clone option*

4. A volume is created as a target volume for this clone in the same pool as the source volume. The FlashCopy mapping is created and started. You can check the FlashCopy progress in the Progress column or in the Running Tasks Status column. After the FlashCopy clone is created, the mapping is removed and the new cloned volume becomes available, as shown in Figure 11-31.



| Volume Name | Status | Progress | Capacity | Group |
|---|---|---|---|---|
| DS 3400_0000000000000002 | | | 7.00 GiB | |
| DS 3400_0000000000000002_01 | | | 7.00 GiB | |
| DS 3400_0000000000000002_02 | | | 7.00 GiB | |
| DS 3400_0000000000000002_02_01 | | | 7.00 GiB | |
| DS 3400_0000000000000003 | | | 20.00 GiB | |
| DS 3400_0000000000000004 | | | 10.00 GiB | |

*Figure 11-31 Clone created and FlashCopy relationship removed*

## 11.4.4 Single-click backup

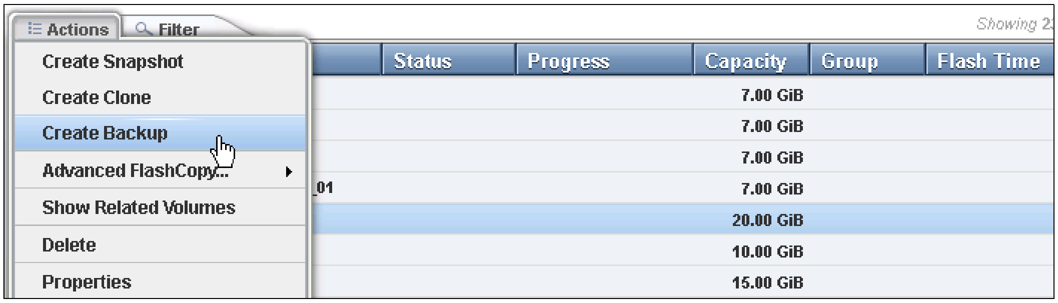
The backup creates a point-in-time replica of the production data. After the copy completes, the backup view can be refreshed from the production data, with minimal copying of data from the production volume to the backup volume.

The backup preset uses the following parameters:

- Background Copy rate: 50
- Incremental: Yes
- Delete after completion: No
- Cleaning rate: 50
- Primary copy source pool: Target pool

To create and start a backup, complete the following steps:

1. From the System pane, move the mouse pointer over Copy Services in the dynamic menu and then click **FlashCopy**.
2. Select the volume that you want to back up.
3. Click **Actions** → **Create Backup**, as shown in Figure 11-32.

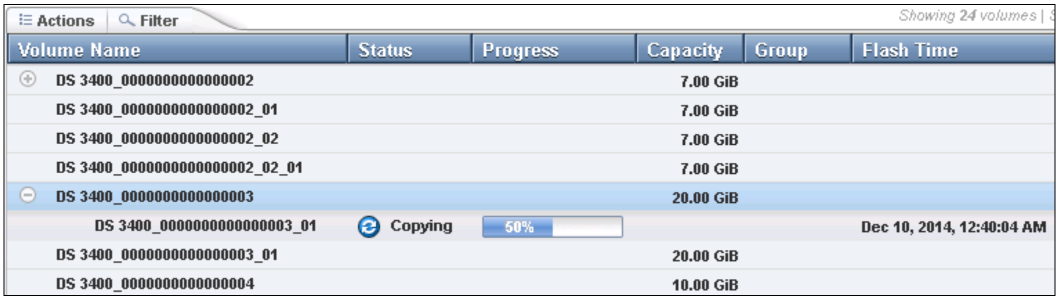


| | Status | Progress | Capacity | Group | Flash Time |
|---|---|---|---|---|---|
| | | | 7.00 GiB | | |
| | | | 7.00 GiB | | |
| | | | 7.00 GiB | | |
| _01 | | | 7.00 GiB | | |
| | | | 20.00 GiB | | |
| | | | 10.00 GiB | | |
| | | | 15.00 GiB | | |

*Figure 11-32 Create Backup option*

4. A volume is created as a target volume for this backup in the same pool as the source volume. The FlashCopy mapping is created and started.

   You can check the FlashCopy progress in the Progress column, as shown in Figure 11-33, or in the Running Tasks Status column.



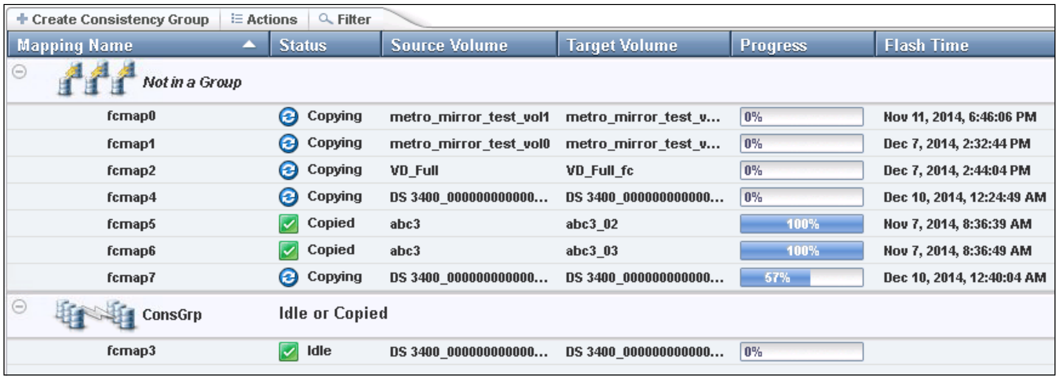
| Volume Name | Status | Progress | Capacity | Group | Flash Time |
|---|---|---|---|---|---|
| DS 3400_0000000000000002 | | | 7.00 GiB | | |
| DS 3400_0000000000000002_01 | | | 7.00 GiB | | |
| DS 3400_0000000000000002_02 | | | 7.00 GiB | | |
| DS 3400_0000000000000002_02_01 | | | 7.00 GiB | | |
| DS 3400_0000000000000003 | | | 20.00 GiB | | |
| DS 3400_0000000000000003_01 | Copying | 50% | | | Dec 10, 2014, 12:40:04 AM |
| DS 3400_0000000000000003_01 | | | 20.00 GiB | | |
| DS 3400_0000000000000004 | | | 10.00 GiB | | |

*Figure 11-33 Backup created and started*

## 11.4.5 Creating a FlashCopy Consistency Group

To create a FlashCopy Consistency Group in the GUI, complete the following steps:

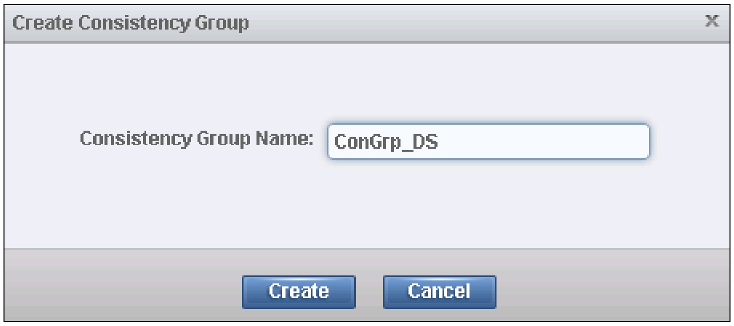
1. From the System pane, move the mouse pointer over Copy Services in the dynamic menu and then click **Consistency Groups**. The Consistency Groups pane opens, as shown in Figure 11-34.



| Mapping Name | Status | Source Volume | Target Volume | Progress | Flash Time |
|---|---|---|---|---|---|
| *Not in a Group* | | | | | |
| fcmap0 | Copying | metro_mirror_test_vol1 | metro_mirror_test_v... | 0% | Nov 11, 2014, 6:46:06 PM |
| fcmap1 | Copying | metro_mirror_test_vol0 | metro_mirror_test_v... | 0% | Dec 7, 2014, 2:32:44 PM |
| fcmap2 | Copying | VD_Full | VD_Full_fc | 0% | Dec 7, 2014, 2:44:04 PM |
| fcmap4 | Copying | DS 3400_000000000000... | DS 3400_000000000000... | 0% | Dec 10, 2014, 12:24:49 AM |
| fcmap5 | Copied | abc3 | abc3_02 | 100% | Nov 7, 2014, 8:36:39 AM |
| fcmap6 | Copied | abc3 | abc3_03 | 100% | Nov 7, 2014, 8:36:49 AM |
| fcmap7 | Copying | DS 3400_000000000000... | DS 3400_000000000000... | 57% | Dec 10, 2014, 12:40:04 AM |
| ConsGrp | Idle or Copied | | | | |
| fcmap3 | Idle | DS 3400_000000000000... | DS 3400_000000000000... | 0% | |

*Figure 11-34 Consistency Groups pane*

2. Click **Create Consistency Group** and enter the FlashCopy Consistency Group name that you want to use and click **Create** (Figure 11-35 on page 454).

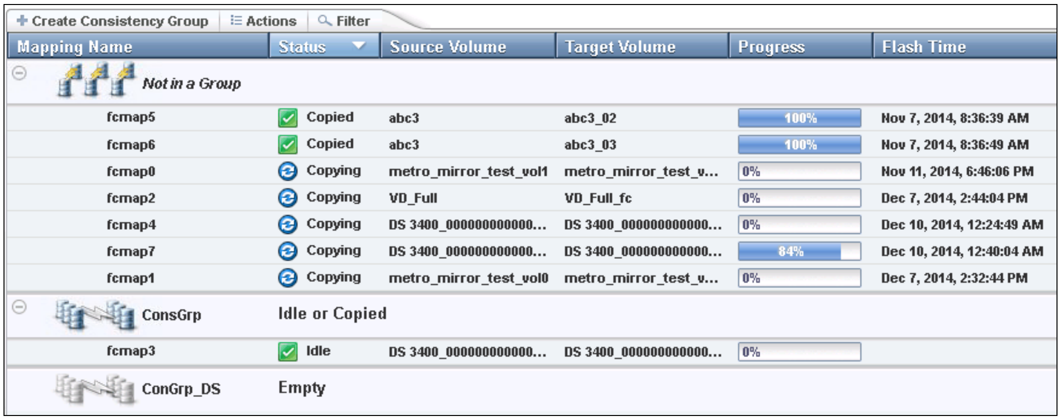


Consistency Group Name: ConGrp_DS

*Figure 11-35 Create Consistency Group window*

> **Consistency Group name:** You can use the letters A - Z and a - z, the numbers 0 - 9, and the underscore (_) character. The volume name can be 1 - 63 characters.

Figure 11-36 shows the result.



| Mapping Name | Status | Source Volume | Target Volume | Progress | Flash Time |
|---|---|---|---|---|---|
| Not in a Group | | | | | |
| fcmap5 | ✓ Copied | abc3 | abc3_02 | 100% | Nov 7, 2014, 8:36:39 AM |
| fcmap6 | ✓ Copied | abc3 | abc3_03 | 100% | Nov 7, 2014, 8:36:49 AM |
| fcmap0 | Copying | metro_mirror_test_vol1 | metro_mirror_test_v... | 0% | Nov 11, 2014, 6:46:06 PM |
| fcmap2 | Copying | VD_Full | VD_Full_fc | 0% | Dec 7, 2014, 2:44:04 PM |
| fcmap4 | Copying | DS 3400_000000000000... | DS 3400_000000000000... | 0% | Dec 10, 2014, 12:24:49 AM |
| fcmap7 | Copying | DS 3400_000000000000... | DS 3400_000000000000... | 84% | Dec 10, 2014, 12:40:04 AM |
| fcmap1 | Copying | metro_mirror_test_vol0 | metro_mirror_test_v... | 0% | Dec 7, 2014, 2:32:44 PM |
| ConsGrp | Idle or Copied | | | | |
| fcmap3 | ✓ Idle | DS 3400_000000000000... | DS 3400_000000000000... | 0% | |
| ConGrp_DS | Empty | | | | |

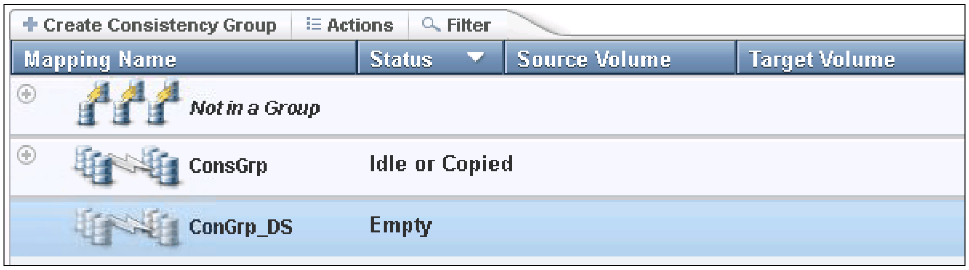
*Figure 11-36 New Consistency Group*

## 11.4.6 Creating FlashCopy mappings in a Consistency Group

In this section, we describe how to create FlashCopy mappings for volumes and their related targets. The source and target volumes were created before this operation.

Complete the following steps:

1. From the System pane, move the mouse pointer over Copy Services in the dynamic menu and then click **Consistency Groups**. The Consistency Groups pane opens, as shown in Figure 11-37 on page 455.
2. Select in which Consistency Group (Figure 11-37) you want to create the FlashCopy mapping. If you prefer not to create a FlashCopy mapping in a Consistency Group, select **Not in a Group**.



| Mapping Name | Status | Source Volume | Target Volume |
|---|---|---|---|
| *Not in a Group* | | | |
| ConsGrp | Idle or Copied | | |
| ConGrp_DS | Empty | | |

Figure 11-37 Consistency Group selection

3. If you select a new Consistency Group, click **Actions** → **Create FlashCopy Mapping**, as shown in Figure 11-38.

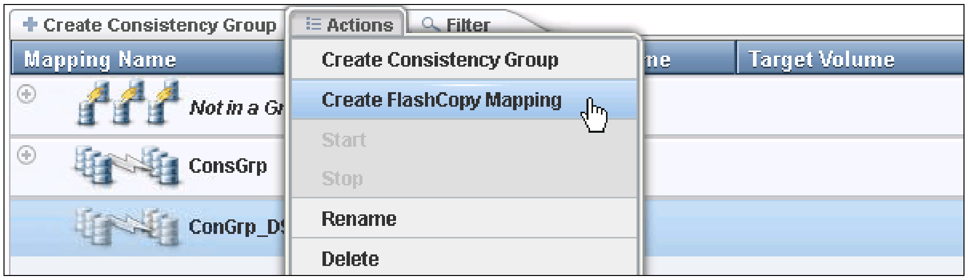
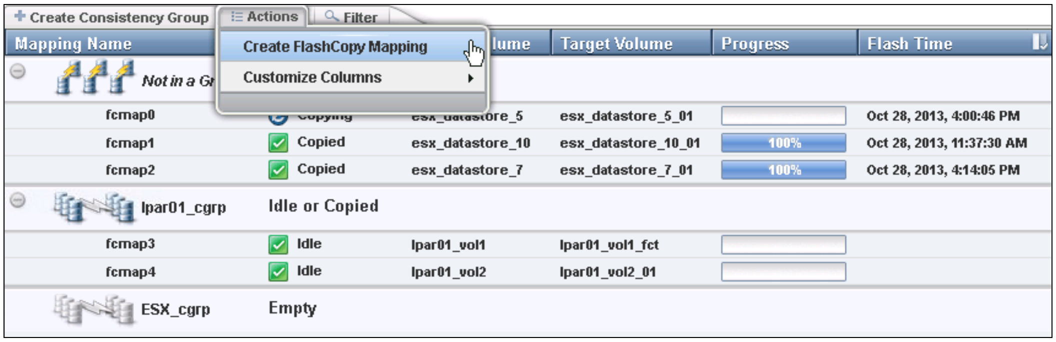


Figure 11-38 Create FlashCopy Mapping action for a Consistency Group

4. If you did not select a Consistency Group, click **Create FlashCopy Mapping**, as shown in Figure 11-39.

**Consistency Groups:** If no Consistency Group is defined, the mapping is a stand-alone mapping. It can be prepared and started without affecting other mappings. All mappings in the same Consistency Group must have the same status to maintain the consistency of the group.



| Mapping Name | | lume | Target Volume | Progress | Flash Time |
|---|---|---|---|---|---|
| *Not in a Gr* | | | | | |
| fcmap0 | Copying | esx_datastore_5 | esx_datastore_5_01 | | Oct 28, 2013, 4:00:46 PM |
| fcmap1 | ✓ Copied | esx_datastore_10 | esx_datastore_10_01 | 100% | Oct 28, 2013, 11:37:30 AM |
| fcmap2 | ✓ Copied | esx_datastore_7 | esx_datastore_7_01 | 100% | Oct 28, 2013, 4:14:05 PM |
| lpar01_cgrp | Idle or Copied | | | | |
| fcmap3 | ✓ Idle | lpar01_vol1 | lpar01_vol1_fct | | |
| fcmap4 | ✓ Idle | lpar01_vol2 | lpar01_vol2_01 | | |
| ESX_cgrp | Empty | | | | |

Figure 11-39 Create FlashCopy Mapping

5. The Create FlashCopy Mapping window opens, as shown in Figure 11-40 on page 456. In this window, you must create the relationships between the source volumes (the volumes that are copied) and the target volumes (the volumes that receive the copy). A mapping can be created between any two volumes in a clustered system.

**Important:** The source volume and the target volume must be of equal size.

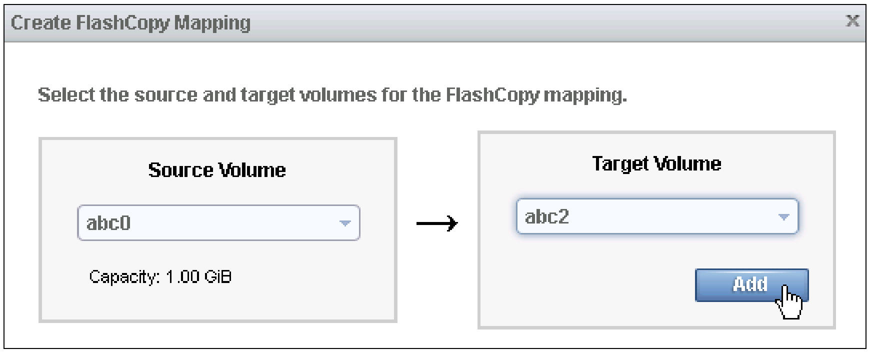


*Figure 11-40 Create FlashCopy Mapping window*

**Tip:** The volumes do not have to be in the same I/O Group or storage pool.

6. Select a volume in the Source Volume column by using the drop-down list. Then, select a volume in the Target Volume column by using the drop-down list. Click **Add**, as shown in Figure 11-40. Repeat this step to create other relationships.

   To remove a relationship that was created, click ✖.

**Important:** The source and target volumes must be of equal size. Therefore, only the targets with the appropriate size are shown for a source volume.

7. Click **Next** after all of the relationships that you want to create are shown (Figure 11-41).

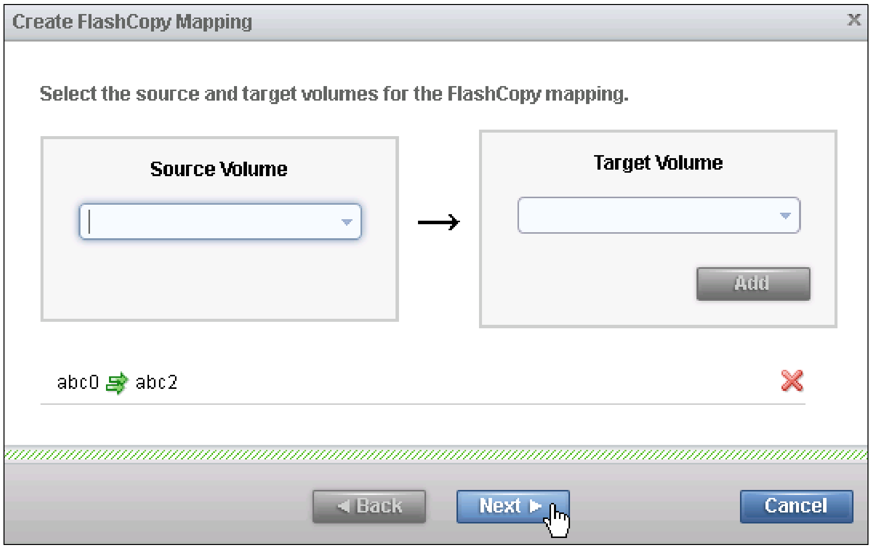


*Figure 11-41 Create FlashCopy Mapping with the relationships that were created*

8. In the next window, you must select one FlashCopy preset. The GUI provides the following presets to simplify common FlashCopy operations, as shown in Figure 11-42 on page 457:

   – Snapshot: Creates a copy-on-write point-in-time copy.

   – Clone: Creates an exact replica of the source volume on a target volume. The copy can be changed without affecting the original volume.

– Backup: Creates a FlashCopy mapping that can be used to recover data or objects if the system experiences data loss. These backups can be copied multiple times from the source and target volumes.

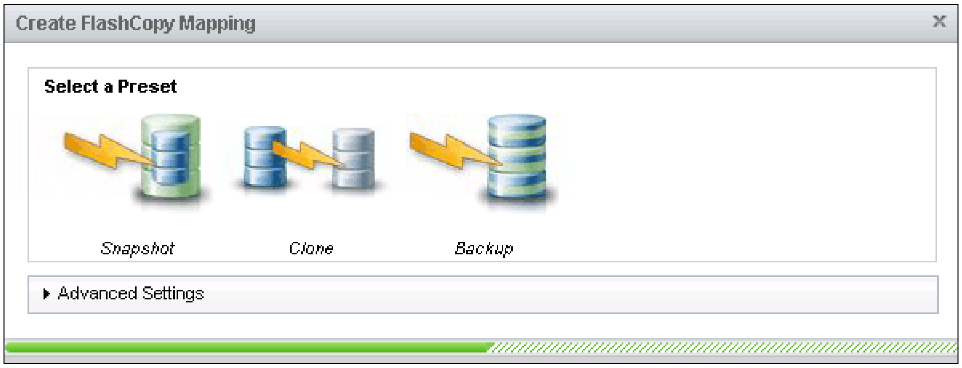
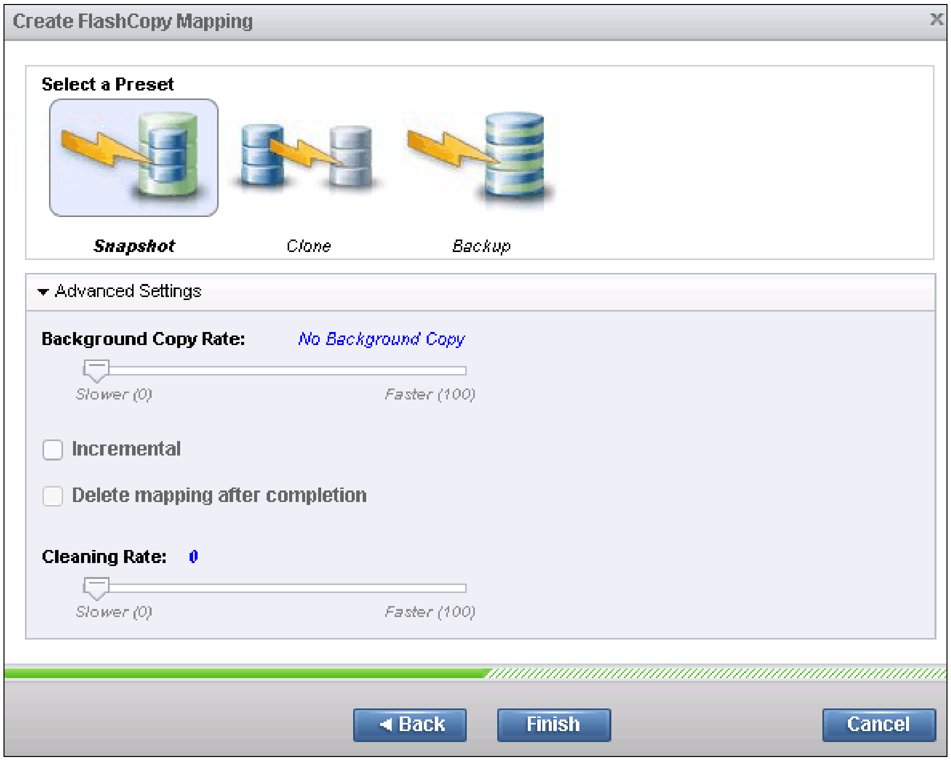


*Figure 11-42 Create FlashCopy Mapping window*

Whichever preset you select, you can customize various advanced options. To access these settings, click **Advanced Settings**.

If you prefer not to customize these settings, go directly to step 9.

You can customize the following advanced setting options, as shown in Figure 11-43:



Background Copy Rate: No Background Copy

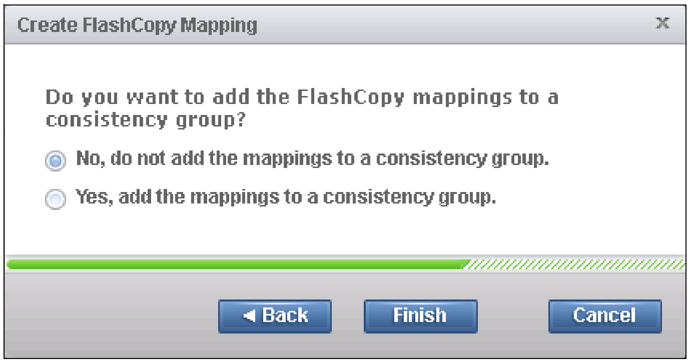
□ Incremental

□ Delete mapping after completion

Cleaning Rate: 0

*Figure 11-43 Create FlashCopy Mapping: Advanced Settings*

9. If you do not want to create these FlashCopy mappings from a Consistency Group (step 3 on page 455), you must confirm your choice by selecting **No, do not add the mappings to a consistency group**, as shown in Figure 11-44.



Do you want to add the FlashCopy mappings to a consistency group?

● No, do not add the mappings to a consistency group.

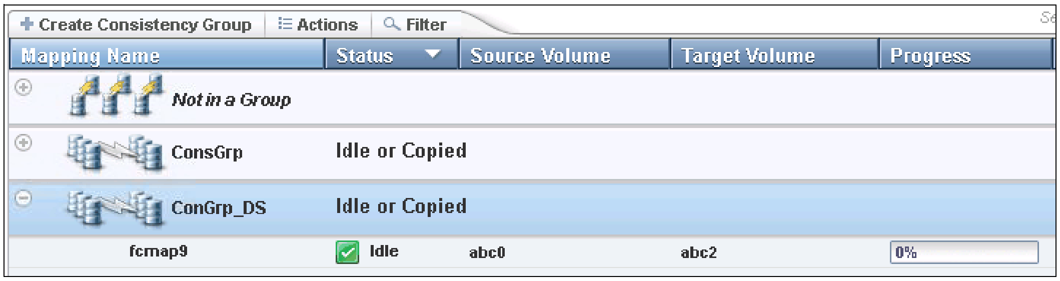
○ Yes, add the mappings to a consistency group.

Figure 11-44 Do not add the mappings to a Consistency Group

10. Click **Finish**.

11. Check the result of this FlashCopy mapping in the Consistency Groups window, as shown in Figure 11-45.

    For each FlashCopy mapping relationship that you created, a mapping name is automatically generated that starts with `fcmapX` where *X* is an available number. If necessary, you can rename these mappings. For more information, see Renaming FlashCopy mapping.



| Mapping Name | Status | Source Volume | Target Volume | Progress |
|---|---|---|---|---|
| *Not in a Group* | | | | |
| ConsGrp | Idle or Copied | | | |
| ConGrp_DS | Idle or Copied | | | |
| fcmap9 | ✓ Idle | abc0 | abc2 | 0% |

Figure 11-45 Create FlashCopy mappings result

> **Tip:** You can start FlashCopy from the IBM Spectrum Virtualize GUI. However, if you plan to handle many FlashCopy mappings or Consistency Groups periodically, or at varying times, creating a script by using the operating system shell CLI might be more convenient.

## 11.4.7 Showing related volumes

Complete the following steps to show related volumes for a specific FlashCopy mapping:

1. From the System pane, move the mouse pointer over Copy Services in the dynamic menu and then click FlashCopy, Consistency Groups, or FlashCopy Mappings.
2. Select the volume (from the FlashCopy pane only) or the FlashCopy mapping that you want to view in this Consistency Group.
3. Click **Actions** → **Show Related Volumes**, as shown in Figure 11-46.

**Tip:** You can also right-click a FlashCopy mapping and select **Show Related Volumes**.

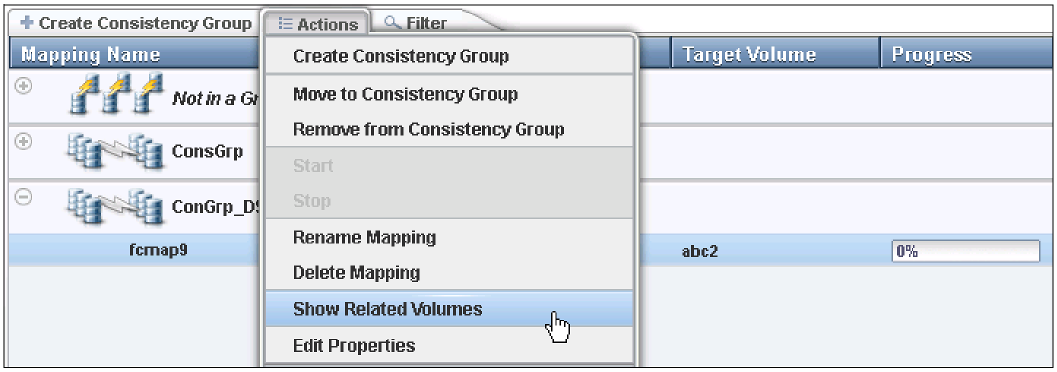


*Figure 11-46 Show Related Volumes*

In the Related Volumes window (Figure 11-47), you can see the related mapping for a volume. If you click one of these volumes, you can see its properties.

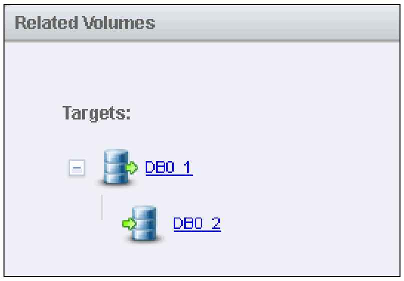


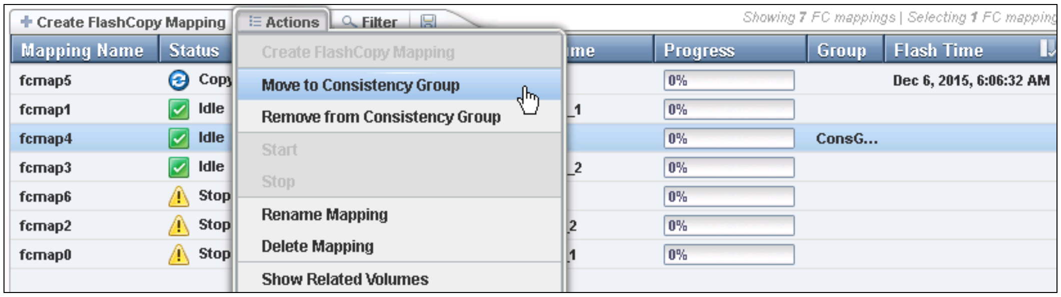
*Figure 11-47 Related Volumes*

## 11.4.8 Moving a FlashCopy mapping to a Consistency Group

Complete the following steps to move a FlashCopy mapping to the Consistency Group:

1. From the System pane, move the mouse pointer over Copy Services in the dynamic menu and then click FlashCopy, Consistency Groups, or FlashCopy Mappings.
2. Select the FlashCopy mapping that you want to move to a Consistency Group or the FlashCopy mapping for which you want to change the Consistency Group.
3. Click **Actions** → **Move to Consistency Group**, as shown in Figure 11-48 on page 460.

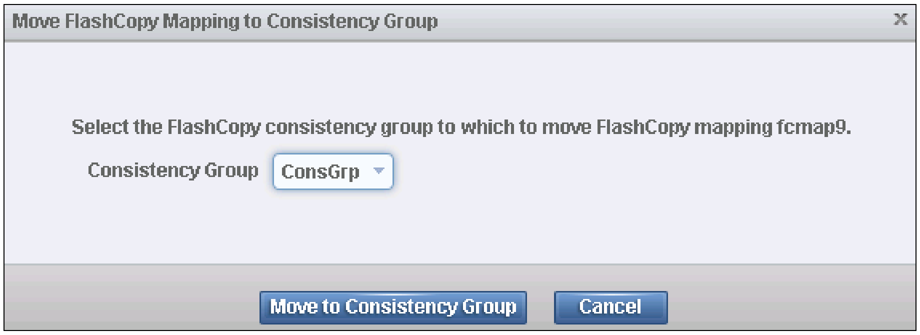
**Tip:** You can also right-click a FlashCopy mapping and select **Move to Consistency Group**.



| Mapping Name | Status | ...me | Progress | Group | Flash Time |
|---|---|---|---|---|---|
| fcmap5 | Copy |  | 0% |  | Dec 6, 2015, 6:06:32 AM |
| fcmap1 | Idle | _1 | 0% |  |  |
| fcmap4 | Idle |  | 0% | ConsG... |  |
| fcmap3 | Idle | _2 | 0% |  |  |
| fcmap6 | Stop |  | 0% |  |  |
| fcmap2 | Stop | 2 | 0% |  |  |
| fcmap0 | Stop | 1 | 0% |  |  |

Figure 11-48 Move to Consistency Group action

4. In the Move FlashCopy Mapping to Consistency Group window, select the Consistency Group for this FlashCopy mapping by using the drop-down list (Figure 11-49).


Move FlashCopy Mapping to Consistency Group
Move to Consistency Group
Cancel

Select the FlashCopy consistency group to which to move FlashCopy mapping fcmap9.

Consistency Group: ConsGrp

Figure 11-49 Move FlashCopy mapping to Consistency Group window

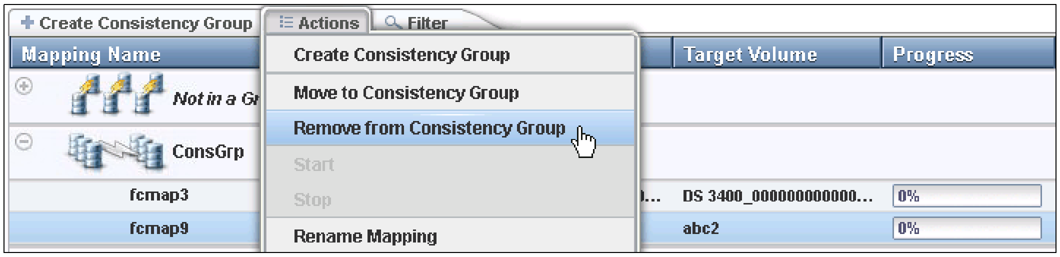
5. Click **Move to Consistency Group** to confirm your changes.

## 11.4.9 Removing a FlashCopy mapping from a Consistency Group

Complete the following steps to remove a FlashCopy mapping from a Consistency Group:

1. From the System pane, move the mouse pointer over Copy Services in the dynamic menu and then click FlashCopy, Consistency Groups, or FlashCopy Mappings.
2. Select the FlashCopy mapping that you want to remove from a Consistency Group.
3. Click **Actions** → **Remove from Consistency Group**, as shown in Figure 11-50.

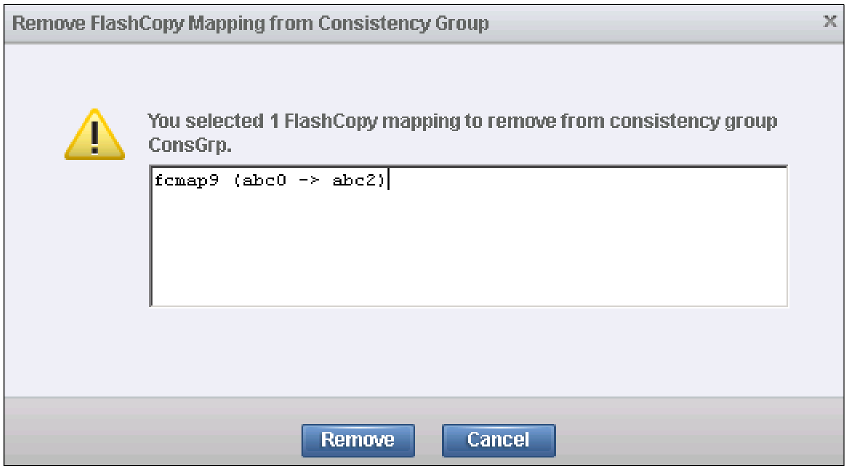
> **Tip:** You can also right-click a FlashCopy mapping and select **Remove from Consistency Group**.



| Mapping Name | Target Volume | Progress |
|---|---|---|
| Not in a Gr |  |  |
| ConsGrp |  |  |
| fcmap3 | DS 3400_000000000000... | 0% |
| fcmap9 | abc2 | 0% |

Figure 11-50 Remove from Consistency Group action

4. In the Remove FlashCopy Mapping from Consistency Group window, click **Remove**, as shown in Figure 11-51.



You selected 1 FlashCopy mapping to remove from consistency group ConsGrp.

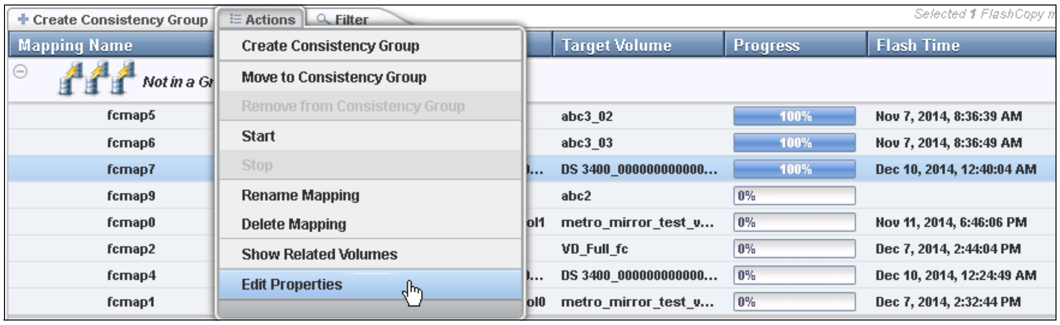
```
fcmap9 (abc0 -> abc2)
```

*Figure 11-51 Remove FlashCopy Mapping from Consistency Group*

## 11.4.10 Modifying a FlashCopy mapping

Complete the following steps to modify a FlashCopy mapping:

1. From the System pane, move the mouse pointer over Copy Services in the dynamic menu and then click FlashCopy, Consistency Groups, or FlashCopy Mappings.
2. In the table, select the FlashCopy mapping that you want to modify.
3. Click **Actions** → **Edit Properties**, as shown in Figure 11-52.

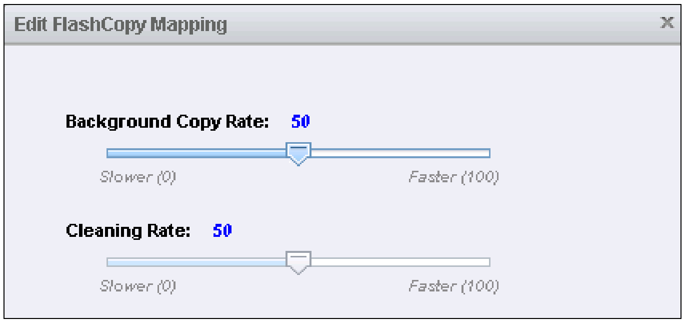


| Mapping Name | | Target Volume | Progress | Flash Time |
|---|---|---|---|---|
| Not in a Gr | | | | |
| fcmap5 | | abc3_02 | 100% | Nov 7, 2014, 8:36:39 AM |
| fcmap6 | | abc3_03 | 100% | Nov 7, 2014, 8:36:49 AM |
| fcmap7 | ... | DS 3400_000000000000... | 100% | Dec 10, 2014, 12:40:04 AM |
| fcmap9 | | abc2 | 0% | |
| fcmap0 | ol1 | metro_mirror_test_v... | 0% | Nov 11, 2014, 6:46:06 PM |
| fcmap2 | | VD_Full_fc | 0% | Dec 7, 2014, 2:44:04 PM |
| fcmap4 | ... | DS 3400_000000000000... | 0% | Dec 10, 2014, 12:24:49 AM |
| fcmap1 | ol0 | metro_mirror_test_v... | 0% | Dec 7, 2014, 2:32:44 PM |

*Figure 11-52 Edit Properties*

**Tip:** You can also right-click a FlashCopy mapping and select **Edit Properties**.

4. In the Edit FlashCopy Mapping window, you can modify the following parameters for a selected FlashCopy mapping, as shown in Figure 11-53 on page 462:
   - Background Copy Rate: This option determines the priority that is given to the copy process. A faster rate increases the priority of the process, which might affect the performance of other operations.
   - Cleaning Rate: This option minimizes the amount of time that a mapping is in the stopping state. If the mapping is not complete, the target volume is offline while the mapping is stopping.



Edit FlashCopy Mapping

Background Copy Rate: 50

Cleaning Rate: 50

Figure 11-53 Edit FlashCopy Mapping

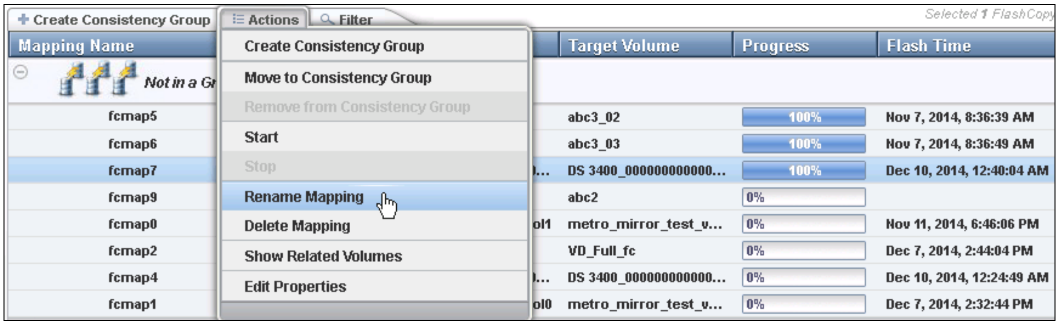
5. Click **Save** to confirm your changes.

## 11.4.11 Renaming FlashCopy mapping

Complete the following steps to rename a FlashCopy mapping:

1. From the System pane, move the mouse pointer over Copy Services in the dynamic menu and then click Consistency Groups or FlashCopy Mappings.
2. In the table, select the FlashCopy mapping that you want to rename.
3. Click **Actions** → **Rename Mapping**, as shown in Figure 11-54.

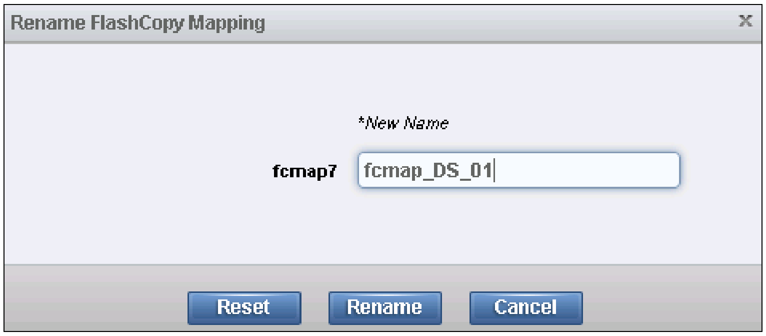
> **Tip:** You can also right-click a FlashCopy mapping and select **Rename Mapping**.



| Mapping Name | | Target Volume | Progress | Flash Time |
|---|---|---|---|---|
| Not in a Gr | | | | |
| fcmap5 | | abc3_02 | 100% | Nov 7, 2014, 8:36:39 AM |
| fcmap6 | | abc3_03 | 100% | Nov 7, 2014, 8:36:49 AM |
| fcmap7 | ... | DS 3400_000000000000... | 100% | Dec 10, 2014, 12:40:04 AM |
| fcmap9 | | abc2 | 0% | |
| fcmap0 | ol1 | metro_mirror_test_v... | 0% | Nov 11, 2014, 6:46:06 PM |
| fcmap2 | | VD_Full_fc | 0% | Dec 7, 2014, 2:44:04 PM |
| fcmap4 | ... | DS 3400_000000000000... | 0% | Dec 10, 2014, 12:24:49 AM |
| fcmap1 | ol0 | metro_mirror_test_v... | 0% | Dec 7, 2014, 2:32:44 PM |

Figure 11-54 Rename Mapping action

4. In the Rename FlashCopy Mapping window, enter the new name that you want to assign to the FlashCopy mapping and click **Rename**, as shown in Figure 11-55.



Rename FlashCopy Mapping

*New Name

fcmap7 fcmap_DS_01

Figure 11-55 Renaming a FlashCopy mapping

> **FlashCopy mapping name:** You can use the letters A - Z and a - z, the numbers 0 - 9, and the underscore (_) character. The FlashCopy mapping name can be 1 - 63 characters.

## 11.4.12 Renaming Consistency Group

To rename a Consistency Group, complete the following steps:

1. From the System pane, move the mouse pointer over Copy Services in the dynamic menu and then click **Consistency Groups**.
2. From the left pane, select the Consistency Group that you want to rename. Then, select **Actions** → **Rename**, as shown in Figure 11-56.

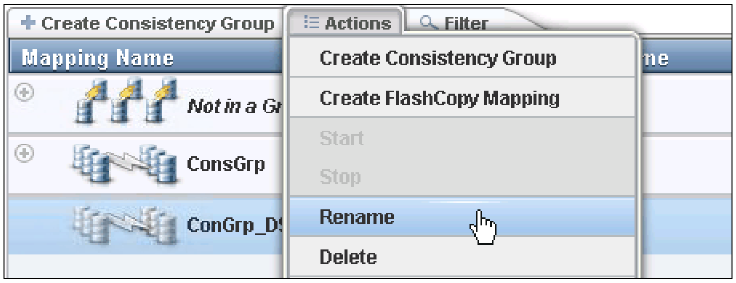
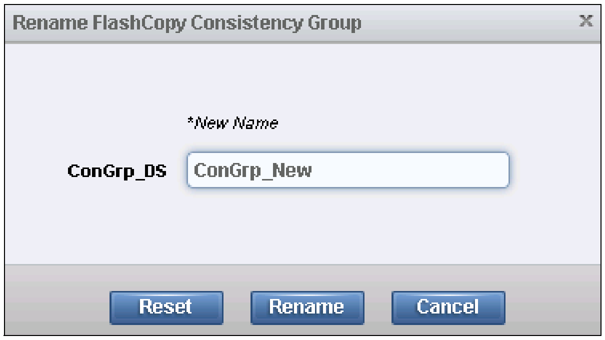


*Figure 11-56 Renaming a Consistency Group*

3. Enter the new name that you want to assign to the Consistency Group and click **Rename**, as shown in Figure 11-57.



*New Name
ConGrp_DS ConGrp_New

*Figure 11-57 Changing the name for a Consistency Group*

> **Consistency Group name:** The name can consist of the letters A - Z and a - z, the numbers 0 - 9, the dash (-), and the underscore (_) character. The name can be 1 - 63 characters. However, the name cannot start with a number, a dash, or an underscore.

The new Consistency Group name is displayed in the Consistency Group pane.

## 11.4.13 Deleting FlashCopy mapping

Complete the following steps to delete a FlashCopy mapping:

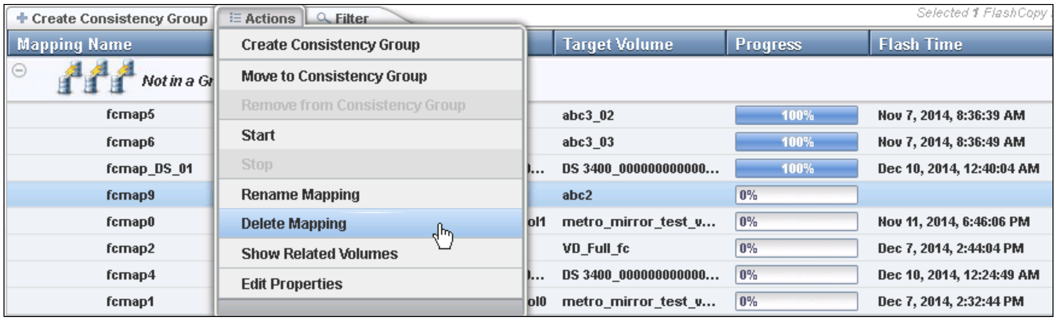
1. From the System pane, move the mouse pointer over Copy Services in the dynamic menu and then click the FlashCopy, Consistency Groups, or FlashCopy Mappings icon.

2. In the table, select the FlashCopy mapping that you want to delete.

   > **Selecting multiple FlashCopy mappings**: To select multiple FlashCopy mappings, hold down Ctrl and click the other entries that you want to delete. This capability is only available in the Consistency Groups pane and the FlashCopy Mappings pane.

3. Click **Actions** → **Delete Mapping**, as shown in Figure 11-58.

   > **Tip**: You can also right-click a FlashCopy mapping and select **Delete Mapping**.



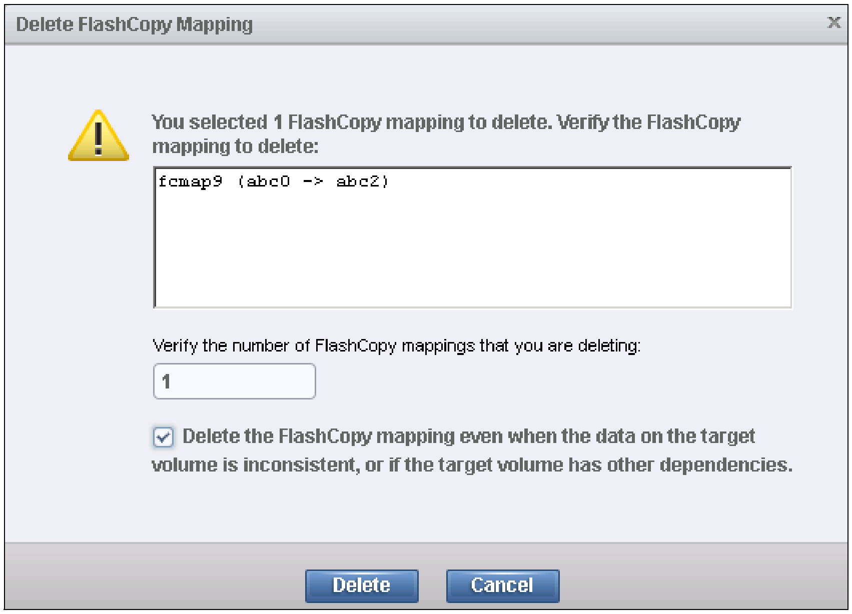
| Mapping Name | | Target Volume | Progress | Flash Time |
|---|---|---|---|---|
| fcmap5 | | abc3_02 | 100% | Nov 7, 2014, 8:36:39 AM |
| fcmap6 | | abc3_03 | 100% | Nov 7, 2014, 8:36:49 AM |
| fcmap_DS_01 | )... | DS 3400_000000000000... | 100% | Dec 10, 2014, 12:40:04 AM |
| fcmap9 | | abc2 | 0% | |
| fcmap0 | ol1 | metro_mirror_test_v... | 0% | Nov 11, 2014, 6:46:06 PM |
| fcmap2 | | VD_Full_fc | 0% | Dec 7, 2014, 2:44:04 PM |
| fcmap4 | )... | DS 3400_000000000000... | 0% | Dec 10, 2014, 12:24:49 AM |
| fcmap1 | ol0 | metro_mirror_test_v... | 0% | Dec 7, 2014, 2:32:44 PM |

*Figure 11-58 Selecting the Delete Mapping option*

4. The Delete FlashCopy Mapping window opens, as shown in Figure 11-59. In the “Verify the number of FlashCopy mappings that you are deleting” field, you must enter the number of volumes that you want to remove. This verification was added to help avoid deleting the wrong mappings.

   If you still have target volumes that are inconsistent with the source volumes and you want to delete these FlashCopy mappings, select **Delete the FlashCopy mapping even when the data on the target volume is inconsistent, or if the target volume has other dependencies**.

   Click **Delete**, as shown in Figure 11-59.



You selected 1 FlashCopy mapping to delete. Verify the FlashCopy mapping to delete:

```
fcmap9 (abc0 -> abc2)
```

Verify the number of FlashCopy mappings that you are deleting:

1

☑ Delete the FlashCopy mapping even when the data on the target volume is inconsistent, or if the target volume has other dependencies.

*Figure 11-59 Delete FlashCopy Mapping*

## 11.4.14 Deleting FlashCopy Consistency Group

**Important:** Deleting a Consistency Group does not delete the FlashCopy mappings.

Complete the following steps to delete a FlashCopy Consistency Group:

1. From the System pane, move the mouse pointer over Copy Services in the dynamic menu and then click **Consistency Groups**.
2. Select the FlashCopy Consistency Group that you want to delete.
3. Click **Actions** → **Delete**, as shown in Figure 11-60.

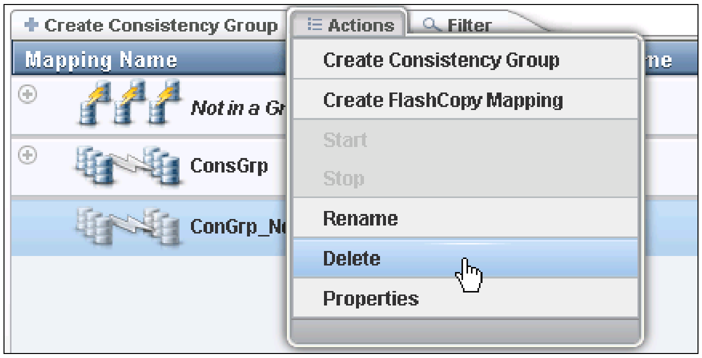
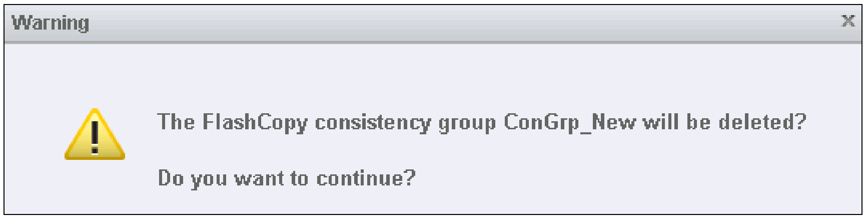


*Figure 11-60 Delete Consistency Group action*

4. The Warning window opens, as shown in Figure 11-61. Click **Yes**.



The FlashCopy consistency group ConGrp_New will be deleted?

Do you want to continue?

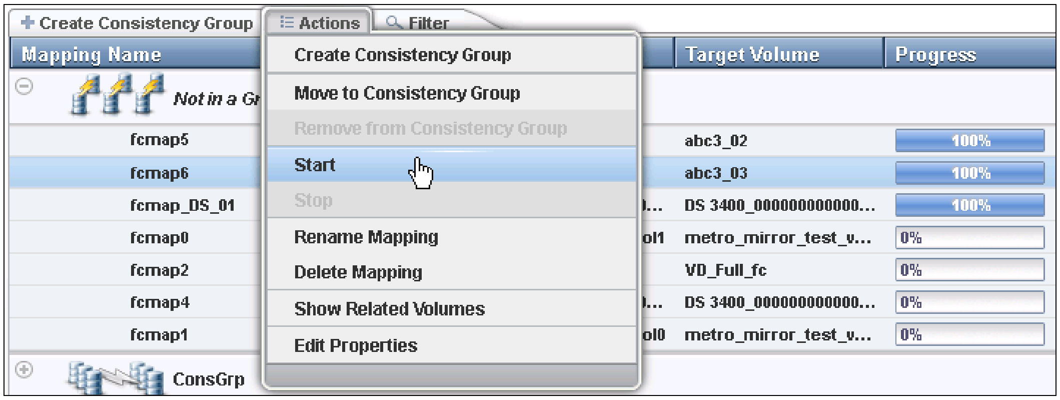
*Figure 11-61 Warning window*

## 11.4.15 Starting FlashCopy process

When the FlashCopy mapping is created, the copy process can be started. Only mappings that are not members of a Consistency Group can be started individually. Complete the following steps:

1. From the System pane, move the mouse pointer over Copy Services in the dynamic menu and then select **FlashCopy Mappings**.
2. In the table, choose the FlashCopy mapping that you want to start.
3. Click **Actions** → **Start** (as shown in Figure 11-62) to start the FlashCopy process.

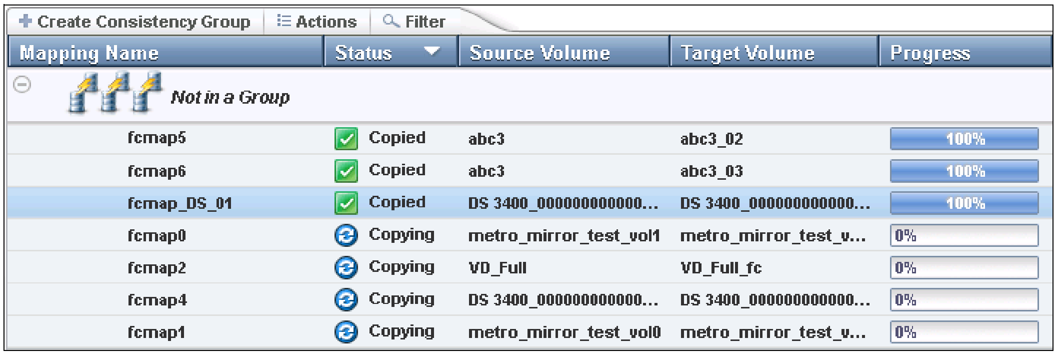
**Tip:** You can also right-click a FlashCopy mapping and select **Start**.



| Mapping Name | | Target Volume | Progress |
|---|---|---|---|
| fcmap5 | | abc3_02 | 100% |
| fcmap6 | | abc3_03 | 100% |
| fcmap_DS_01 | )... | DS 3400_000000000000... | 100% |
| fcmap0 | ol1 | metro_mirror_test_v... | 0% |
| fcmap2 | | VD_Full_fc | 0% |
| fcmap4 | )... | DS 3400_000000000000... | 0% |
| fcmap1 | ol0 | metro_mirror_test_v... | 0% |

*Figure 11-62 Start the FlashCopy process action*

4. You can check the FlashCopy progress in the Progress column of the table or in the Running Tasks status area. After the task completes, the FlashCopy mapping status is in a Copied state, as shown in Figure 11-63.



| Mapping Name | Status | Source Volume | Target Volume | Progress |
|---|---|---|---|---|
| fcmap5 | ✓ Copied | abc3 | abc3_02 | 100% |
| fcmap6 | ✓ Copied | abc3 | abc3_03 | 100% |
| fcmap_DS_01 | ✓ Copied | DS 3400_000000000000... | DS 3400_000000000000... | 100% |
| fcmap0 | Copying | metro_mirror_test_vol1 | metro_mirror_test_v... | 0% |
| fcmap2 | Copying | VD_Full | VD_Full_fc | 0% |
| fcmap4 | Copying | DS 3400_000000000000... | DS 3400_000000000000... | 0% |
| fcmap1 | Copying | metro_mirror_test_vol0 | metro_mirror_test_v... | 0% |

*Figure 11-63 Checking the FlashCopy progress*

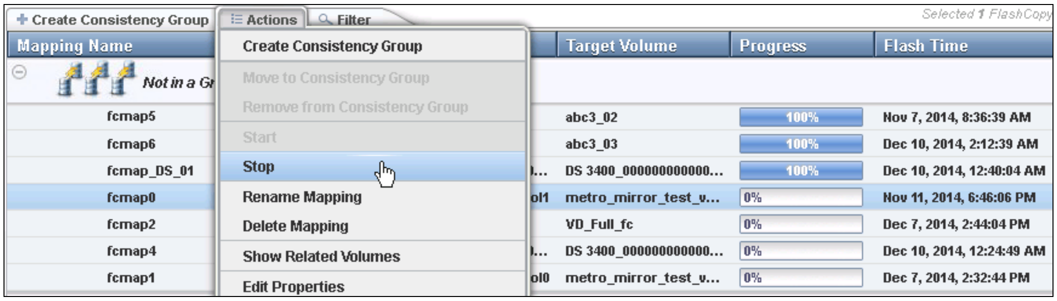
## 11.4.16 Stopping FlashCopy process

When a FlashCopy copy process is stopped, the target volume becomes invalid and it is set offline by the system. The FlashCopy mapping copy must be retriggered to bring the target volume online again.

> **Important:** Stop a FlashCopy copy process only when the data on the target volume is useless, or if you want to modify the FlashCopy mapping. When a FlashCopy mapping is stopped, the target volume becomes invalid and it is set offline by the system.

Complete the following steps to stop a FlashCopy copy process:

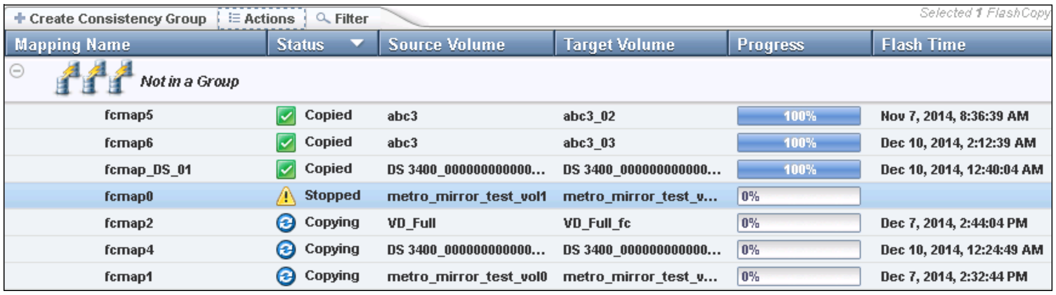
1. From the System pane, move the mouse pointer over Copy Services in the dynamic menu and select **FlashCopy Mappings**.
2. Choose the FlashCopy mapping that you want to stop.
3. Click **Actions** → **Stop** (as shown in Figure 11-64) to stop the FlashCopy Consistency Group copy process.



| Mapping Name | | Target Volume | Progress | Flash Time |
|---|---|---|---|---|
| Not in a Gr | | | | |
| fcmap5 | | abc3_02 | 100% | Nov 7, 2014, 8:36:39 AM |
| fcmap6 | | abc3_03 | 100% | Dec 10, 2014, 2:12:39 AM |
| fcmap_DS_01 | )... | DS 3400_000000000000... | 100% | Dec 10, 2014, 12:40:04 AM |
| fcmap0 | ol1 | metro_mirror_test_v... | 0% | Nov 11, 2014, 6:46:06 PM |
| fcmap2 | | VD_Full_fc | 0% | Dec 7, 2014, 2:44:04 PM |
| fcmap4 | )... | DS 3400_000000000000... | 0% | Dec 10, 2014, 12:24:49 AM |
| fcmap1 | ol0 | metro_mirror_test_v... | 0% | Dec 7, 2014, 2:32:44 PM |

*Figure 11-64 Stopping the FlashCopy copy process*

The FlashCopy Mapping status changes to Stopped, as shown in Figure 11-65.



| Mapping Name | Status | Source Volume | Target Volume | Progress | Flash Time |
|---|---|---|---|---|---|
| Not in a Group | | | | | |
| fcmap5 | Copied | abc3 | abc3_02 | 100% | Nov 7, 2014, 8:36:39 AM |
| fcmap6 | Copied | abc3 | abc3_03 | 100% | Dec 10, 2014, 2:12:39 AM |
| fcmap_DS_01 | Copied | DS 3400_000000000000... | DS 3400_000000000000... | 100% | Dec 10, 2014, 12:40:04 AM |
| fcmap0 | Stopped | metro_mirror_test_vol1 | metro_mirror_test_v... | 0% | |
| fcmap2 | Copying | VD_Full | VD_Full_fc | 0% | Dec 7, 2014, 2:44:04 PM |
| fcmap4 | Copying | DS 3400_000000000000... | DS 3400_000000000000... | 0% | Dec 10, 2014, 12:24:49 AM |
| fcmap1 | Copying | metro_mirror_test_vol0 | metro_mirror_test_v... | 0% | Dec 7, 2014, 2:32:44 PM |

*Figure 11-65 FlashCopy Mapping status*

## 11.5 Transparent Cloud Tiering

Introduced in V7.8, Transparent Cloud Tiering is a licensed function of IBM Spectrum Virtualize running on IBM Storwize V7000, that uses IBM FlashCopy mechanisms to produce a *point-in-time* snapshot of the data. Transparent Cloud Tiering helps to increase the flexibility to protect and transport data to public or private cloud infrastructure. This technology is built on top of IBM Spectrum Virtualize software capabilities. Transparent Cloud Tiering leverages the cloud to store snapshot targets and provides alternatives to restore snapshots from the private and public cloud of an entire volume or set of volumes.

Transparent Cloud Tiering can help to solve business needs that require duplication of data of your source volume. Volumes can remain online and active while you create snapshot copies of the data sets. Transparent Cloud Tiering operates below the host operating system and its cache. Therefore, the copy is not apparent to the host.

IBM Spectrum Virtualize V7.8 has built-in software algorithms that allows the Transparent Cloud Tiering function to securely interact, for example with Information Dispersal Algorithms (IDA) which is essentially the interface to IBM cloud object storage.

*Object storage* is a general term that refers to the entity in which Cloud Object Storage (COS) organize, manage and store with units of storage, or just "objects". To transform these snapshots of traditional data into object storage, the storage nodes and the IDA ingest the data and transforms into a number of metadata and slices, such that object can be read using a subset of those slices. Once an object storage is stored as COS, the objects have to be manipulated or managed as whole unit; therefore objects can't be accessed or updated partially.

IBM Spectrum Virtualize uses internal software components to support HTTP-based REST application programming interface (API) to interact with external cloud service provider or private cloud.

Access the following URL to learn more about the IBM Cloud Object Storage portfolio:

https://ibm.biz/Bdsc7m

## 11.5.1 Considerations for using Transparent Cloud Tiering

Transparent Cloud Tiering can help to address certain business needs. When considering to use Transparent Cloud Tiering, you must adopt a combination of business and technical view of the challenges and determine if Transparent Cloud Tiering can solve the needs.

The use of Transparent Cloud Tiering can assists business to manipulate data as follows:

- Creating a consistent snapshot of dynamically changing data
- Creating a consistent snapshot of production data to facilitate data movement or migration between systems running at different locations
- Creating a snapshot of production data sets for application development and testing
- Creating a snapshot of production data sets for quality assurance
- Securely data tiering to off-premises cloud providers

From the technical standpoint, ensure you evaluate the network capacity and bandwidth requirements to support your data migration to off-premises infrastructure. To maximize productivity, you must match your amount of data that needs to be transmitted off cloud plus the your network capacity.

From the security standpoint, ensure your on-premises or off-premises cloud infrastructure, supports the your requirements in terms of methods and level of encryption.

Regardless of your business needs, Transparent Cloud Tiering within the IBM Spectrum Virtualize may provide opportunities to manage the exponential data growth and to manipulate data at low cost.

Today, many Cloud Service Providers offers a number of *storage-as-services* solutions such as content repository, backup and archive. Combining all of these services, your IBM Spectrum Virtualize can help you solve many challenges related to rapidly data growth, scalability and manageability at attractive costs.

## 11.5.2 Transparent Cloud Tiering as a backup solution and data migration

Transparent Cloud Tiering can also be used as a backup and data migration solution. In certain conditions, can be easily applied to eliminate the downtime that is associated with the needs to import and export data.

When Transparent Cloud Tiering is applied as your backup strategy, IBM Spectrum Virtualize uses the same FlashCopy functions to produce *point-in-time* snapshot of an entire volume or set of volumes.

To ensure the integrity of the snapshot that is made, it might be necessary to flush the host operating system and application cache for any outstanding reads or writes before the snapshot is performed. Failing to flush the host operating system and application cache, may produce inconsistent and useless data.

Many operating systems and applications provide mechanism to stop I/O operations and ensure that all data is flushed from host cache. If these mechanisms are available, they can be used combining with snapshots operation. When these mechanism are not available, it might be necessary to flush the cache manually by quiescing the application and unmounting the file system or logical drives.

When choosing cloud objects storage as backup solution, organizations have to be aware the objects storage must managed as a whole, backup and restore of individual files, folders and partitions, are not possible.

To interact with cloud service providers or private cloud, the IBM Spectrum Virtualize requires interaction to the right architecture and specific properties. Conversely, cloud service providers has assure attractive prices per object storage stored in cloud and easy to use interface. Normally, cloud providers offer low cost prices for object storage space and charges are only applied for the cloud outbound traffic.

## 11.5.3 Restore using Transparent Cloud Tiering

Transparent Cloud Tiering may also be used to restore data from any existing snapshot stored in cloud providers. When the cloud accounts, technical and security requirements are met, the storage objects in the cloud can be used as recovering data solution. The recovery method is similar to backup, instead, the reverse direction is applied. Transparent Cloud Tiering running on IBM Spectrum Virtualize, queries for object storage stored in cloud infrastructure and allows user to restore the objects into a new volume or set of volumes.

This approach can be used for various applications, such as recovering your production database application after an errant batch process that caused extensive damage.

> **Note:** You should always consider the bandwidth characteristics and network capabilities when choosing to use Transparent Cloud Tiering

The restore of individual files using Transparent Cloud Tiering is not possible. As mentioned, objects storage in cloud are unlike a file or block, objects storage must be managed as whole unit piece of storage, not partially. Cloud objects storage are accessible using an HTTP-based REST application programming interface (API).

## 11.5.4 Transparent Cloud Tiering restrictions

This section describes the list of restrictions that must be considered before using Transparent Cloud Tiering.

- Because the object storage are normally accessed via HTTP protocol, on top of TPC/IP stack, all traffic that is associated to cloud service, flows through the node management ports.
- The size of cloud-enabled volumes, cannot change. If the size of the volume change, new snapshot must be created, consequently, new object storage is constructed.
- Transparent Cloud Tiering cannot be applied to volumes that is part of traditional copy services, such as Flash Copy, Metro Mirror, Global Mirror and Hyper Swap.
- Volume containing two physical copies in two different storage pools cannot be part of Transparent Cloud Tiering.
- Cloud Tiering snapshots cannot be taken from volume that is part of migration activity across storage pools.
- Because of VVols are managed by specific VMware application, these volumes are not candidate to Transparent Cloud Tiering.

- File System volumes, such as volumes provisioned by IBM Storwize V7000 Unified platform, are not qualified for Transparent Cloud Tiering.

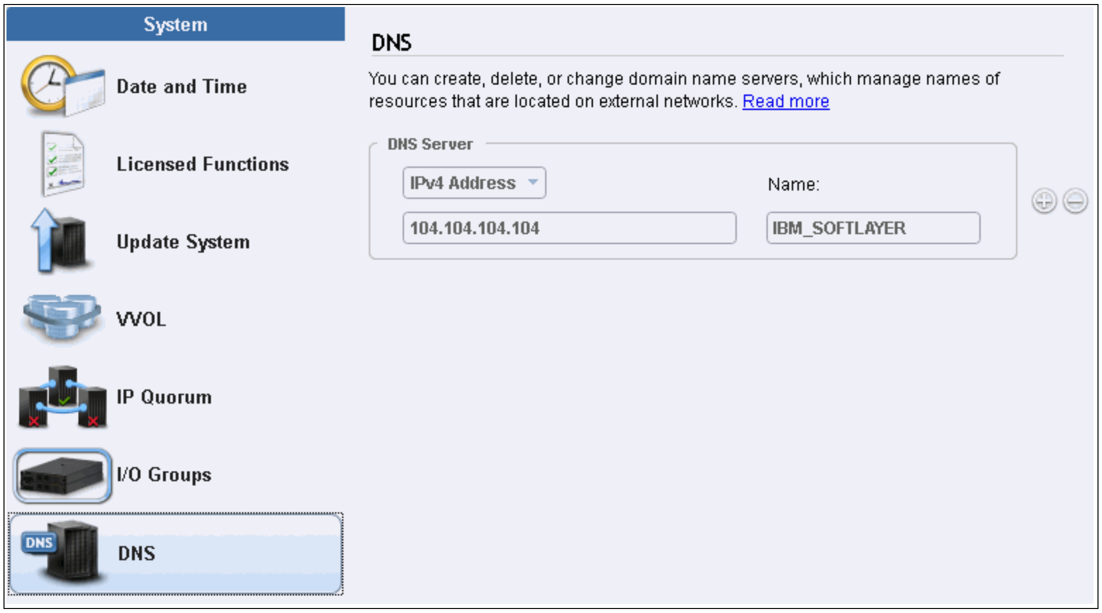
# 11.6 Implementing Transparent Cloud Tiering

This section describes the steps and requirements to implement Transparent Cloud Tiering using your IBM Spectrum Virtualize V7.8.

## 11.6.1 DNS Configuration

Because most of the cloud object storage are managed and accessible via HTTP protocol, DNS (Domain Name System) setting is an important requirement to ensure consistent resolution of domain names to internet resources.

Using your IBM Spectrum Virtualize management GUI, go to **Settings** → **System** → **DNS** and insert your DNS IPv4 or IPv6. The DNS name can be any of your choice that is used as reference. Click **Save** after you complete as shown in Figure 11-66.



DNS

You can create, delete, or change domain name servers, which manage names of resources that are located on external networks. Read more

DNS Server

IPv4 Address: 104.104.104.104

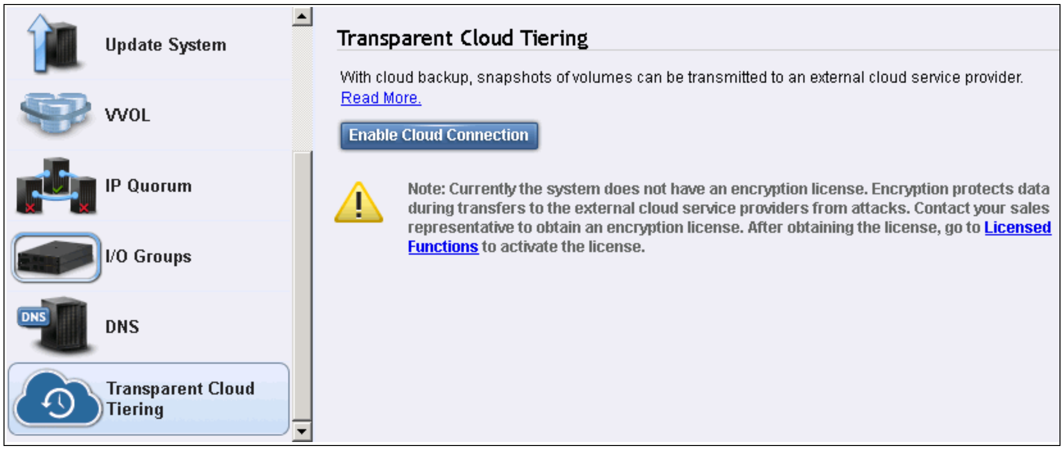
Name: IBM_SOFTLAYER

*Figure 11-66 DNS settings*

## 11.6.2 Enabling Transparent Cloud Tiering

After you complete the DNS settings, you can enable Transparent Cloud Tiering function in your IBM Spectrum Virtualize, by completing the following steps:

1. Using the IBM Spectrum Virtualize GUI, navigate to **Settings** → **System** → **Transparent Cloud Tiering** and click **Enable Cloud Connection** as shown in Figure 11-67 on page 471.



Transparent Cloud Tiering

With cloud backup, snapshots of volumes can be transmitted to an external cloud service provider. Read More.

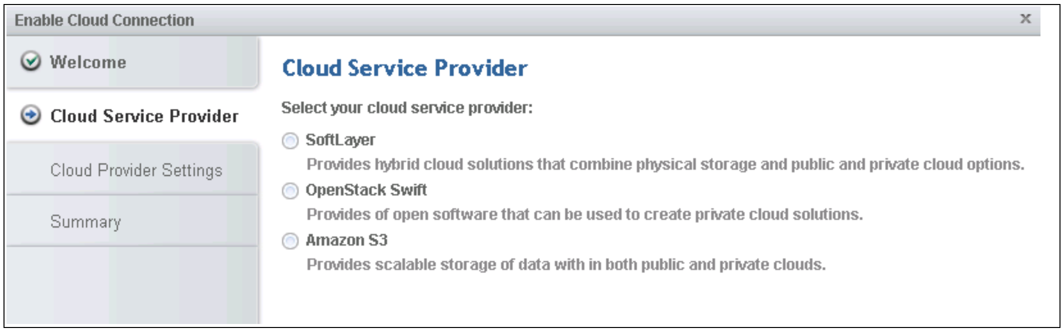
Note: Currently the system does not have an encryption license. Encryption protects data during transfers to the external cloud service providers from attacks. Contact your sales representative to obtain an encryption license. After obtaining the license, go to Licensed Functions to activate the license.

*Figure 11-67 Enabling cloud connection*

2. The Transparent Cloud Tiering wizard starts and shows the welcome warning. It is highly recommended to implement encryption before enabling cloud connecting. Encryption protects your data from attacks, during the transfer to the external cloud service. Because the HTTP protocol is used to connect to cloud infrastructure, it is likely to initiate transactions using the Internet. For purposes of this writing, our system does not have encryption enabled. Click **Next** to continue and you must select one of three cloud service providers:

- Softlayer
- OpenStack Switch
- Amazon S3

Figure 11-68 shows the options available:



Cloud Service Provider

Select your cloud service provider:

○ SoftLayer
Provides hybrid cloud solutions that combine physical storage and public and private cloud options.

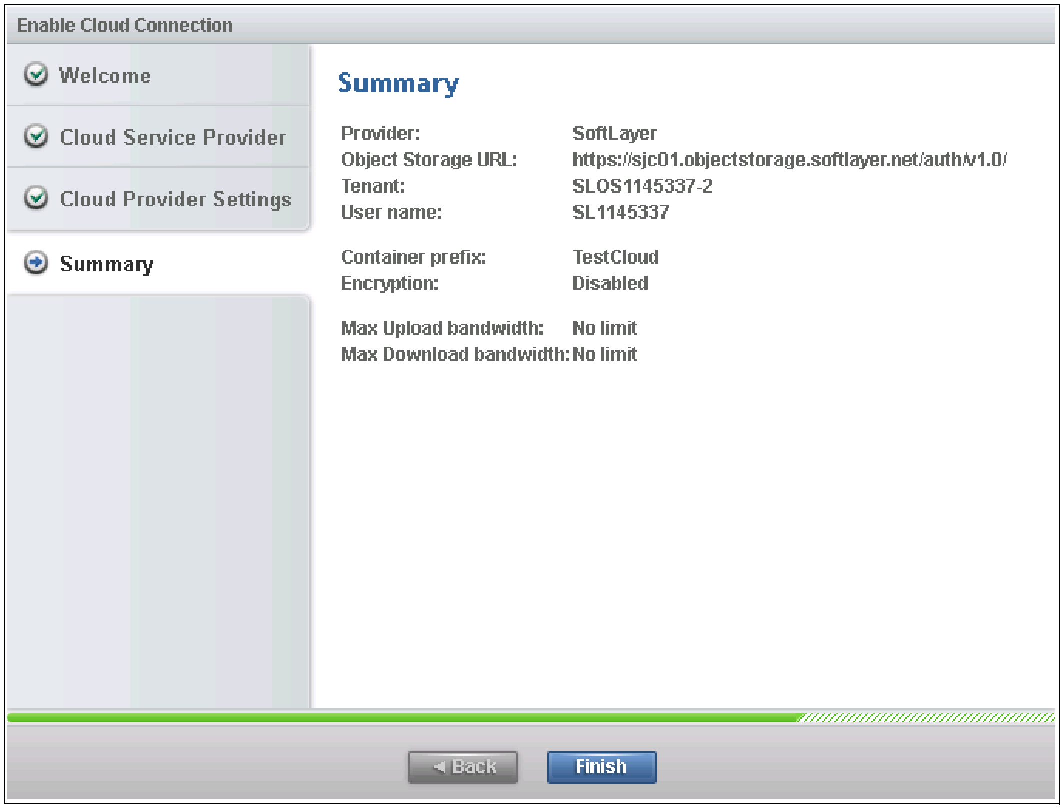
○ OpenStack Swift
Provides of open software that can be used to create private cloud solutions.

○ Amazon S3
Provides scalable storage of data with in both public and private clouds.

*Figure 11-68 Cloud service providers*

3. In the next window, you must complete the settings of the Cloud Provider, credentials and security access keys. The required settings may change depending on your cloud service provider.

4. Review your settings and click **Finish** as shown in Figure 11-69.



Summary

Provider: SoftLayer
Object Storage URL: https://sjc01.objectstorage.softlayer.net/auth/v1.0/
Tenant: SLOS1145337-2
User name: SL1145337

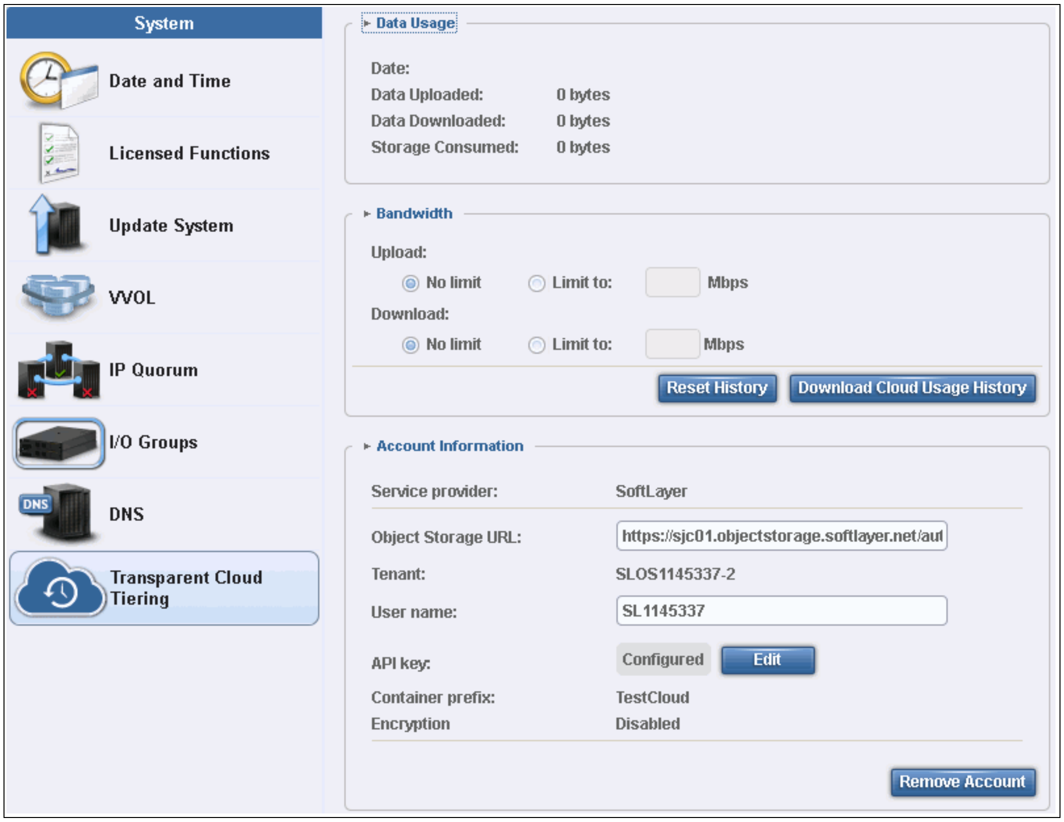
Container prefix: TestCloud
Encryption: Disabled

Max Upload bandwidth: No limit
Max Download bandwidth: No limit

*Figure 11-69 Cloud Settings*

The cloud credentials can be viewed and updated at any time by using the function icons in left side of the GUI and choosing **Settings** → **Systems** → **Transparent Cloud Tiering.** From this panel, you can also verify the status, the data usage statistics as well as the upload and download bandwidth limits set to support this functionality.

In the account information panel, you can visualize your cloud account information. This panel also enables to remove the account. See an example shown in Figure 11-70 on page 473.



Data Usage

Date:
Data Uploaded: 0 bytes
Data Downloaded: 0 bytes
Storage Consumed: 0 bytes

Bandwidth

Upload:
● No limit ○ Limit to: Mbps

Download:
● No limit ○ Limit to: Mbps

Account Information

Service provider: SoftLayer
Object Storage URL: https://sjc01.objectstorage.softlayer.net/aut
Tenant: SLOS1145337-2
User name: SL1145337
API key: Configured
Container prefix: TestCloud
Encryption Disabled

*Figure 11-70 Cloud settings summary*

## 11.6.3 Creating and Managing Cloud Snapshots

To manage the cloud snapshots, the IBM Spectrum Virtualize provides a new section in the GUI named **Cloud Volumes**. This section allows you to add the volumes that are going to be part of the Transparent Cloud Tiering. As described in “Transparent Cloud Tiering restrictions”, cloud snapshot is only available for volumes that do not have relationship to the list of restrictions previously mentioned. Any volume can be added to the Cloud volumes however, snapshots will only work for volumes that are not related to any other copy service.

Navigate to **Volumes** and click **Cloud Volumes** as shown in Figure 11-71.

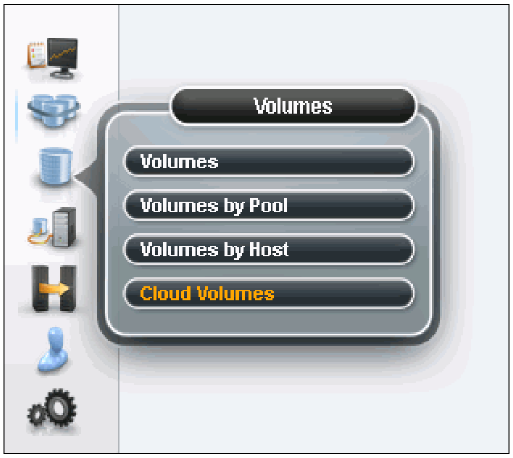
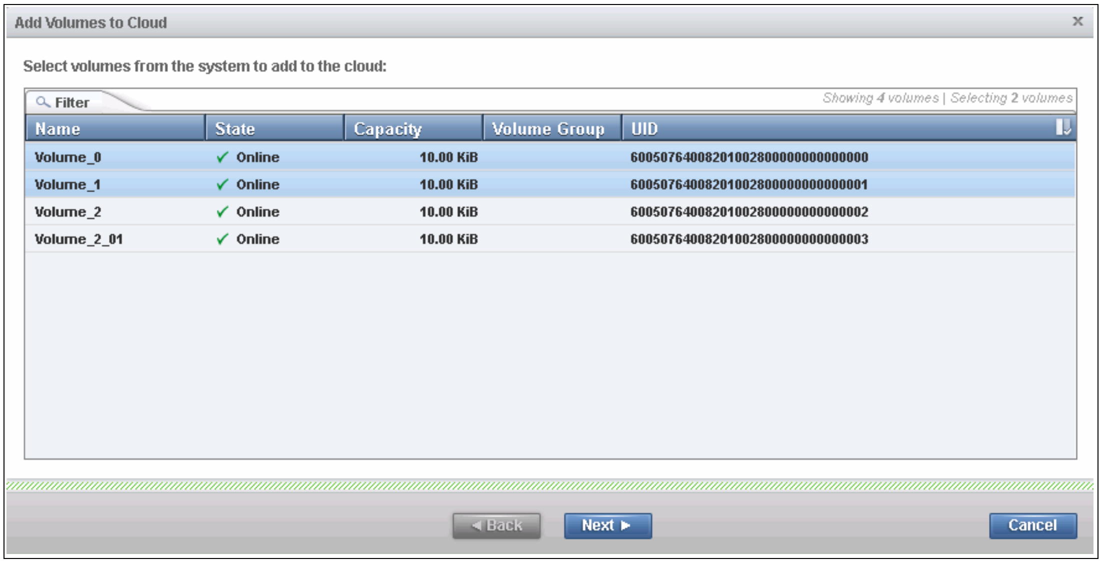


*Figure 11-71 Cloud volumes option*

After **Cloud Volumes** is selected, a new window opens and GUI allows you to select the volume(s) that you need to enable cloud snapshot. If you had enabled cloud snapshot previously, the cloud-enabled volumes are listed in this window. Select the volumes you want to enable cloud-snapshot and click **Next** as shown in Figure 11-72.



Add Volumes to Cloud

Select volumes from the system to add to the cloud:

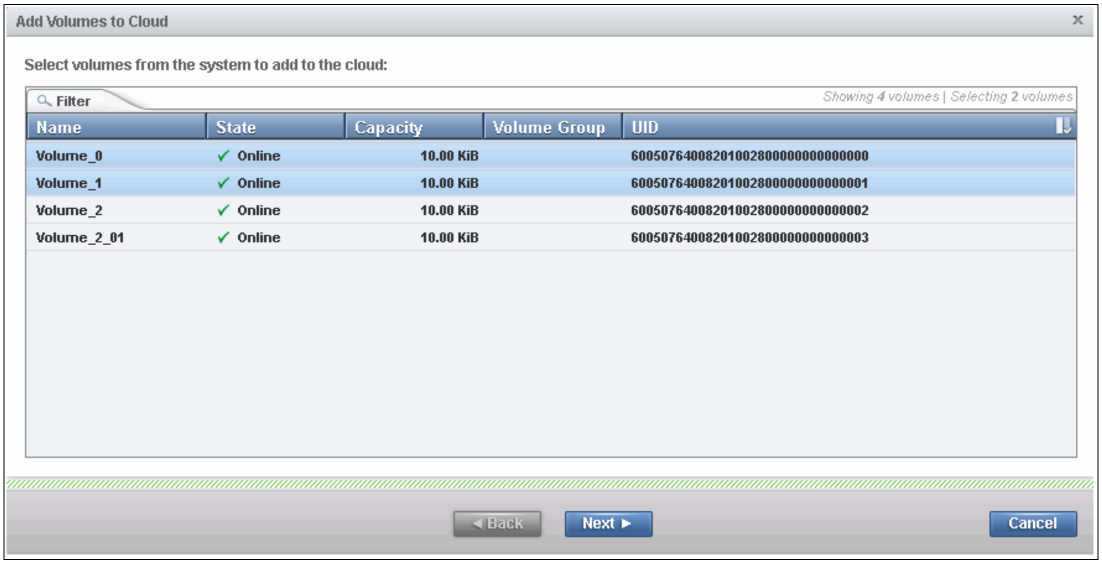
Showing 4 volumes | Selecting 2 volumes

| Name | State | Capacity | Volume Group | UID |
|---|---|---|---|---|
| Volume_0 | ✓ Online | 10.00 KiB | | 60050764008201002800000000000000 |
| Volume_1 | ✓ Online | 10.00 KiB | | 60050764008201002800000000000001 |
| Volume_2 | ✓ Online | 10.00 KiB | | 60050764008201002800000000000002 |
| Volume_2_01 | ✓ Online | 10.00 KiB | | 60050764008201002800000000000003 |

*Figure 11-72 Selecting volumes to enable cloud snapshot*

You must ensure to select the volumes that meet the system requirements described in section “Transparent Cloud Tiering restrictions” otherwise the procedure to perform snapshots of the cloud volume fails.

For purposes of this writing, we are selecting the volumes as shown in Figure 11-73.



Add Volumes to Cloud

Select volumes from the system to add to the cloud:

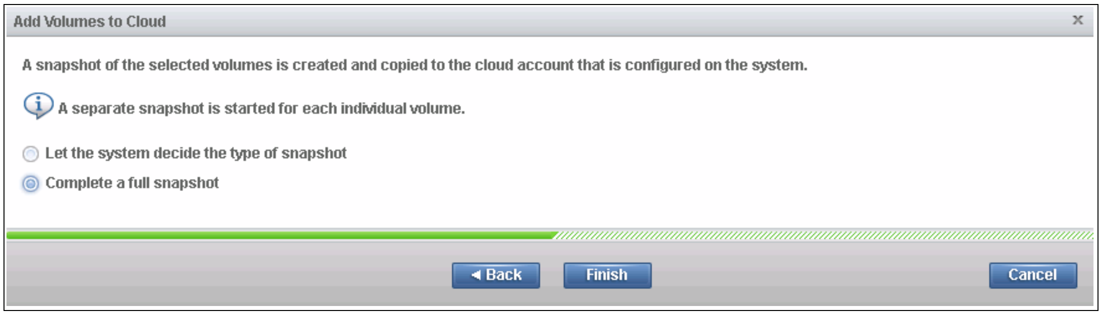
Showing 4 volumes | Selecting 2 volumes

| Name | State | Capacity | Volume Group | UID |
|---|---|---|---|---|
| Volume_0 | ✓ Online | 10.00 KiB | | 60050764008201002800000000000000 |
| Volume_1 | ✓ Online | 10.00 KiB | | 60050764008201002800000000000001 |
| Volume_2 | ✓ Online | 10.00 KiB | | 60050764008201002800000000000002 |
| Volume_2_01 | ✓ Online | 10.00 KiB | | 60050764008201002800000000000003 |

*Figure 11-73 Adding volumes to cloud snapshot*

After you click **Next** the IBM Spectrum Virtualize GUI provides two options for user to select. If the first option is select, the system decides what type of snapshot is created based on previous objects for each selected volume. The second options creates a full snap of the selected volume(s).

Select the second option and click **Finish** as shown in Figure 11-74.



Add Volumes to Cloud

A snapshot of the selected volumes is created and copied to the cloud account that is configured on the system.

A separate snapshot is started for each individual volume.

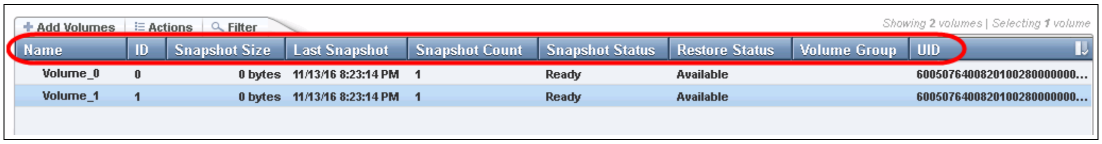
○ Let the system decide the type of snapshot
● Complete a full snapshot

*Figure 11-74 Adding volumes and full snapshot*

The system creates a full snapshot of the selected volumes. Depending on the size of each volume, the GUI shows the cloud snapshot as a running tasks in the lower left of the main window. The **Cloud Volumes** window shows a complete information about the volumes and the snapshots of taken. There are a number of useful information that helps administrators to manage the volumes the snapshots of them. The GUI shows the following info:

- Name of the vdisk (volume name)
- ID of the vdisk assigned by the IBM Spectrum Virtualize
- The size all snapshots taken off the volume
- The date and time that the last snapshot was created
- The number of snapshots taken for every volume
- The snapshot status
- The restore status
- The volume group for a set of volumes
- The vdisk UID

The Figure 11-75 on page 475 shows an example

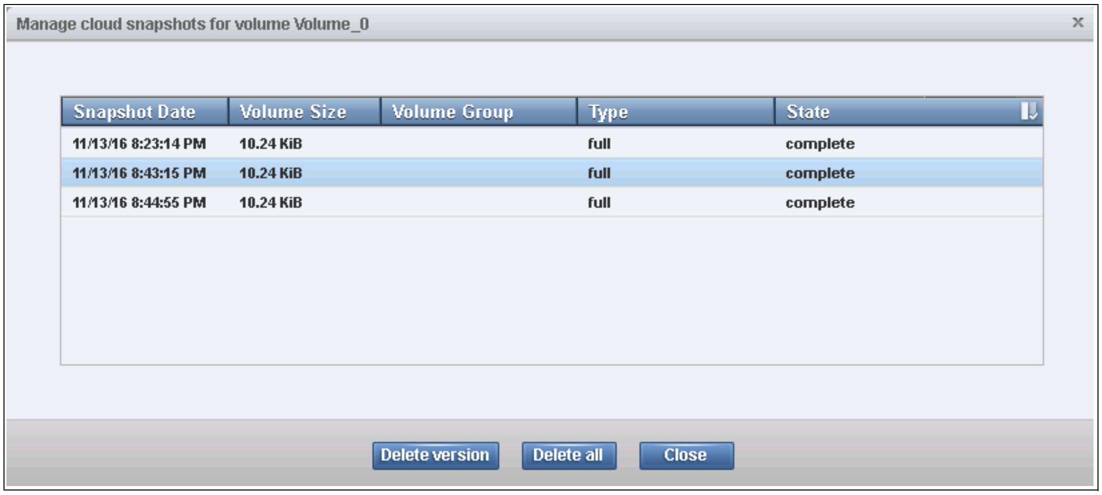


| Name | ID | Snapshot Size | Last Snapshot | Snapshot Count | Snapshot Status | Restore Status | Volume Group | UID |
|---|---|---|---|---|---|---|---|---|
| Volume_0 | 0 | 0 bytes | 11/13/16 8:23:14 PM | 1 | Ready | Available | | 6005076400820100280000000... |
| Volume_1 | 1 | 0 bytes | 11/13/16 8:23:14 PM | 1 | Ready | Available | | 6005076400820100280000000... |

*Figure 11-75 Cloud volumes snapshots*

The **Actions** menu in the same window, allows user to create and to manage snapshots. Also, allows users to cancel, to disable and to restore snapshots to volumes. If **Manage Cloud Snapshot** is selected, the user can view the list of snapshot taken for any particular date and time and delete snapshots objects as example. From this window, an administrator can delete old snapshots that could be no longer valid as data has updated.

Figure 11-76 on page 476 shows the options available under **Manage Cloud Snapshot** option.



| Snapshot Date | Volume Size | Volume Group | Type | State |
|---|---|---|---|---|
| 11/13/16 8:23:14 PM | 10.24 KiB | | full | complete |
| 11/13/16 8:43:15 PM | 10.24 KiB | | full | complete |
| 11/13/16 8:44:55 PM | 10.24 KiB | | full | complete |

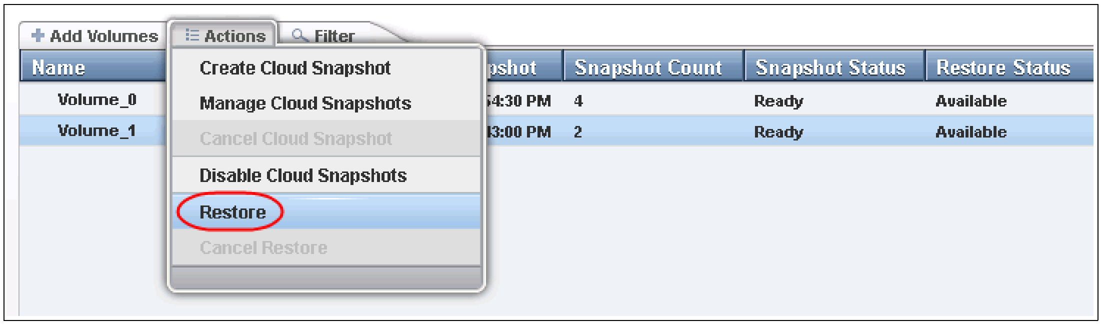
*Figure 11-76 Manage cloud snapshot option*

## 11.6.4 Restore Cloud Snapshots

The restore option allows the IBM Spectrum Virtualize to restore snapshots from cloud to production volume or new volume. When using this method, you must remember the system restore the cloud snapshot to the volume and overwrites the volume entirely.

If the cloud account is shared among systems, the IBM Spectrum Virtualize queries the snapshots stored in the cloud and allows user to restore to a new volume.

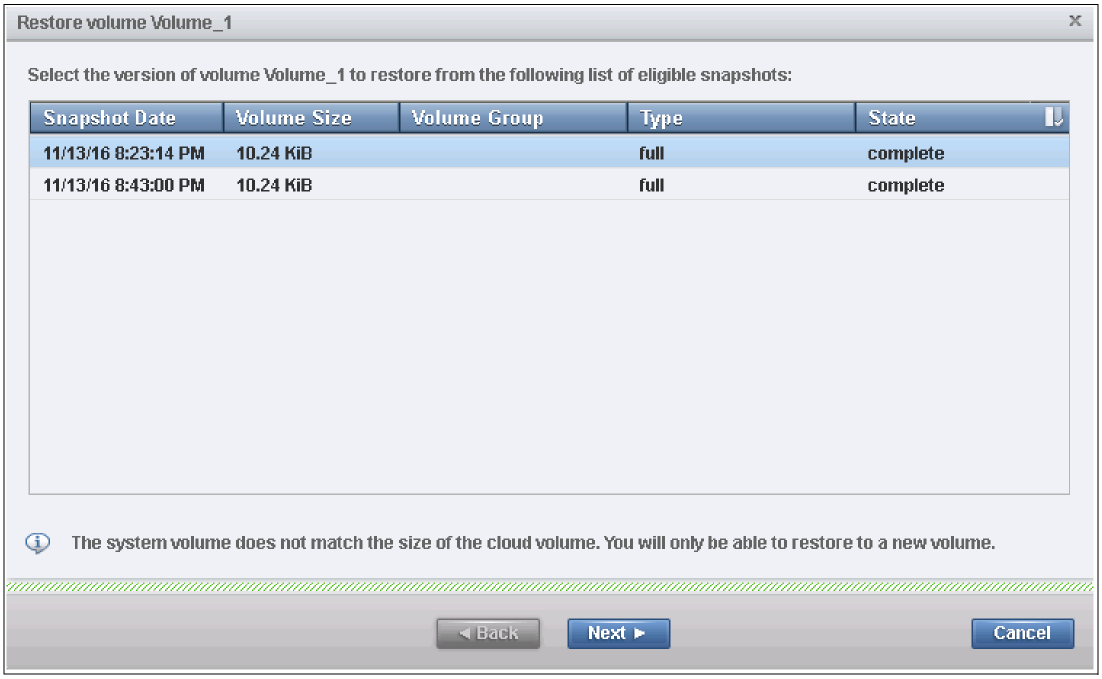
To restore the volume from cloud snapshot using the IBM Spectrum Virtualize, go to **Cloud Volumes** window and select the volume which you want to restore from cloud and click **Action** → **Restore** as shown in Figure 11-77 on page 476.



| Name | pshot | Snapshot Count | Snapshot Status | Restore Status |
|---|---|---|---|---|
| Volume_0 | 54:30 PM | 4 | Ready | Available |
| Volume_1 | 13:00 PM | 2 | Ready | Available |

*Figure 11-77 Restore volume from cloud snapshot*

The GUI shows all available snapshots and enables you to choose the eligible snapshot from the list as shown in Figure 11-78



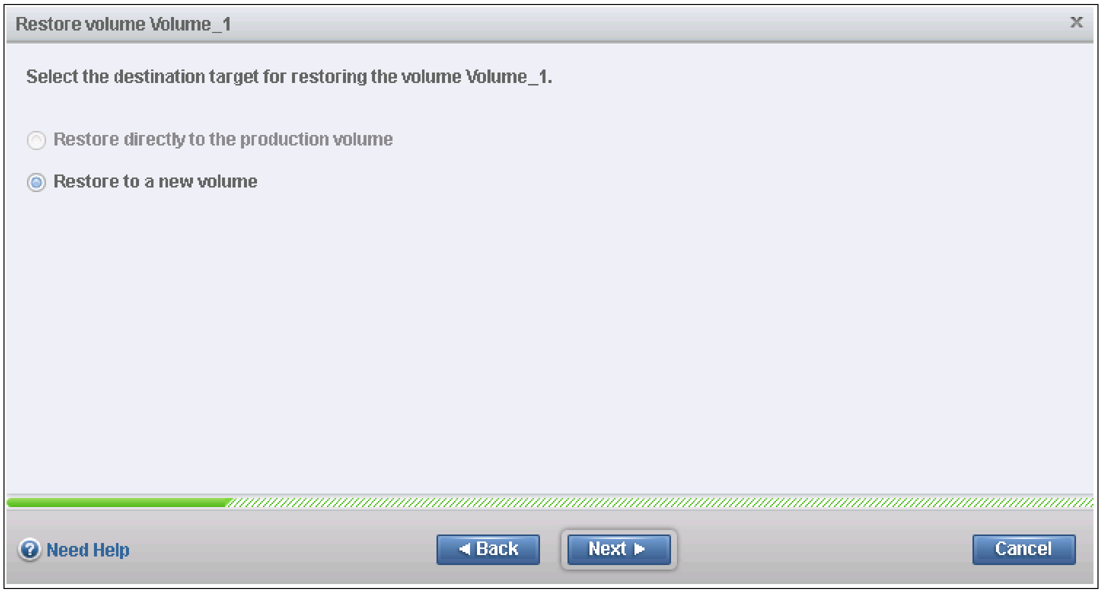
Select the version of volume Volume_1 to restore from the following list of eligible snapshots:

| Snapshot Date | Volume Size | Volume Group | Type | State |
|---|---|---|---|---|
| 11/13/16 8:23:14 PM | 10.24 KiB | | full | complete |
| 11/13/16 8:43:00 PM | 10.24 KiB | | full | complete |

The system volume does not match the size of the cloud volume. You will only be able to restore to a new volume.

Figure 11-78 Restoring volume snapshot from cloud

Select the snapshot version of the volume you wish to restore, and click **Next.** The IBM Spectrum Virtualize GUI provides two options to restore the snapshot from cloud. You may restore the snapshot from cloud directly to the production volume or, the system will create a new volume to receive the data from the snapshot object in the cloud. In both situation, the system will overwrite the volume using snapshot version stored in the cloud. The figure Figure 11-79 shows both options.



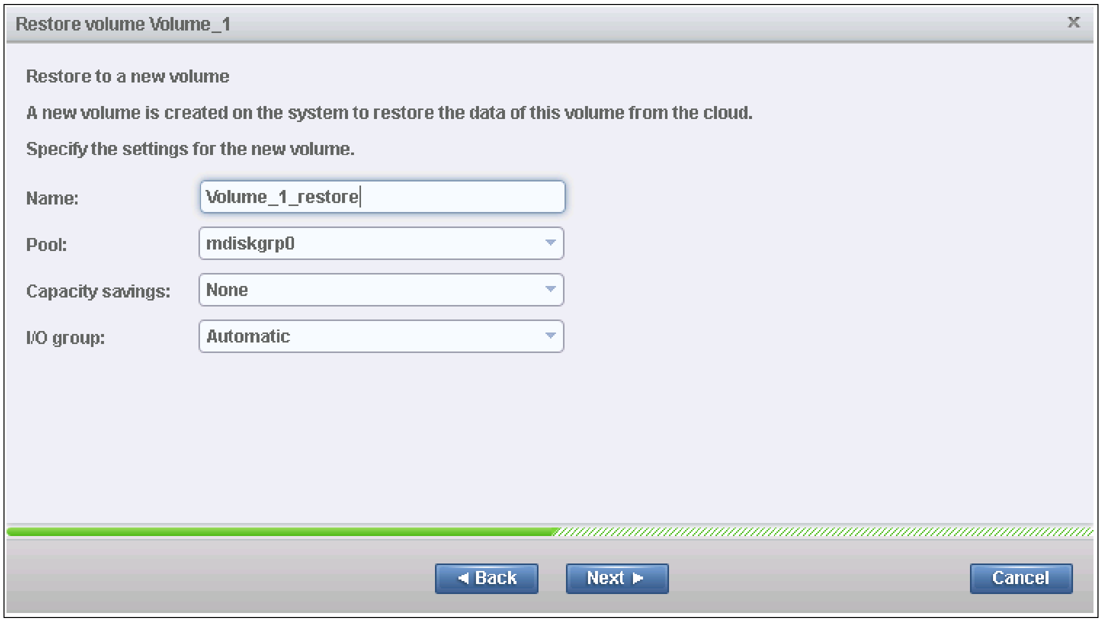
Select the destination target for restoring the volume Volume_1.

○ Restore directly to the production volume

● Restore to a new volume

Figure 11-79 Options to restore volume from cloud

Select **Restore to a new volume** and click **Next.** The system will guide you through the process to restore the volume and will provide input options such as Name, Storage Pool, option to Capacity Savings characteristic (Compressed or Thin-provisioned) and I/O group as shown in Figure 11-80. Click **Next** after you complete the settings for the new volume.



Restore volume Volume_1

Restore to a new volume

A new volume is created on the system to restore the data of this volume from the cloud.

Specify the settings for the new volume.

Name: Volume_1_restore

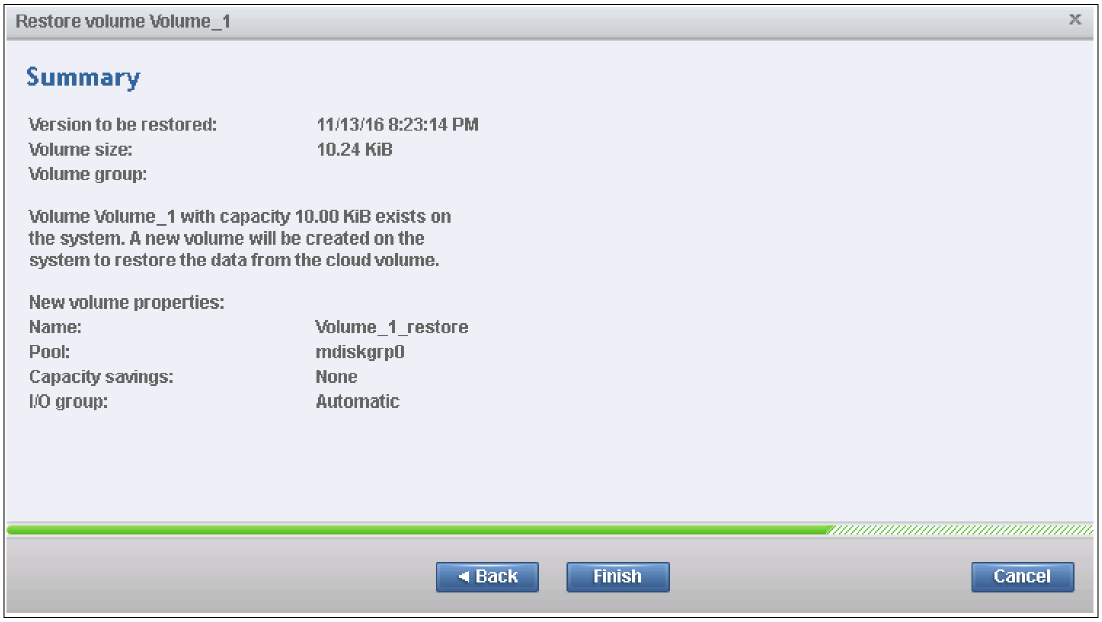
Pool: mdiskgrp0

Capacity savings: None

I/O group: Automatic

*Figure 11-80 Options to restore snapshots from cloud*

The next window shows the summary of the new volume that the IBM Spectrum Virtualize will create to restore the snapshot stored in the cloud. The restore acts as reverse and the cloud object storage acts as source and the newly volume as target.



Restore volume Volume_1

Summary

Version to be restored: 11/13/16 8:23:14 PM
Volume size: 10.24 KiB
Volume group:

Volume Volume_1 with capacity 10.00 KiB exists on the system. A new volume will be created on the system to restore the data from the cloud volume.

New volume properties:
Name: Volume_1_restore
Pool: mdiskgrp0
Capacity savings: None
I/O group: Automatic

*Figure 11-81 Summary of the restore option*

After you click **Finish** the system creates a new volume and the selected snapshot version is then restored to the new volume. Once the system completes the restore operation, go to **Volume** window and verify the newly volume used to restore the whole data set from cloud infrastructure.

## 11.7 Volume mirroring and migration options

*Volume mirroring* is a simple RAID 1-type function that enables a volume to remain online even when the storage pool backing it becomes inaccessible. Volume mirroring is designed to protect the volume from storage infrastructure failures by seamless mirroring between storage pools.

Volume mirroring is provided by a specific volume mirroring function in the I/O stack, and it cannot be manipulated like a FlashCopy or other types of copy volumes. However, this feature provides migration functionality, which can be obtained by splitting the mirrored copy from the source or by using the *migrate to* function. Volume mirroring cannot control backend storage mirroring or replication.

With volume mirroring, host I/O completes when both copies are written, and this feature is enhanced with a tunable latency tolerance. This tolerance provides an option to give preference to losing the redundancy between the two copies. This tunable timeout value is `Latency` or `Redundancy`.

The `Latency` tuning option, which is set with **`chvdisk -mirrorwritepriority latency`**, is the default. It prioritizes host I/O latency, which yields a preference to host I/O over availability. However, you might need to give preference to redundancy in your environment when availability is more important than I/O response time. Use the **`chvdisk -mirror writepriority redundancy`** command to set the redundancy option.

Regardless of which option you choose, volume mirroring can provide extra protection for your environment.

Migration offers the following options:

- Export to Image mode. By using this option, you can move storage from `managed` mode to `image` mode, which is useful if you are using the IBM Storwize as a migration device. For example, vendor A’s product cannot communicate with vendor B’s product, but you must migrate existing data from vendor A to vendor B. By using Export to image mode, you can migrate data by using Copy Services functions and then return control to the native array while maintaining access to the hosts.
- Import to Image mode. By using this option, you can import an existing storage MDisk or logical unit number (LUN) with its existing data from an external storage system without putting metadata on it so that the existing data remains intact. After you import it, all copy services functions can be used to migrate the storage to other locations while the data remains accessible to your hosts.
- Volume migration by using volume mirroring and then by using Split into New Volume. By using this option, you can use the available RAID 1 functionality. You create two copies of data that initially has a set relationship (one volume with two copies, one primary and one secondary) but then break the relationship (two volumes, both primary and no relationship between them) to make them independent copies of data.

  You can use this to migrate data between storage pools and devices. You might use this option if you want to move volumes to multiple storage pools. Each volume can have two copies at a time, which means you can add only one copy to the original volume, and then you have to split those copies to create another copy of the volume.
- Volume migration by using move to another pool. By using this option, you can move any volume between storage pools without any interruption to the host access. This option is a quicker version of the “Volume Mirroring and Split into New Volume” option. You might use this option if you want to move volumes in a single step, or you do not have a volume mirror copy already.

**Migration:** While these migration methods do not disrupt access, you must take a brief outage to install the host drivers for your IBM Storwize V7000 if you do not already have them installed.

With volume mirroring, you can move data to different MDisks within the same storage pool or move data between different storage pools. Using volume mirroring over volume migration is beneficial because with volume mirroring, storage pools do not need to have the same extent size as is the case with volume migration.

**Note:** Volume mirroring does not create a second volume before you split copies. Volume mirroring adds a second copy of the data under the same volume, so you end up having one volume presented to the host with two copies of data connected to this volume. Only splitting copies creates another volume, and then both volumes have only one copy of the data.

Starting with V7.3 and the introduction of the new cache architecture, mirrored volume performance has been significantly improved. Now, lower cache is beneath the volume mirroring layer, which means that both copies have their own cache.

This approach helps in cases of having copies of different types, for example generic and compressed, because now both copies use its independent cache and performs its own read prefetch. Destaging of the cache can now be done independently for each copy, so one copy does not affect performance of a second copy.

Also, because the Storwize destage algorithm is MDisk aware, it can tune or adapt the destaging process, depending on MDisk type and usage, for each copy independently.

# 11.8 Native IP replication

Before we describe Remote Copy features that benefit from the use of multiple IBM Storwize V7000 systems, it is important to describe the partnership option introduced with V7.2 native IP replication.

## 11.8.1 Native IP replication technology

Remote Mirroring over IP communication is supported on the IBM SAN Volume Controller and Storwize Family systems by using Ethernet communication links. The IBM Spectrum Virtualize Software IP replication uses innovative *Bridgeworks SANSlide* technology to optimize network bandwidth and utilization. This new function enables the use of a lower-speed and lower-cost networking infrastructure for data replication.

Bridgeworks’ SANSlide technology, which is integrated into the IBM Spectrum Virtualize Software, uses artificial intelligence to help optimize network bandwidth use and adapt to changing workload and network conditions. This technology can improve remote mirroring network bandwidth usage up to three times, which can enable clients to deploy a less costly network infrastructure, or speed up remote replication cycles to enhance disaster recovery effectiveness.

With an Ethernet network data flow the data transfer can slow down over time. This condition occurs because of the latency that is caused by waiting for the acknowledgment of each set of packets that are sent. The next packet set cannot be sent until the previous packet is acknowledged, as shown in Figure 11-82.

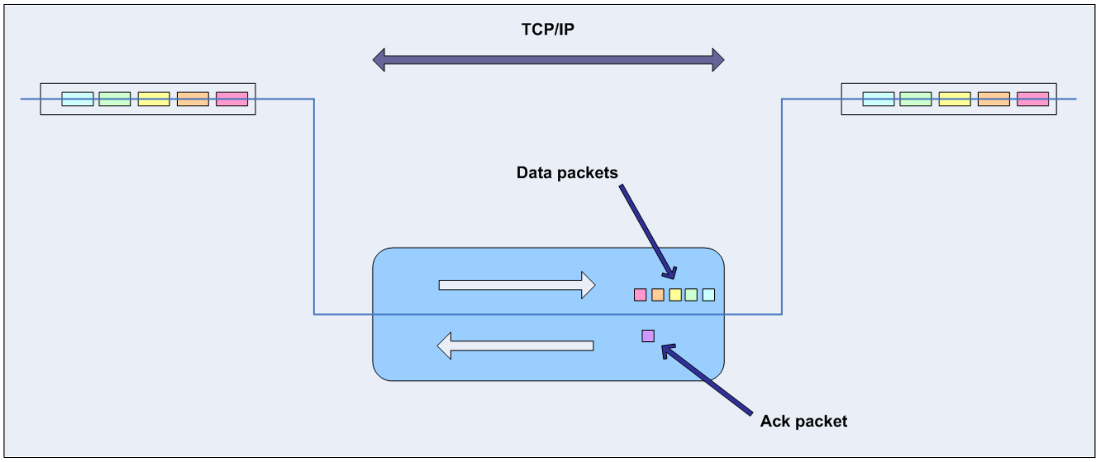


*Figure 11-82 Typical Ethernet network data flow*

However, by using the embedded IP replication, this behavior can be eliminated with the enhanced parallelism of the data flow by using multiple virtual connections (VC) that share IP links and addresses. The artificial intelligence engine can dynamically adjust the number of VCs, receive window size, and packet size as appropriate to maintain optimum performance. While the engine is waiting for one VC's ACK, it sends more packets across other VCs. If packets are lost from any VC, data is automatically retransmitted, as shown in Figure 11-83.

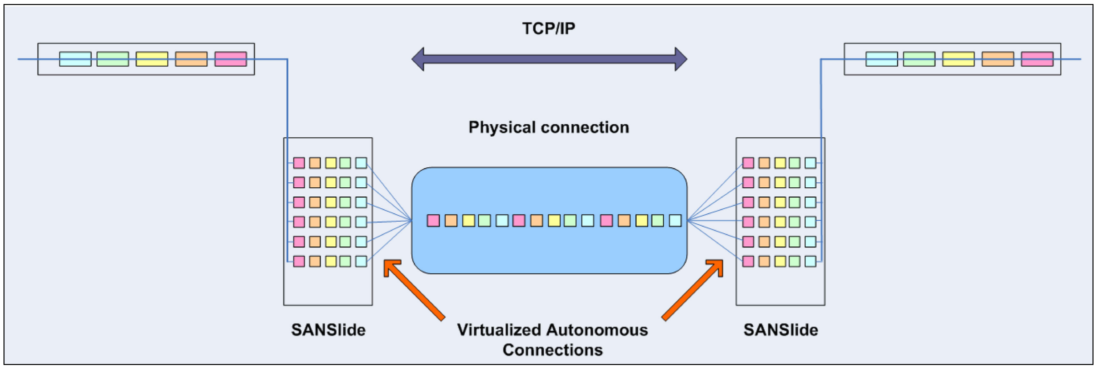


*Figure 11-83 Optimized network data flow by using Bridgeworks SANSlide technology*

For more information about this technology, see *IBM Storwize V7000 and SANSlide Implementation*, REDP-5023.

With native IP partnership, the following Copy Services features are supported:

- Metro Mirror (MM)

  Referred to as *synchronous replication*, MM provides a consistent copy of a source virtual disk on a target virtual disk. Data is written to the target virtual disk synchronously after it is written to the source virtual disk so that the copy is continuously updated.

- Global Mirror (GM) and GM with Change Volumes

  Referred to as *asynchronous replication*, GM provides a consistent copy of a source virtual disk on a target virtual disk. Data is written to the target virtual disk asynchronously so that the copy is continuously updated. However, the copy might not contain the last few updates if a disaster recovery (DR) operation is performed. An added extension to GM is GM with Change Volumes. GM with Change Volumes is the preferred method for use with native IP replication.

### 11.8.2 IBM Storwize System Layers

An IBM Storwize family system can be in one of the two layers: the *replication* layer or the *storage* layer. The system layer affects how the system interacts with IBM Storwize V7000 systems and IBM SAN Volume Controller systems. The IBM SAN Volume Controller is always set to replication layer and this parameter is unchangeable.

In storage layer, an Storwize family system has the following characteristics and requirements:

- The system can perform MM and GM replication with other storage-layer systems.
- The system can provide external storage for replication-layer systems or IBM SAN Volume Controller.
- The system cannot use a storage-layer system as external storage.

In replication layer, an IBM SAN Volume Controller or an IBM Storwize family system has the following characteristics and requirements:

- The system can perform MM and GM replication with other replication-layer systems or IBM SAN Volume Controller.
- The system cannot provide external storage for a replication-layer system or an IBM SAN Volume Controller.
- The system can use a storage-layer system as external storage.

A Storwize family system is in the storage layer by default, but the layer can be changed. For example, you might want to change a Storwize V7000 to a replication layer if you want to virtualize Storwize V3700 systems or other Storwize V7000 systems.

> **Note:** Before you change the system layer, the following conditions must be met:
>
> - No host object can be configured with worldwide port names (WWPNs) from a Storwize family system.
> - No system partnerships can be defined.
> - No Storwize family system can be visible on the SAN fabric.
>
> The layer can be changed during normal host I/O.

In your IBM Storwize system, use the **lssystem** command to check the current system layer, as shown in Example 11-2.

*Example 11-2 Output from lssystem command showing the system layer*

```
IBM_Storwize:ITSO_7K:superuser>lssystem
id 000001002140020E
name ITSO_V7K
...
```

```
lines omited for brevity
...
easy_tier_acceleration off
has_nas_key no
layer replication
...
```

> **Note:** Consider the following rules for creating remote partnerships between the IBM SAN Volume Controller and Storwize Family systems:
>
> - An IBM SAN Volume Controller is always in the replication layer.
> - By default, the IBM Storwize systems are in the storage layer but can be changed to the replication layer.
> - A system can form partnerships only with systems in the same layer.
> - Starting in software V6.4, an IBM SAN Volume Controller or Storwize system in the replication layer can virtualize an IBM Storwize in the storage layer.

### 11.8.3 IP partnership limitations

The following prerequisites and assumptions must be considered before IP partnership between two IBM Spectrum Virtualize systems can be established:

- The IBM SAN Volume Controller or IBM Storwize systems are successfully installed with V7.2 or later code levels.
- The systems must have the necessary licenses that enable remote copy partnerships to be configured between two systems. No separate license is required to enable IP partnership.
- The storage SANs are configured correctly and the correct infrastructure to support the Spectrum Virtualize systems in remote copy partnerships over IP links is in place.
- The two systems must be able to ping each other and perform the discovery.
- The maximum number of partnerships between the local and remote systems, including both IP and Fibre Channel (FC) partnerships, is limited to the current maximum that is supported, which is three partnerships (four systems total).
- Only a single partnership over IP is supported.
- A system can have simultaneous partnerships over FC and IP, but with separate systems. The FC zones between two systems must be removed before an IP partnership is configured.
- IP partnerships are supported on both 10 gigabits per second (Gbps) links and 1 Gbps links. However, the intermix of both on a single link is not supported.
- The maximum supported round-trip time is 80 milliseconds (ms) for 1 Gbps links.
- The maximum supported round-trip time is 10 ms for 10 Gbps links.
- The minimum supported link bandwidth is 10 Mbps.
- The inter-cluster heartbeat traffic uses 1 Mbps per link.
- Only nodes from two I/O Groups can have ports that are configured for an IP partnership.
- Migrations of remote copy relationships directly from FC-based partnerships to IP partnerships are not supported.
- IP partnerships between the two systems can be over IPv4 or IPv6 only, but not both.

- Virtual LAN (VLAN) tagging of the IP addresses that are configured for remote copy is supported starting with V7.4.
- Management IP and Internet SCSI (iSCSI) IP on the same port can be in a different network starting with V7.4.
- An added layer of security is provided by using Challenge Handshake Authentication Protocol (CHAP) authentication.
- Transmission Control Protocol (TCP) ports 3260 and 3265 are used for IP partnership communications. Therefore, these ports must be open in firewalls between the systems.
- Only a single Remote Copy data session per physical link can be established. It is intended that only one connection (for sending/receiving Remote Copy data) is made for each independent physical link between the systems.

> **Note:** A physical link is the physical IP link between the two sites, A (local) and B (remote). Multiple IP addresses on local system A could be connected (by Ethernet switches) to this physical link. Similarly, multiple IP address on remote system B could be connected (by Ethernet switches) to the same physical link. At any point in time, only a single IP address on cluster A can form an RC data session with an IP address on cluster B.

- The maximum throughput is restricted based on the use of 1 Gbps or 10 Gbps Ethernet ports, and varies based on distance (for example, round-trip latency) and quality of communication link (for example, packet loss):
  - One 1 Gbps port might transfer up to 110 megabytes per second (MBps) unidirectional, 190 MBps bidirectional
  - Two 1 Gbps ports might transfer up to 220 MBps unidirectional, 325 MBps bidirectional
  - One 10 Gbps port might transfer up to 240 MBps unidirectional, 350 MBps bidirectional
  - Two 10 Gbps port might transfer up to 440 MBps unidirectional, 600 MBps bidirectional

> **Note:** The Bandwidth setting definition when the IP partnerships are created changed. Previously, the bandwidth setting defaulted to 50 MB, and was the maximum transfer rate from the primary site to the secondary site for initial sync/resyncs of volumes.
>
> The Link Bandwidth setting is now configured by using megabits (Mb) not MB. You set the Link Bandwidth setting to a value that the communication link can sustain, or to what is allocated for replication. The Background Copy Rate setting is now a percentage of the Link Bandwidth. The Background Copy Rate setting determines the available bandwidth for the initial sync and resyncs or for GM with Change Volumes.

## 11.8.4 VLAN support

Starting with V7.4, VLAN tagging is supported for both iSCSI host attachment and IP replication. Hosts and remote-copy operations can connect to the system through Ethernet ports. Each traffic type has different bandwidth requirements, which can interfere with each other if they share the same IP connections. VLAN tagging creates two separate connections on the same IP network for different types of traffic. The system supports VLAN configuration on both IPv4 and IPv6 connections.

When the VLAN ID is configured for the IP addresses that are used for either iSCSI host attach or IP replication, the appropriate VLAN settings on the Ethernet network and servers must be configured correctly in order not to experience connectivity issues. After the VLANs are configured, changes to the VLAN settings will disrupt iSCSI and IP replication traffic to and from the partnerships.

During the VLAN configuration for each IP address, the VLAN settings for the local and failover ports on two nodes of an I/O Group can differ. To avoid any service disruption, switches must be configured so the failover VLANs are configured on the local switch ports and the failover of IP addresses from a failing node to a surviving node succeeds. If failover VLANs are not configured on the local switch ports, there are no paths to IBM Storwize V7000 nodes during a node failure and the replication will fail.

Consider the following requirements and procedures when implementing VLAN tagging:

- VLAN tagging is supported for IP partnership traffic between two systems.
- VLAN provides network traffic separation at the layer 2 level for Ethernet transport.
- VLAN tagging by default is disabled for any IP address of a node port. You can use the CLI or GUI to optionally set the VLAN ID for port IPs on both systems in the IP partnership.
- When a VLAN ID is configured for the port IP addresses that are used in remote copy port groups, appropriate VLAN settings on the Ethernet network must also be properly configured to prevent connectivity issues.

Setting VLAN tags for a port is disruptive. Therefore, VLAN tagging requires that you stop the partnership first before you configure VLAN tags. Then, restart again when the configuration is complete.

## 11.8.5 IP partnership and terminology

The IP partnership terminology and abbreviations that are used are listed in Table 11-7.

*Table 11-7 Terminology for IP partnership*

| IP partnership terminology | Description |
|---|---|
| Remote copy group or Remote copy port group | The following numbers group a set of IP addresses that are connected to the same physical link. Therefore, only IP addresses that are part of the same remote copy group can form remote copy connections with the partner system:<br>► 0 – Ports that are not configured for remote copy<br>► 1 – Ports that belong to remote copy port group 1<br>► 2 – Ports that belong to remote copy port group 2<br>Each IP address can be shared for iSCSI host attach and remote copy functionality. Therefore, appropriate settings must be applied to each IP address. |
| IP partnership | Two systems that are partnered to perform remote copy over native IP links. |
| FC partnership | Two systems that are partnered to perform remote copy over native Fibre Channel links. |
| Failover | Failure of a node within an I/O group causes the volume access to go through the surviving node. The IP addresses fail over to the surviving node in the I/O group. When the configuration node of the system fails, management IPs also fail over to an alternative node. |

| IP partnership terminology | Description |
|---|---|
| Failback | When the failed node rejoins the system, all failed over IP addresses are failed back from the surviving node to the rejoined node, and virtual disk access is restored through this node. |
| linkbandwidthmbits | Aggregate bandwidth of all physical links between two sites in Mbps. |
| IP partnership or partnership over native IP links | These terms are used to describe the IP partnership feature. |
| Discovery | Process by which two IBM Spectrum Virtualize systems exchange information about their IP address configuration. For IP-based partnerships, only IP addresses configured for Remote Copy are discovered.<br>For example, the first Discovery takes place when the user is running the `mkippartnership` CLI command. Subsequent Discoveries can take place as a result of user activities (configuration changes) or as a result of hardware failures (for example, node failure, ports failure, and so on). |

## 11.8.6 States of IP partnership

The different partnership states in IP partnership are listed in Table 11-8.

*Table 11-8 States of IP partnership*

| State | Systems connected | Support for active remote copy I/O | Comments |
|---|---|---|---|
| Partially_Configured_Local | No | No | This state indicates that the initial discovery is complete. |
| Fully_Configured | Yes | Yes | Discovery successfully completed between two systems, and the two systems can establish remote copy relationships. |
| Fully_Configured_Stopped | Yes | Yes | The partnership is stopped on the system. |
| Fully_Configured_Remote_Stopped | Yes | No | The partnership is stopped on the remote system. |
| Not_Present | Yes | No | The two systems cannot communicate with each other. This state is also seen when data paths between the two systems are not established. |
| Fully_Configured_Exceeded | Yes | No | There are too many systems in the network, and the partnership from the local system to remote system is disabled. |
| Fully_Configured_Excluded | No | No | The connection is excluded because of too many problems, or either system cannot support the I/O work load for the Metro Mirror and Global Mirror relationships. |

The following steps must be completed to establish two systems in the IP partnerships:

1. The administrator configures the CHAP secret on both the systems. This step is not mandatory, and users can choose to not configure the CHAP secret.
2. The administrator configures the system IP addresses on both local and remote systems so that they can discover each other over the network.
3. If you want to use VLANs, configure your LAN switches and Ethernet ports to use VLAN tagging (for more information about VLAN tagging, see VLAN support).
4. The administrator configures the systems ports on each node in both of the systems by using the GUI (or the **`cfgportip`** CLI command), and completes the following steps:
   a. Configure the IP addresses for remote copy data.
   b. Add the IP addresses in the respective remote copy port group.
   c. Define whether the host access on these ports over iSCSI is allowed.
5. The administrator establishes the partnership with the remote system from the local system where the partnership state then changes to the `Partially_Configured_Local` state.
6. The administrator establishes the partnership from the remote system with the local system, and if successful, the partnership state then changes to the `Fully_Configured` state, which implies that the partnerships over the IP network were successfully established. The partnership state momentarily remains in the `Not_Present` state before moving to the `Fully_Configured` state.
7. The administrator creates MM, GM, and GM with Change Volume relationships.

**Partnership consideration**: When the partnership is created, no master or auxiliary status is defined or implied. The partnership is equal. The concepts of *master or auxiliary* and *primary or secondary* apply to volume relationships only, not to system partnerships.

## 11.8.7 Remote copy groups

This section describes remote copy groups (or remote copy port groups) and different ways to configure the links between the two remote systems. The two IBM Spectrum Virtualize systems can be connected to each other over one link or, at most, two links. To address the requirement to enable the systems to know about the physical links between the two sites, the concept of remote copy port groups was introduced.

Remote copy port group ID is a numerical tag associated with an IP port of IBM Storwize V7000 to indicate which physical IP link it is connected to. Multiple nodes could be connected to the same physical long-distance link, and must therefore share the same remote copy port group id.

In scenarios where there are two physical links between the local and remote clusters, two remote copy port group IDs must be used to designate which IP addresses are connected to which physical link. This configuration must be done by the system administrator using the GUI or the **`cfgportip`** CLI command.

**Remember:** IP ports on both partners must have been configured with identical remote copy port group IDs for the partnership to be established correctly.

The IBM Storwize V7000 IP addresses that are connected to the same physical link are designated with identical remote copy port groups. The IBM Storwize V7000 system supports three remote copy groups: 0, 1, and 2.

The IBM Storwize V7000 IP addresses are, by default, in remote copy port group 0. Ports in port group 0 are not considered for creating remote copy data paths between two systems. For partnerships to be established over IP links directly, IP ports must be configured in remote copy group 1 if a single inter-site link exists, or in remote copy groups 1 and 2 if two inter-site links exist.

You can assign one IPv4 address and one IPv6 address to each Ethernet port on the system platforms. Each of these IP addresses can be shared between iSCSI host attach and the IP partnership. The user must configure the required IP address (IPv4 or IPv6) on an Ethernet port with a remote copy port group.

The administrator might want to use IPv6 addresses for remote copy operations and use IPv4 addresses on that same port for iSCSI host attach. This configuration also implies that for two systems to establish an IP partnership, both systems must have IPv6 addresses that are configured.

Administrators can choose to dedicate an Ethernet port for IP partnership only. In that case, host access must be explicitly disabled for that IP address and any other IP address that is configured on that Ethernet port.

**Note:** To establish an IP partnership, each IBM Storwize node must have only a single remote copy port group that is configured, 1 or 2. The remaining IP addresses must be in remote copy port group 0.

## 11.8.8 Supported configurations

The following supported configurations for IP partnership that were in the first release are described in this section.

- Two 2-node systems in IP partnership over a single inter-site link, as shown in Figure 11-84 (configuration 1).

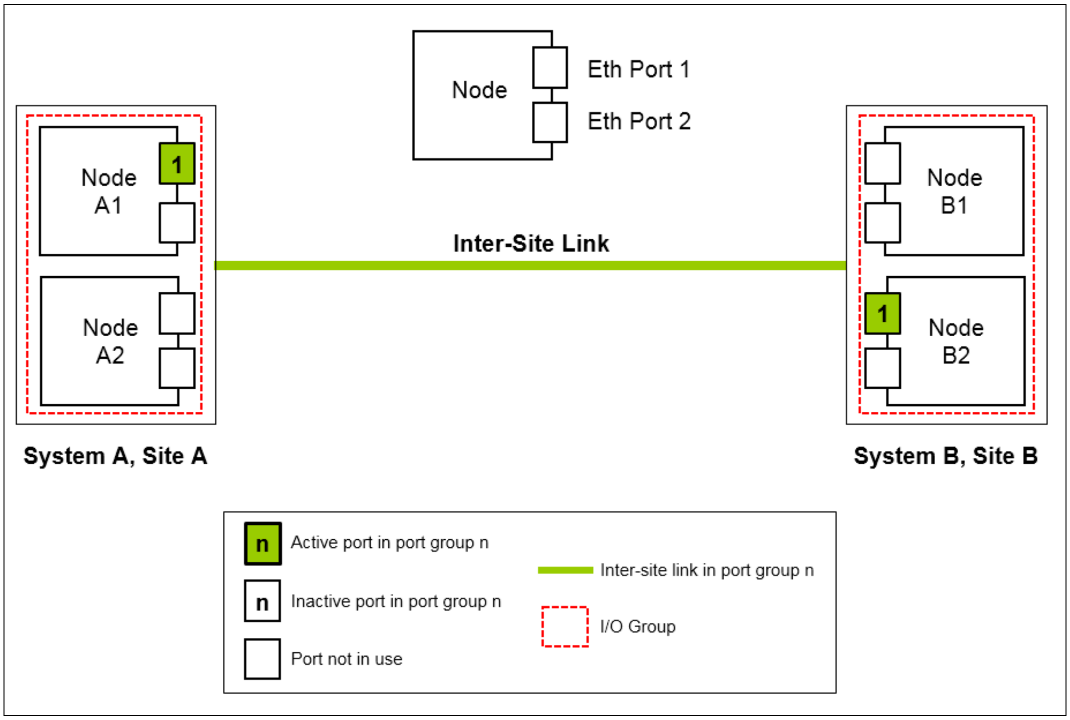


*Figure 11-84 Single link with only one remote copy port group that is configured in each system*

As shown in Figure 11-84, there are two systems: System A and System B. A single remote copy port group 1 is created on Node A1 on System A and on Node B2 on System B (an administrator might choose to configure the remote copy port group on Node B1 on System B rather than Node B2), because there is only a single inter-site link to facilitate the IP partnership traffic.

At any time, only the IP addresses that are configured in remote copy port group 1 on the nodes in System A and System B participate in establishing data paths between the two systems after the IP partnerships are created. In this configuration, there are no failover ports that are configured on the partner node in the same I/O group.

This configuration has the following characteristics:

- Only one node in each system has a remote copy port group that is configured, and there are no failover ports configured.
- If the Node A1 in System A or the Node B2 in System B were to encounter some failure, the IP partnership stops and enters the `Not_Present` state until the failed nodes recover.
- After the nodes recover, the IP ports fail back, the IP partnership recovers, and the partnership state goes to the `Fully_Configured` state.
- If the inter-site system link fails, the IP partnerships transition to the `Not_Present` state.
- This configuration is not recommended, because it is not resilient to node failures.

- Two 2-node systems in IP partnership over a single inter-site link (with failover ports configured), as shown in Figure 11-85 on page 490 (configuration 2).

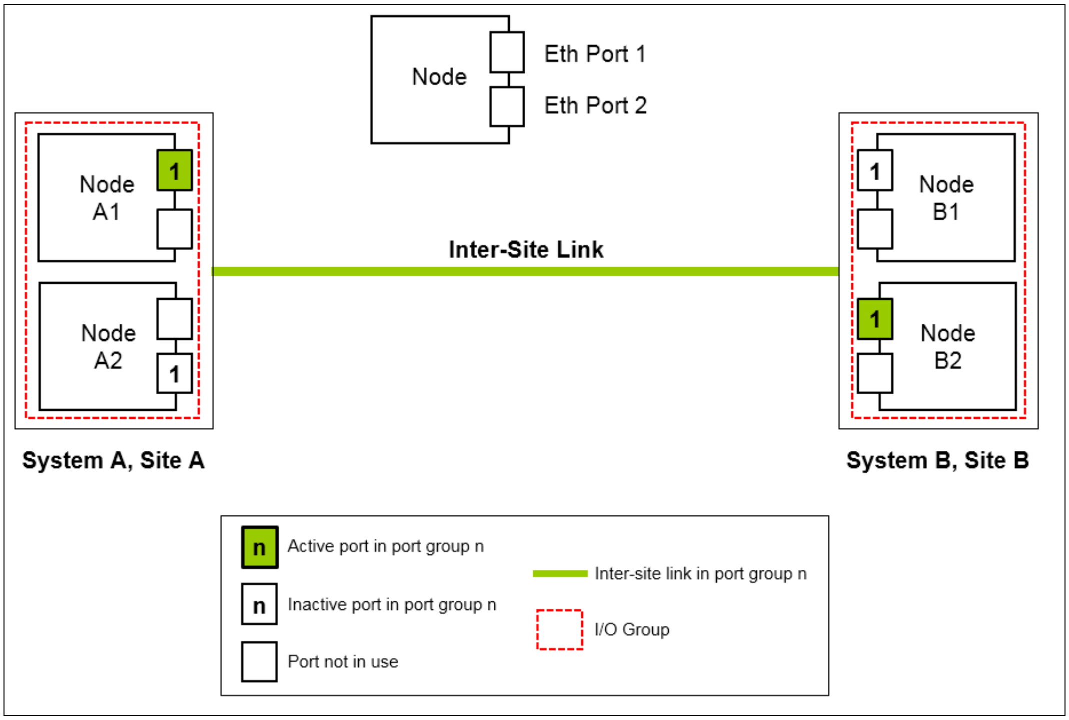


*Figure 11-85 One remote copy group on each system and nodes with failover ports configured*

As shown in Figure 11-85, there are two systems: System A and System B. A single remote copy port group 1 is configured on two Ethernet ports, one each on Node A1 and Node A2 on System A. Similarly, a single remote copy port group is configured on two Ethernet ports on Node B1 and Node B2 on System B.

Although there are two ports on each system that are configured for remote copy port group 1, only one Ethernet port in each system actively participates in the IP partnership process. This selection is determined by a path configuration algorithm that is designed to choose data paths between the two systems to optimize performance.

The other port on the partner node in the I/O Group behaves as a standby port that is used if there is a node failure. If Node A1 fails in System A, IP partnership continues servicing replication I/O from Ethernet Port 2, because a failover port is configured on Node A2 on Ethernet Port 2.

However, it might take some time for discovery and path configuration logic to reestablish paths post failover, and this can cause partnerships to change to `Not_Present` for that time. The details of the particular IP port that is actively participating in IP partnership is provided in the **`lsportip`** output (reported as `used`).

This configuration has the following characteristics:

– Each node in the I/O group has the same remote copy port group that is configured, but only one port in that remote copy port group is active at any time at each system.

– If the Node A1 in System A or the Node B2 in System B fails in the respective systems, IP partnerships rediscovery is triggered and continues servicing the I/O from the failover port.

– The discovery mechanism that is triggered because of failover might introduce a delay where the partnerships momentarily transition to the `Not_Present` state and recover.

- Two 4-node systems in IP partnership over a single inter-site link (with failover ports configured), as shown in Figure 11-86 on page 491 (configuration 3).

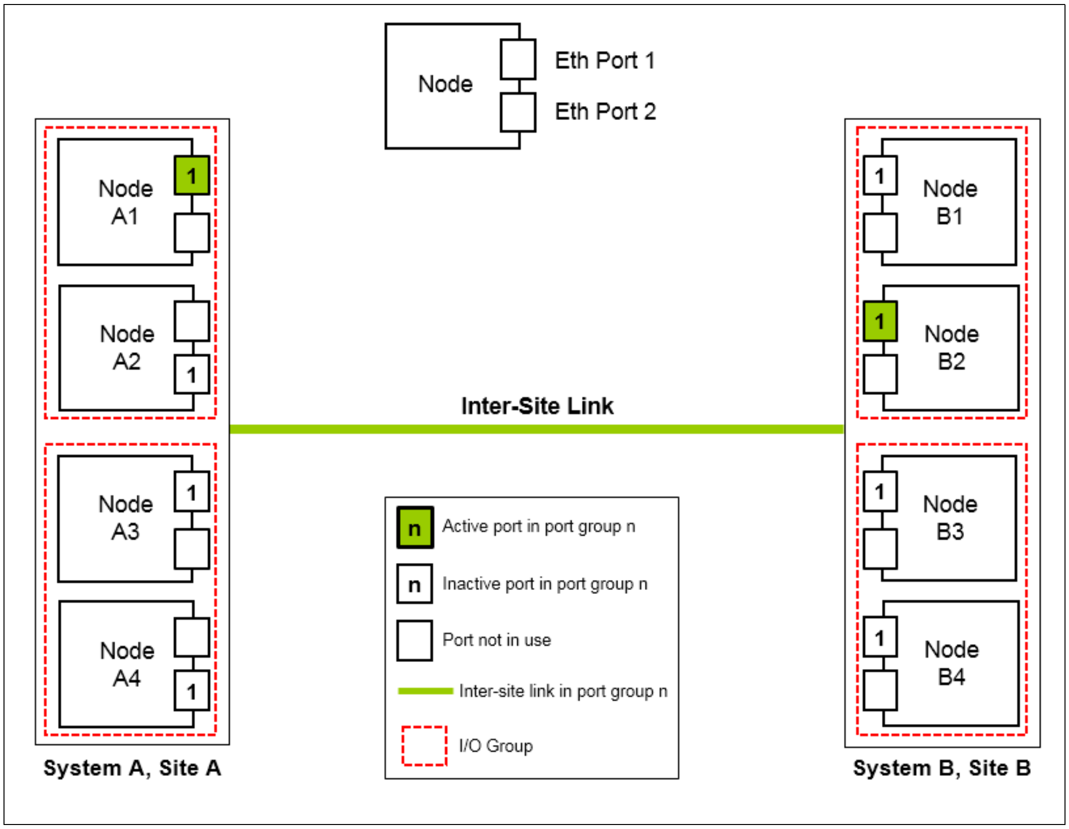


*Figure 11-86 Multinode systems single inter-site link with only one remote copy port group*

As shown in Figure 11-86, there are two 4-node systems: System A and System B. A single remote copy port group 1 is configured on nodes A1, A2, A3, and A4 on System A, Site A; and on nodes B1, B2, B3, and B4 on System B, Site B. Although there are four ports that are configured for remote copy group 1, only one Ethernet port in each remote copy port group on each system actively participates in the IP partnership process.

Port selection is determined by a path configuration algorithm. The other ports play the role of standby ports.

If Node A1 fails in System A, the IP partnership selects one of the remaining ports that is configured with remote copy port group 1 from any of the nodes from either of the two I/O groups in System A. However, it might take some time (generally seconds) for discovery and path configuration logic to reestablish paths post failover, and this process can cause partnerships to transition to the `Not_Present` state.

This result causes remote copy relationships to stop, and the administrator might need to manually verify the issues in the event log and start the relationships or remote copy consistency groups, if they do not autorecover. The details of the particular IP port actively participating in the IP partnership process is provided in the **`lsportip`** view (reported as `used`).

This configuration has the following characteristics:

- Each node has the remote copy port group that is configured in both I/O groups. However, only one port in that remote copy port group remains active and participates in IP partnership on each system.
- If the Node A1 in System A or the Node B2 in System B were to encounter some failure in the system, IP partnerships discovery is triggered and it continues servicing the I/O from the failover port.

  - The discovery mechanism that is triggered because of failover might introduce a delay wherein the partnerships momentarily transition to the `Not_Present` state and then recover.
  - The bandwidth of the single link is used completely.
- Eight-node system in IP partnership with four-node system over single inter-site link, as shown in Figure 11-87 (configuration 4).

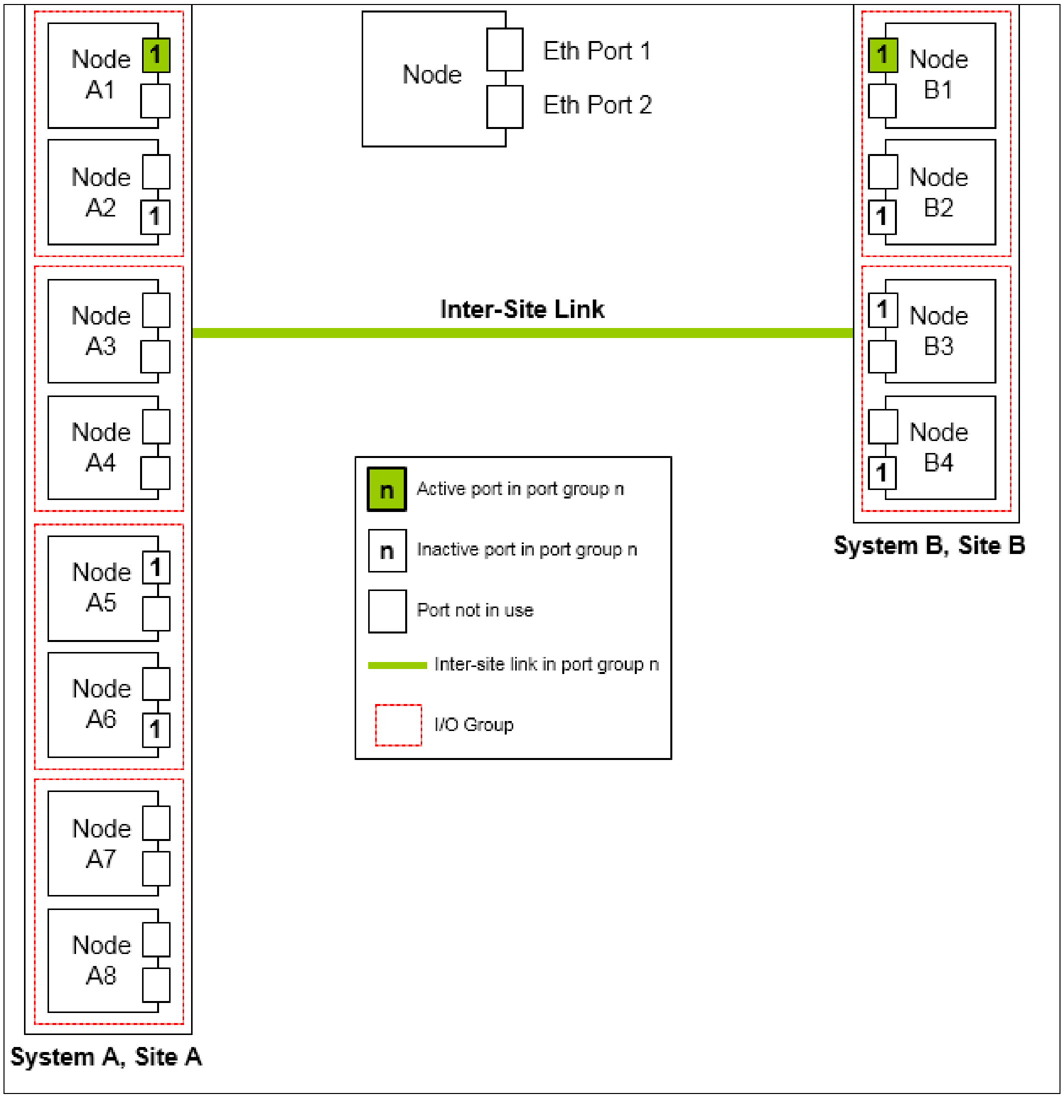


*Figure 11-87 Multinode systems single inter-site link with only one remote copy port group*

As shown in Figure 11-87, there is an eight-node system (System A in Site A) and a four-node system (System B in Site B). A single remote copy port group 1 is configured on nodes A1, A2, A5, and A6 on System A at Site A and similarly, a single remote copy port group 1 is configured on nodes B1, B2, B3, and B4 on System B.

Although there are four I/O groups (eight nodes) in System A, any two I/O groups at maximum are supported to be configured for IP partnerships. If Node A1 fails in System A, IP partnership continues using one of the ports that is configured in remote copy port group from any of the nodes from either of the two I/O groups in System A.

However, it might take some time for discovery and path configuration logic to reestablish paths post-failover, and this delay might cause partnerships to transition to the `Not_Present` state.

This process can lead to remote copy relationships stopping, and the administrator must manually start them if the relationships do not auto-recover. The details of which particular IP port is actively participating in IP partnership process is provided in **`lsportip`** output (reported as `used`).

This configuration has the following characteristics:

- Each node has the remote copy port group that is configured in both the I/O groups that are identified for participating in IP Replication. However, only one port in that remote copy port group remains active on each system and participates in IP Replication.
- If the Node A1 in System A or the Node B2 in System B fails in the system, the IP partnerships trigger discovery and continue servicing the I/O from the failover ports.
- The discovery mechanism that is triggered because of failover might introduce a delay wherein the partnerships momentarily transition to the `Not_Present` state and then recover.
- The bandwidth of the single link is used completely.

- Two 2-node systems with two inter-site links, as shown in Figure 11-88 (configuration 5).

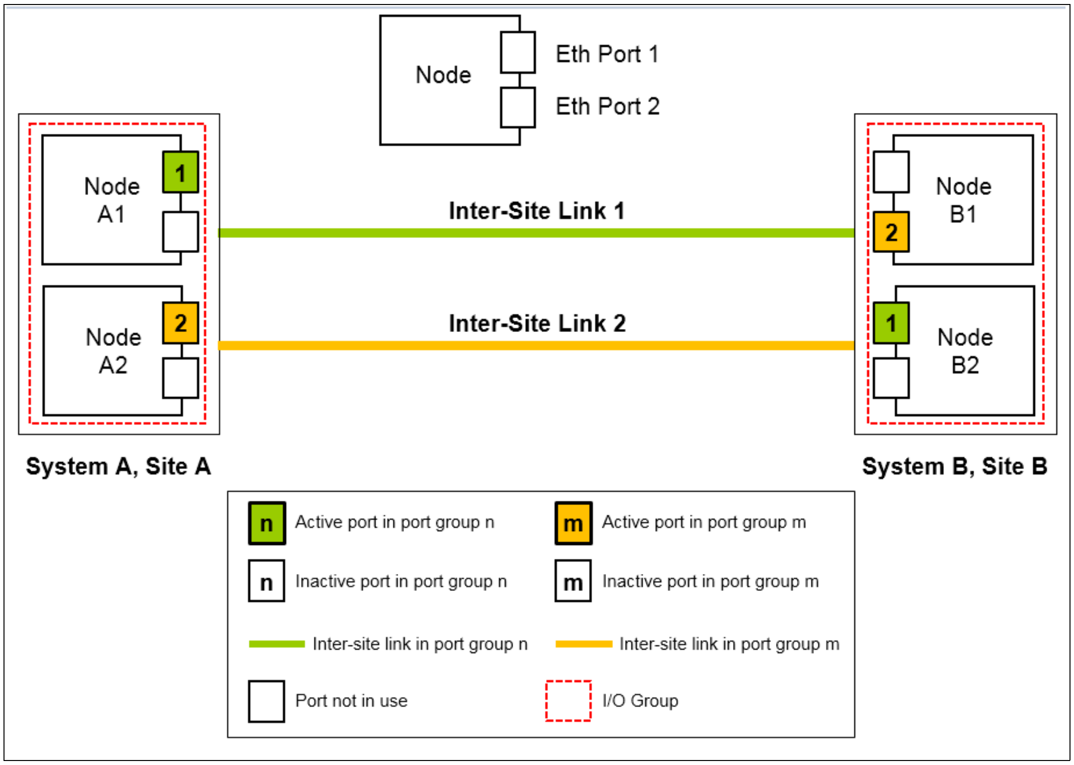

Node
Eth Port 1
Eth Port 2
Node A1
1
Node A2
2
Inter-Site Link 1
Inter-Site Link 2
Node B1
2
Node B2
1
System A, Site A
System B, Site B
n
Active port in port group n
m
Active port in port group m
n
Inactive port in port group n
m
Inactive port in port group m
Inter-site link in port group n
Inter-site link in port group m
Port not in use
I/O Group

*Figure 11-88 Dual links with two remote copy groups on each system configured*

As shown in Figure 11-88 on page 493, remote copy port groups 1 and 2 are configured on the nodes in System A and System B because there are two inter-site links available. In this configuration, the failover ports are not configured on partner nodes in the I/O group. Rather, the ports are maintained in different remote copy port groups on both of the nodes and they remain active and participate in IP partnership by using both of the links.

However, if either of the nodes in the I/O group fail (that is, if Node A1 on System A fails), the IP partnership continues only from the available IP port that is configured in remote copy port group 2. Therefore, the effective bandwidth of the two links is reduced to 50%, because only the bandwidth of a single link is available until the failure is resolved.

This configuration has the following characteristics:

– There are two inter-site links and two remote copy port groups are configured.

– Each node has only one IP port in remote copy port group 1 or 2.

– Both the IP ports in the two remote copy port groups participate simultaneously in IP partnerships. Therefore, both of the links are used.

– During node failure or link failure, the IP partnership traffic continues from the other available link and the port group. Therefore, if two links of 10 Mbps each are available and you have 20 Mbps of effective link bandwidth, bandwidth is reduced to 10 Mbps only during a failure.

– After the node failure or link failure is resolved and failback happens, the entire bandwidth of both of the links is available as before.

- Two 4-node systems in IP partnership with dual inter-site links, as shown in Figure 11-89 (configuration 6).

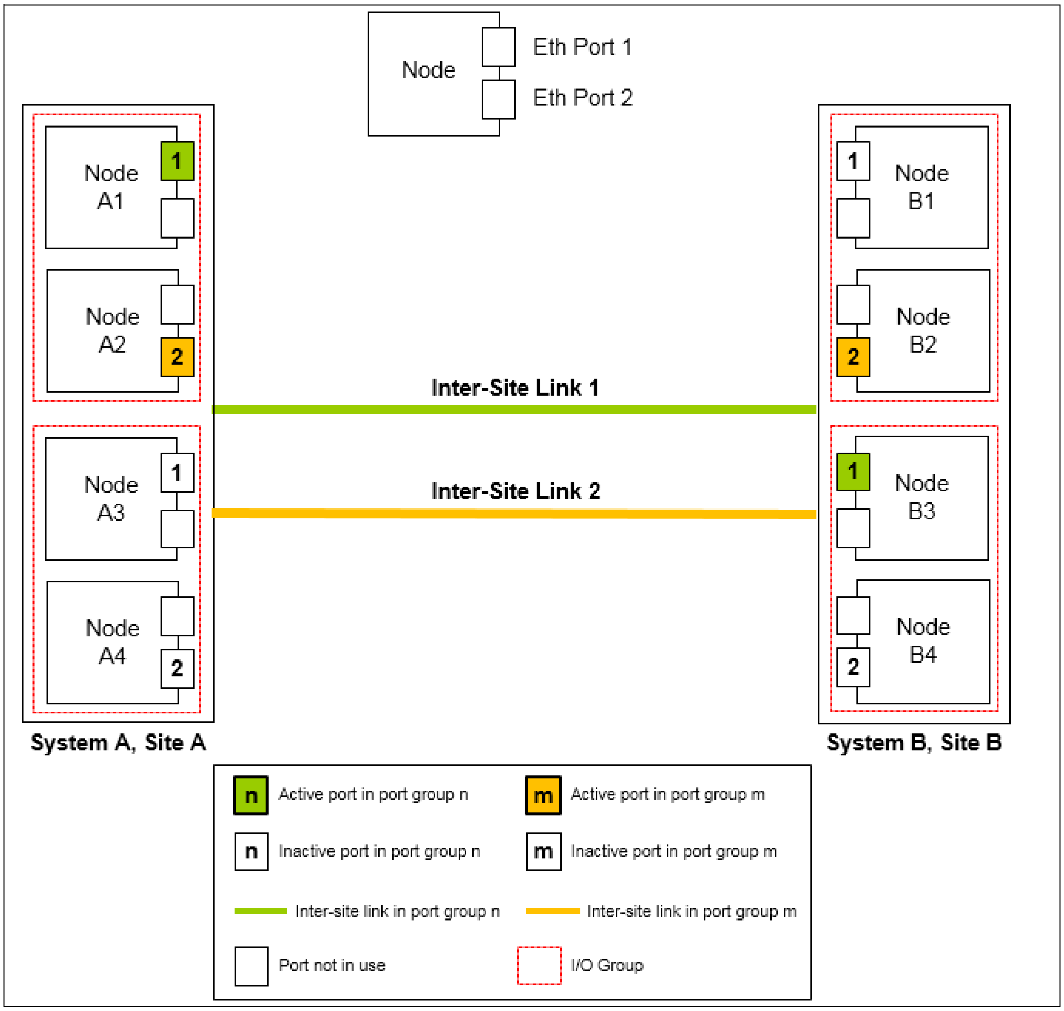

Node
Eth Port 1
Eth Port 2
Node A1
1
Node A2
2
Node A3
1
Node A4
2
Node B1
1
Node B2
2
Node B3
1
Node B4
2
Inter-Site Link 1
Inter-Site Link 2
System A, Site A
System B, Site B
n Active port in port group n
m Active port in port group m
n Inactive port in port group n
m Inactive port in port group m
Inter-site link in port group n
Inter-site link in port group m
Port not in use
I/O Group

*Figure 11-89 Multinode systems with dual inter-site links between the two systems*

As shown in Figure 11-89, there are two 4-node systems: System A and System B. This configuration is an extension of Configuration 5 to a multinode multi-I/O group environment. As seen in this configuration, there are two I/O groups and each node in the I/O group has a single port that is configured in remote copy port groups 1 or 2.

Although there are two ports that are configured in remote copy port groups 1 and 2 on each system, only one IP port in each remote copy port group on each system actively participates in IP partnership. The other ports that are configured in the same remote copy port group act as standby ports in event of failure. Which port in a configured remote copy port group participates in IP partnership at any moment is determined by a path configuration algorithm.

In this configuration, if Node A1 fails in System A, IP partnership traffic continues from Node A2 (that is, remote copy port group 2) and at the same time the failover also causes discovery in remote copy port group 1. Therefore, the IP partnership traffic continues from Node A3 on which remote copy port group 1 is configured. The details of the particular IP port that is actively participating in IP partnership process is provided in the **`lsportip`** output (reported as `used`).

This configuration has the following characteristics:

- Each node has the remote copy port group that is configured in the I/O groups 1 or 2. However, only one port per system in both remote copy port groups remains active and participates in IP partnership.
- Only a single port per system from each configured remote copy port group participates simultaneously in IP partnership. Therefore, both of the links are used.
- During node failure or port failure of a node that is actively participating in IP partnership, IP partnership continues from the alternative port because another port is in the system in the same remote copy port group but in a different I/O Group.
- The pathing algorithm can start discovery of available port in the affected remote copy port group in the second I/O group and pathing is reestablished, which restores the total bandwidth, so both of the links are available to support IP partnership.

- Eight-node system in IP partnership with a four-node system over dual inter-site links, as shown in Figure 11-90 (configuration 7).

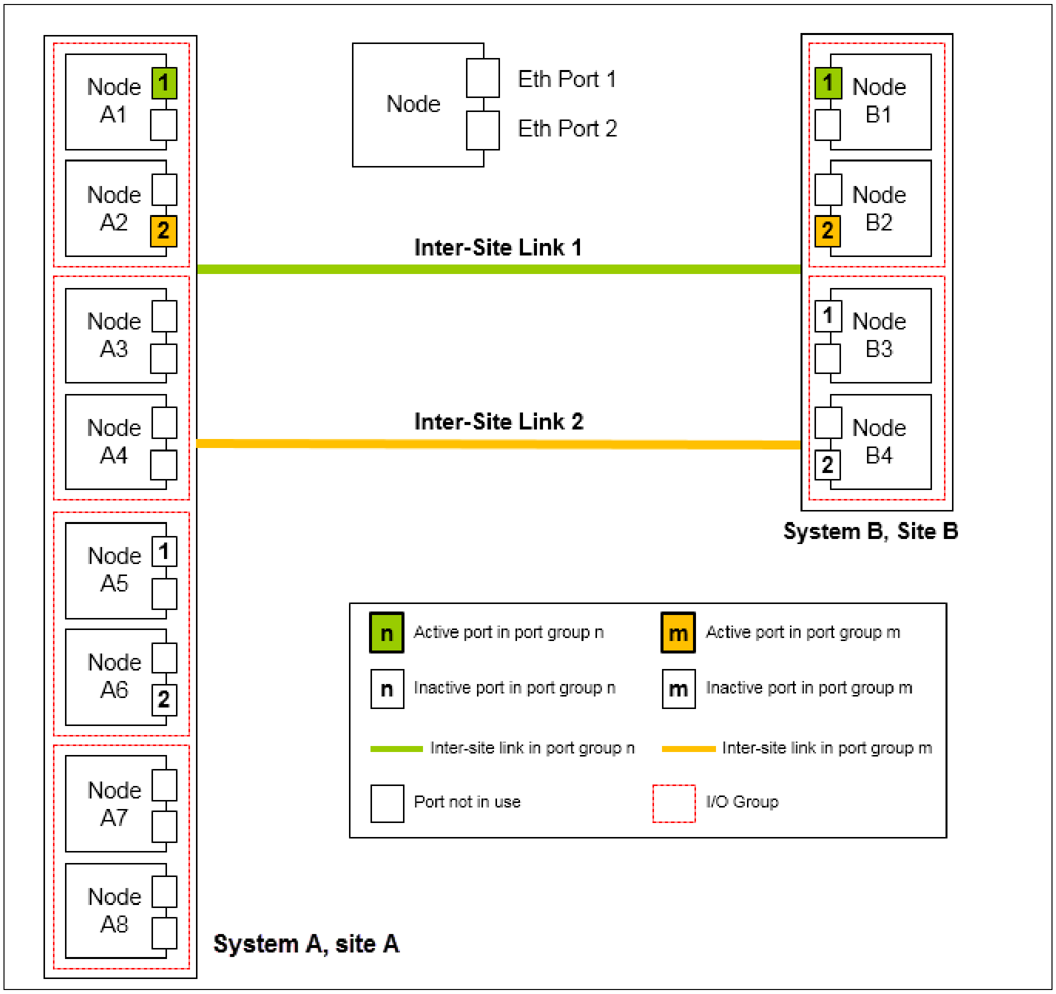


*Figure 11-90 Multinode systems (two I/O groups on each system) with dual inter-site links between the two systems*

As shown Figure 11-90, there is an eight-node System A in Site A and a four-node System B in Site B. Because a maximum of two I/O groups in IP partnership is supported in a system, although there are four I/O groups (eight nodes), nodes from only two I/O groups' are configured with remote copy port groups in System A. The remaining or all of the I/O groups can be configured to be remote copy partnerships over FC.

In this configuration, there are two links and two I/O groups that are configured with remote copy port groups 1 and 2, but path selection logic is managed by an internal algorithm. Therefore, this configuration depends on the pathing algorithm to decide which of the nodes actively participates in IP partnership. Even if Node A5 and Node A6 are configured with remote copy port groups properly, active IP partnership traffic on both of the links might be driven from Node A1 and Node A2 only.

If Node A1 fails in System A, IP partnership traffic continues from Node A2 (that is, remote copy port group 2) and the failover also causes IP partnership traffic to continue from Node A5 on which remote copy port group 1 is configured. The details of the particular IP port actively participating in IP partnership process is provided in the **`lsportip`** output (reported as `used`).

This configuration has the following characteristics:

– There are two I/O Groups with nodes in those I/O groups that are configured in two remote copy port groups, because there are two inter-site links for participating in IP partnership. However, only one port per system in a particular remote copy port group remains active and participates in IP partnership.

- One port per system from each remote copy port group participates in IP partnership simultaneously. Therefore, both of the links are used.
- If a node or port on the node that is actively participating in IP partnership fails, the remote copy (RC) data path is established from that port because another port is available on an alternative node in the system with the same remote copy port group.
- The path selection algorithm starts discovery of available ports in the affected remote copy port group in the alternative I/O groups and paths are reestablished, which restores the total bandwidth across both links.
- The remaining or all of the I/O groups can be in remote copy partnerships with other systems.

- An example of an unsupported configuration for a single inter-site link is shown in Figure 11-91 (configuration 8).

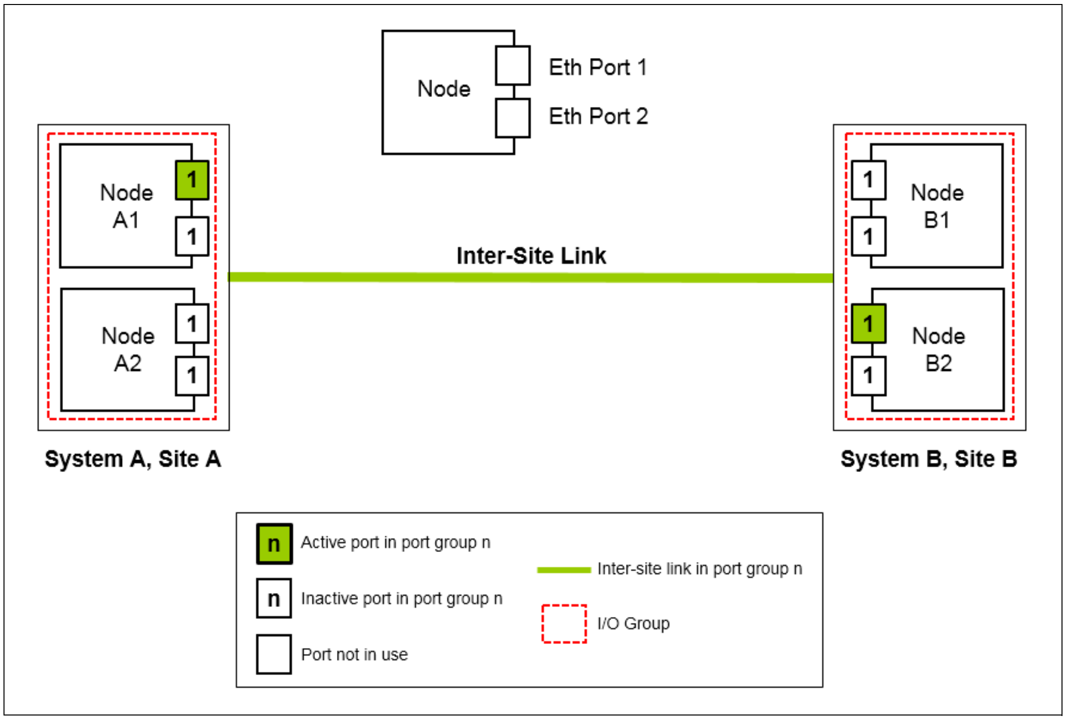


*Figure 11-91 Two node systems with single inter-site link and remote copy port groups configured*

As shown in Figure 11-91, this configuration is similar to Configuration 2, but differs because each node now has the same remote copy port group that is configured on more than one IP port.

On any node, only one port at any time can participate in IP partnership. Configuring multiple ports in the same remote copy group on the same node is not supported.

- An example of an unsupported configuration for a dual inter-site link is shown in Figure 11-92 (configuration 9).

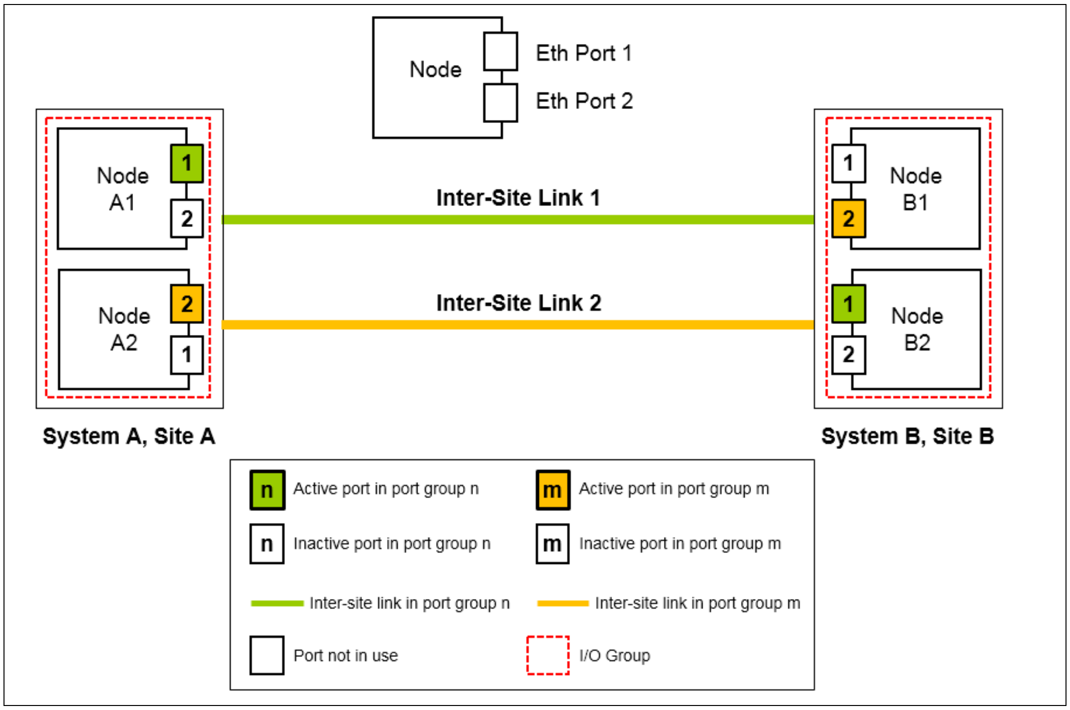

Node
Eth Port 1
Eth Port 2
Node A1
1
2
Node A2
2
1
Inter-Site Link 1
Inter-Site Link 2
Node B1
1
2
Node B2
1
2
System A, Site A
System B, Site B
n Active port in port group n
m Active port in port group m
n Inactive port in port group n
m Inactive port in port group m
Inter-site link in port group n
Inter-site link in port group m
Port not in use
I/O Group

*Figure 11-92 Dual Links with two Remote Copy Port Groups with failover Port Groups configured*

As shown in Figure 11-92, this configuration is similar to Configuration 5, but differs because each node now also has two ports that are configured with remote copy port groups. In this configuration, the path selection algorithm can select a path in a manner such that at times this might cause partnerships to change to the `Not_Present` state and then recover. This results in a configuration restriction, and the use of this configuration is not recommended until the configuration restriction is lifted in future releases.

- An example deployment for configuration 2 with a dedicated inter-site link is shown in Figure 11-93 (configuration 10).

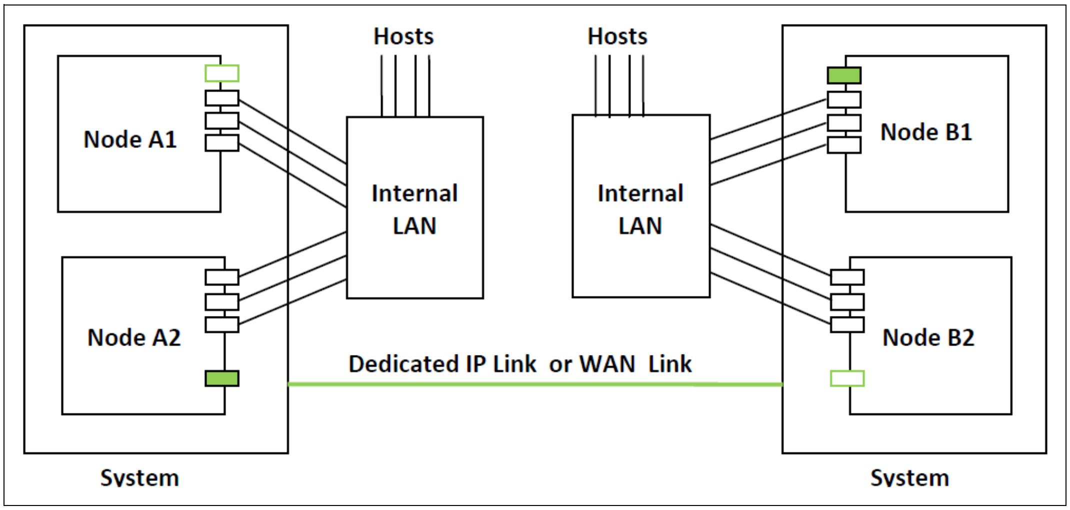


*Figure 11-93 Deployment example*

In this configuration, one port on each node in System A and System B is configured in remote copy group 1 to establish IP partnership and support remote copy relationships.

There is a dedicated inter-site link that is used for IP partnership traffic, and iSCSI host attach is disabled on those ports.

The following configuration steps are used:

a. Configure system IP addresses properly. As such, they can be reached over the inter-site link.

b. Qualify if the partnerships must be created over IPv4 or IPv6, and then assign IP addresses and open firewall ports 3260 and 3265.

c. Configure IP ports for remote copy on both the systems by using the following settings:

- Remote copy group: 1
- Host: No
- Assign IP address

d. Check that the maximum transmission unit (MTU) levels across the network meet the requirements as set (default MTU is 1500 on Storwize V7000).

e. Establish IP partnerships from both of the systems.

f. After the partnerships are in the `Fully_Configured` state, you can create the remote copy relationships.

- Example deployment for the configuration shown in Figure 11-87 on page 492. Ports that are shared with host access are shown in Figure 11-94 (configuration 11).

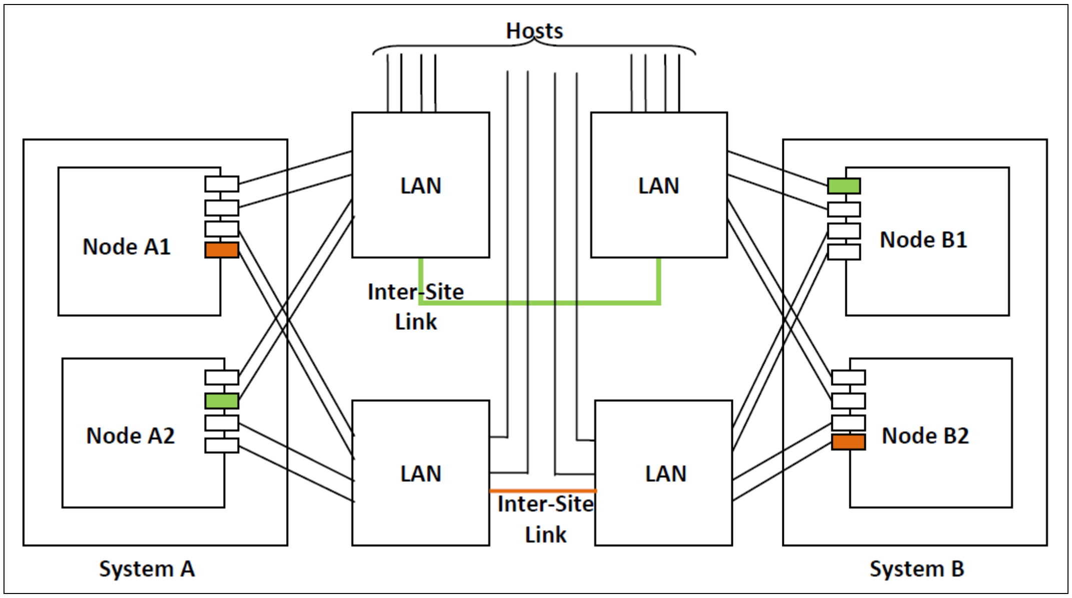


*Figure 11-94 Deployment example*

In this configuration, IP ports are to be shared by both iSCSI hosts and for IP partnership.

The following configuration steps are used:

a. Configure System IP addresses properly so that they can be reached over the inter-site link.

b. Qualify if the partnerships must be created over IPv4 or IPv6, and then assign IP addresses and open firewall ports 3260 and 3265.

c. Configure IP ports for remote copy on System A1 by using the following settings:

- Node 1:

  - Port 1, remote copy port group 1
  - Host: Yes
  - Assign IP address

- Node 2:

  - Port 4, Remote Copy Port Group 2
  - Host: Yes
  - Assign IP address

d. Configure IP ports for remote copy on System B1 by using the following settings:

- Node 1:

  - Port 1, remote copy port group 1
  - Host: Yes
  - Assign IP address

- Node 2:

  - Port 4, remote copy port group 2
  - Host: Yes
  - Assign IP address

e. Check the MTU levels across the network and meet the requirements as set (default MTU is 1500 on IBM Storwize V7000).

f. Establish IP partnerships from both systems.

g. After the partnerships are in the `Fully_Configured` state, you can create the remote copy relationships.

## 11.9 Remote Copy

In this section, we describe the Remote Copy services, which are a synchronous remote copy called Metro Mirror (MM), asynchronous remote copy called Global Mirror (GM), and Global Mirror with Change Volumes. Remote Copy in the IBM Storwize V7000 is similar to Remote Copy in the IBM System Storage DS8000 family at a functional level, but the implementation differs.

The IBM Storwize V7000 provides a single point of control when remote copy is enabled in your network (regardless of the disk subsystems that are used) if those disk subsystems are supported by the IBM Storwize V7000.

The general application of remote copy services is to maintain two real-time synchronized copies of a volume. Often, two copies are geographically dispersed between two IBM Storwize V7000 systems, although it is possible to use MM or GM within a single system (within an I/O Group). If the master copy fails, you can enable an auxiliary copy for I/O operation.

> **Tips:** Intracluster MM/GM uses more resources within the system when compared to an intercluster MM/GM relationship, where resource allocation is shared between the systems. Use intercluster MM/GM when possible. For mirroring volumes in the same system, it is better to use Volume Mirroring or the FlashCopy feature.

A typical application of this function is to set up a dual-site solution that uses two IBM Storwize V7000 systems. The first site is considered the primary or production site, and the second site is considered the backup site or failover site, which is activated when a failure at the first site is detected.

## 11.9.1 Multiple IBM Storwize V7000 system mirroring

Each IBM Storwize V7000 can maintain up to three partner system relationships, which enables as many as four systems to be directly associated with each other. This system partnership capability enables the implementation of disaster recovery (DR) solutions.

> **Note:** For more information about restrictions and limitations of native IP replication, see IP partnership limitations.

Figure 11-95 shows an example of a multiple system mirroring configuration.

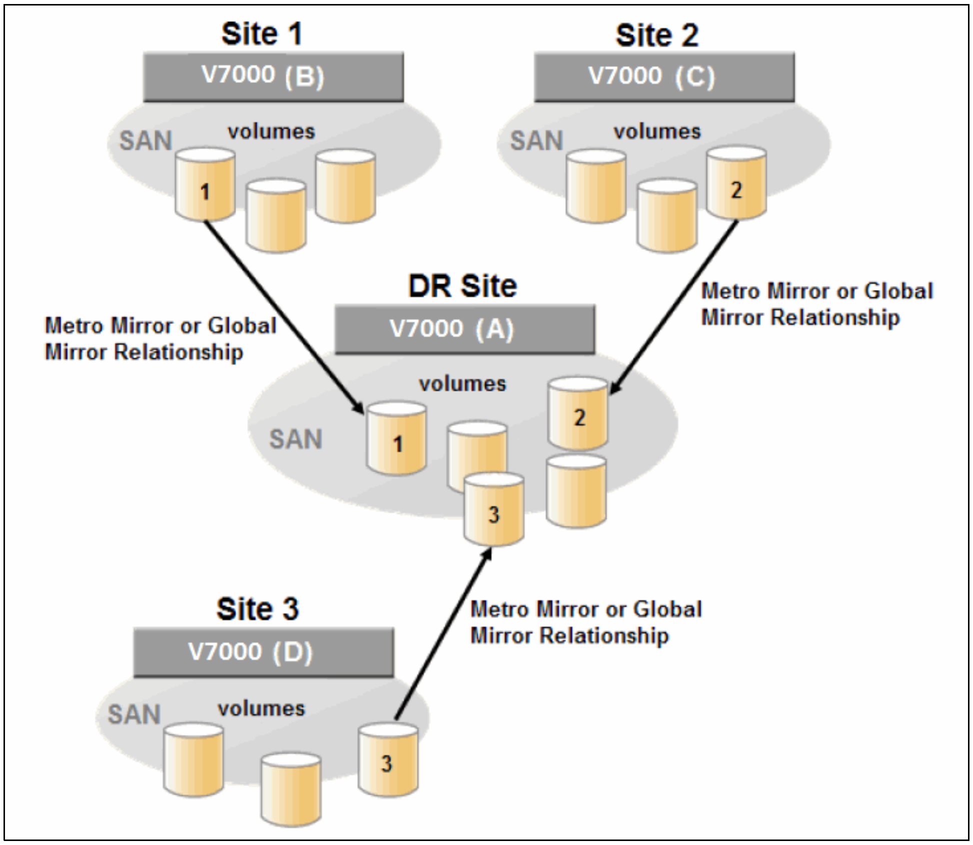

Site 1
V7000 (B)
SAN
volumes
1
Site 2
V7000 (C)
SAN
volumes
2
Metro Mirror or Global Mirror Relationship
DR Site
V7000 (A)
volumes
SAN
1
2
3
Metro Mirror or Global Mirror Relationship
Site 3
V7000 (D)
SAN
volumes
3
Metro Mirror or Global Mirror Relationship

*Figure 11-95 Multiple system mirroring configuration example*

### Supported multiple system mirroring topologies

Multiple system mirroring supports various partnership topologies, as shown in the example in Figure 11-96 on page 502.

The following example is a star topology: A → B, A → C, and A → D.

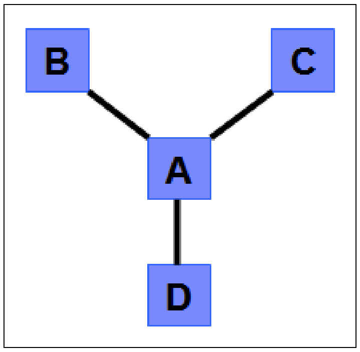


*Figure 11-96 Star topology*

Figure 11-96 shows four systems in a star topology, with System A at the center. System A can be a central DR site for the three other locations.

By using a star topology, you can migrate applications by using a process, such as the one described in the following example:

1. Suspend application at A.
2. Remove the A → B relationship.
3. Create the A → C relationship (or the B → C relationship).
4. Synchronize to system C, and ensure that A → C is established:
   - A → B, A → C, A → D, B → C, B → D, and C → D
   - A → B, A → C, and B → C

Figure 11-97 shows an example of a triangle topology: A → B, A → C, and B → C.

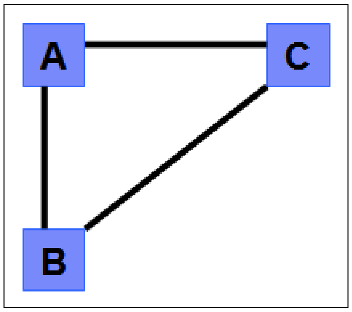


*Figure 11-97 Triangle topology*

Figure 11-98 on page 503 shows an example of an IBM Storwize V7000 fully connected topology: A → B, A → C, A → D, B → D, and C → D.

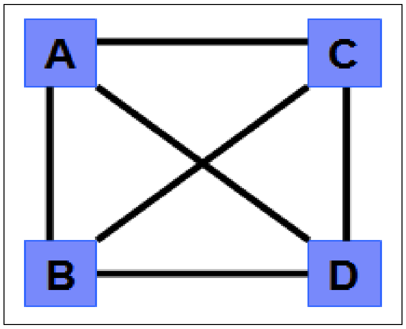


*Figure 11-98 Fully connected topology*

Figure 11-98 is a fully connected mesh in which every system has a partnership to each of the three other systems. This topology enables volumes to be replicated between any pair of systems, for example, A → B, A → C, and B → C.

Figure 11-99 shows a daisy-chain topology.

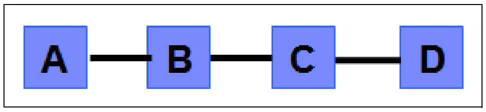


*Figure 11-99 Daisy-chain topology*

Although systems can have up to three partnerships, volumes can be part of only one remote copy relationship, for example, A → B.

**System partnership intermix:** All of the preceding topologies are valid for the intermix of the IBM SAN Volume Controller with the Storwize V7000 if the Storwize V7000 is set to the replication layer and running IBM Spectrum Virtualize code 6.3.0 or later.

## 11.9.2 Importance of write ordering

Many applications that use block storage have a requirement to survive failures, such as loss of power or a software crash, and to not lose data that existed before the failure. Because many applications must perform large numbers of update operations in parallel, maintaining write ordering is key to ensuring the correct operation of applications following a disruption.

An application that performs a high volume of database updates is designed with the concept of dependent writes. With dependent writes, it is important to ensure that an earlier write completed before a later write is started. Reversing or performing the order of writes differently than the application intended can undermine the application’s algorithms and can lead to problems, such as detected or undetected data corruption.

The IBM Spectrum Virtualize Metro Mirror and Global Mirror implementation operates in a manner that is designed to always keep a consistent image at the secondary site. The Global Mirror implementation uses complex algorithms that operate to identify sets of data and number those sets of data in sequence. The data is then applied at the secondary site in the defined sequence.

Operating in this manner ensures that if the relationship is in a `Consistent_Synchronized` state, the Global Mirror target data is at least crash consistent, and supports quick recovery through application crash recovery facilities.

For more information about dependent writes, see Consistency Groups.

### Remote Copy Consistency Groups

A Remote Copy Consistency Group can contain an arbitrary number of relationships up to the maximum number of MM/GM relationships that is supported by the IBM Spectrum Virtualize system. MM/GM commands can be issued to a Remote Copy Consistency Group.

Therefore, these commands can be issued simultaneously for all MM/GM relationships that are defined within that Consistency Group, or to a single MM/GM relationship that is not part of a Remote Copy Consistency Group. For example, when a **`startrcconsistgrp`** command is issued to the Consistency Group, all of the MM/GM relationships in the Consistency Group are started at the same time.

Figure 11-100 shows the concept of Metro Mirror Consistency Groups. The same applies to Global Mirror Consistency Groups.

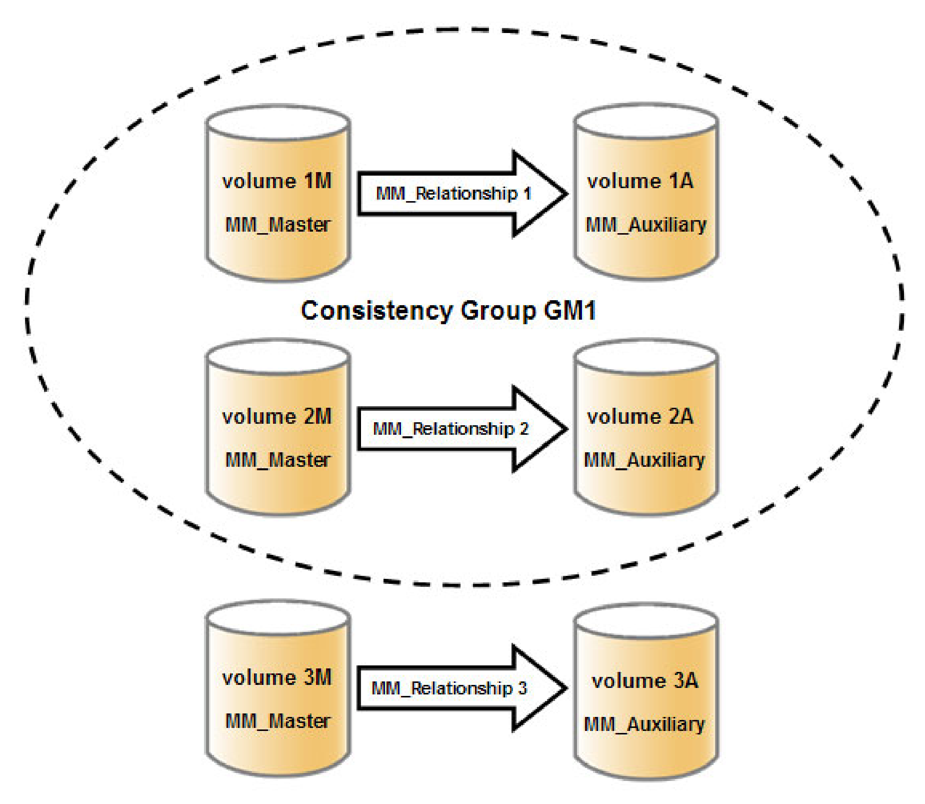


*Figure 11-100 Metro Mirror Consistency Group*

Because the MM_Relationship 1 and 2 are part of the Consistency Group, they can be handled as one entity. The stand-alone MM_Relationship 3 is handled separately.

Certain uses of MM/GM require the manipulation of more than one relationship. Remote Copy Consistency Groups can group relationships so that they are manipulated in unison.

Consider the following points:

- MM/GM relationships can be part of a Consistency Group, or they can be stand-alone and, therefore, are handled as single instances.

- A Consistency Group can contain zero or more relationships. An empty Consistency Group with zero relationships in it has little purpose until it is assigned its first relationship, except that it has a name.
- All relationships in a Consistency Group must have corresponding master and auxiliary volumes.
- All relationships in one Consistency Group must be the same type, for example only Metro Mirror or only Global Mirror.

Although Consistency Groups can be used to manipulate sets of relationships that do not need to satisfy these strict rules, this manipulation can lead to undesired side effects. The rules behind a Consistency Group mean that certain configuration commands are prohibited. These configuration commands are not prohibited if the relationship is not part of a Consistency Group.

For example, consider the case of two applications that are independent, yet they are placed into a single Consistency Group. If an error occurs, synchronization is lost and a background copy process is required to recover synchronization. While this process is progressing, MM/GM rejects attempts to enable access to the auxiliary volumes of either application.

If one application finishes its background copy more quickly than the other application, MM/GM still refuses to grant access to its auxiliary volumes even though it is safe in this case. The MM/GM policy is to refuse access to the entire Consistency Group if any part of it is inconsistent.

Stand-alone relationships and Consistency Groups share a common configuration and state model. All of the relationships in a non-empty Consistency Group have the same state as the Consistency Group.

## 11.9.3 Remote copy intercluster communication

In the traditional Fibre Channel (FC), the intercluster communication between systems in a Metro Mirror and Global Mirror partnership is performed over the SAN. In the following section, we describe this communication path.

For more information about intercluster communication between systems in an IP partnership, see States of IP partnership.

### Zoning

The IBM Storwize V7000 FC ports on each system must communicate with each other to create the partnership. Switch zoning is critical to facilitating intercluster communication.

### Intercluster communication channels

When an IBM Spectrum Virtualize system partnership is defined on a pair of systems, the following intercluster communication channels are established:

- A single control channel, which is used to exchange and coordinate configuration information
- I/O channels between each of these nodes in the systems

These channels are maintained and updated as nodes and links appear and disappear from the fabric, and are repaired to maintain operation where possible. If communication between the systems is interrupted or lost, an event is logged (and the Metro Mirror and Global Mirror relationships stop).

**Alerts:** You can configure the system to raise Simple Network Management Protocol (SNMP) traps to the enterprise monitoring system to alert on events that indicate an interruption in internode communication occurred.

### Intercluster links

All IBM Storwize V7000 node canisters maintain a database of other devices that are visible on the fabric. This database is updated as devices appear and disappear.

Devices that advertise themselves as IBM SAN Volume Controller or Storwize V7000 nodes are categorized according to the system to which they belong. Nodes that belong to the same system establish communication channels between themselves and begin to exchange messages to implement clustering and the functional protocols of IBM Spectrum Virtualize.

Nodes that are in separate systems do not exchange messages after initial discovery is complete, unless they are configured together to perform a remote copy relationship.

The intercluster link carries control traffic to coordinate activity between two systems. The link is formed between one node in each system. The traffic between the designated nodes is distributed among logins that exist between those nodes.

If the designated node fails (or all of its logins to the remote system fail), a new node is chosen to carry control traffic. This node change causes the I/O to pause, but it does not put the relationships in a `ConsistentStopped` state.

**Note:** It is advised to use **`chsystem`** with **`-partnerfcportmask`** to dedicate several FC ports only to system-to-system traffic to ensure that remote copy is not affected by other traffic, such as host-to-node traffic or node-to-node traffic within the same system.

## 11.9.4 Metro Mirror overview

Metro Mirror establishes a synchronous relationship between two volumes of equal size. The volumes in a Metro Mirror relationship are referred to as the master (primary) volume and the auxiliary (secondary) volume. Traditional FC Metro Mirror is primarily used in a metropolitan area or geographical area, up to a maximum distance of 300 km (186.4 miles) to provide synchronous replication of data.

With synchronous copies, host applications write to the master volume, but they do not receive confirmation that the write operation completed until the data is written to the auxiliary volume. This action ensures that both the volumes have identical data when the copy completes. After the initial copy completes, the Metro Mirror function always maintains a fully synchronized copy of the source data at the target site.

Metro Mirror has the following characteristics:

- Zero recovery point objective (RPO)
- Synchronous
- Production application performance that is affected by round-trip latency

Increased distance directly affects host I/O performance because the writes are synchronous. Use the requirements for application performance when you are selecting your Metro Mirror auxiliary location.

Consistency Groups can be used to maintain data integrity for dependent writes, which is similar to FlashCopy Consistency Groups (FlashCopy Consistency Groups are described in Implementing FlashCopy).

The IBM Spectrum Virtualize provides intracluster and intercluster Metro Mirror.

### Intracluster Metro Mirror

Intracluster Metro Mirror performs the intracluster copying of a volume, in which both volumes belong to the same system and I/O Group within the system. Because it is within the same I/O Group, there must be sufficient bitmap space within the I/O Group for both sets of volumes and licensing on the system.

**Important:** Performing Metro Mirror across I/O Groups within a system is not supported.

### Intercluster Metro Mirror

Intercluster Metro Mirror performs intercluster copying of a volume, in which one volume belongs to a system and the other volume belongs to a separate system.

Two IBM Spectrum Virtualize systems must be defined in a partnership, which must be performed on both systems to establish a fully functional Metro Mirror partnership.

By using standard single-mode connections, the supported distance between two systems in a Metro Mirror partnership is 10 km (6.2 miles), although greater distances can be achieved by using extenders. For extended distance solutions, contact your IBM representative.

**Limit:** When a local fabric and a remote fabric are connected for Metro Mirror purposes, the inter-switch link (ISL) hop count between a local node and a remote node cannot exceed seven.

## 11.9.5 Synchronous remote copy

Metro Mirror is a fully synchronous remote copy technique that ensures that writes are committed at both the master and auxiliary volumes before write completion is acknowledged to the host, but only if writes to the auxiliary volumes are possible.

Events, such as a loss of connectivity between systems, can cause mirrored writes from the master volume and the auxiliary volume to fail. In that case, Metro Mirror suspends writes to the auxiliary volume and enables I/O to the master volume to continue to avoid affecting the operation of the master volumes.

Figure 11-100 on page 504 shows how a write to the master volume is mirrored to the cache of the auxiliary volume before an acknowledgment of the write is sent back to the host that issued the write. This process ensures that the auxiliary is synchronized in real time if it is needed in a failover situation.

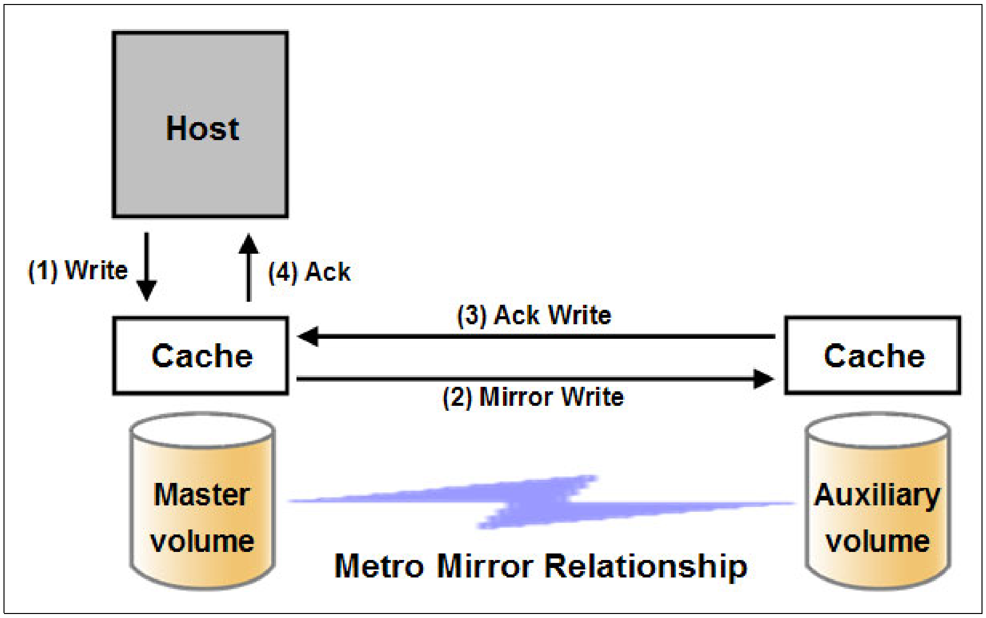


*Figure 11-101 Write on volume in Metro Mirror relationship*

However, this process also means that the application is exposed to the latency and bandwidth limitations (if any) of the communication link between the master and auxiliary volumes. This process might lead to unacceptable application performance, particularly when placed under peak load. Therefore, the use of traditional Fibre Channel Metro Mirror has distance limitations that are based on your performance requirements. The IBM Spectrum Virtualize does not support more than 300 km (186.4 miles).

## 11.9.6 Metro Mirror features

The IBM Spectrum Virtualize Metro Mirror function supports the following features:

- Synchronous remote copy of volumes that are dispersed over metropolitan distances.
- The Metro Mirror relationships between volume pairs, with each volume in a pair that is managed by a Storwize V7000 system or IBM SAN Volume Controller system (requires V6.3.0 or later).
- Supports intracluster Metro Mirror where both volumes belong to the same system (and I/O Group).
- The IBM Spectrum Virtualize supports intercluster Metro Mirror where each volume belongs to a separate systems. You can configure a specific system for partnership with another system. All intercluster Metro Mirror processing occurs between two IBM Spectrum Virtualize systems that are configured in a partnership.
- Intercluster and intracluster Metro Mirror can be used concurrently.
- The IBM Storwize V7000 does not require that a control network or fabric is installed to manage Metro Mirror. For intercluster Metro Mirror, the system maintains a control link between two systems. This control link is used to control the state and coordinate updates at either end. The control link is implemented on top of the same FC fabric connection that the IBM Storwize V7000 uses for Metro Mirror I/O.
- The IBM Spectrum Virtualize implements a configuration model that maintains the Metro Mirror configuration and state through major events, such as failover, recovery, and resynchronization, to minimize user configuration action through these events.

The IBM Spectrum Virtualize supports the resynchronization of changed data so that write failures that occur on the master or auxiliary volumes do not require a complete resynchronization of the relationship.

## 11.9.7 Metro Mirror attributes

The Metro Mirror function in IBM Spectrum Virtualize possesses the following attributes:

- A partnership is created between two IBM Storwize V7000 systems or an IBM SAN Volume Controller system and IBM Storwize V7000 operating in the replication layer (for intercluster Metro Mirror).
- A Metro Mirror relationship is created between two volumes of the same size.
- To manage multiple Metro Mirror relationships as one entity, relationships can be made part of a Metro Mirror Consistency Group, which ensures data consistency across multiple Metro Mirror relationships and provides ease of management.
- When a Metro Mirror relationship is started and when the background copy completes, the relationship becomes consistent and synchronized.
- After the relationship is synchronized, the auxiliary volume holds a copy of the production data at the primary, which can be used for DR.
- The auxiliary volume is in read-only mode when relationship is active.
- To access the auxiliary volume, the Metro Mirror relationship must be stopped with the access option enabled, before write I/O is allowed to the auxiliary.
- The remote host server is mapped to the auxiliary volume, and the disk is available for I/O.

## 11.9.8 Practical use of Metro Mirror

The master volume is the production volume, and updates to this copy are mirrored in real time to the auxiliary volume. The contents of the auxiliary volume that existed when the relationship was created are destroyed.

> **Switching copy direction:** The copy direction for a Metro Mirror relationship can be switched so that the auxiliary volume becomes the master, and the master volume becomes the auxiliary, which is similar to the FlashCopy restore option. However, although the FlashCopy target volume can operate in read/write mode, the target volume of the started remote copy is always in read-only mode.

While the Metro Mirror relationship is active, the auxiliary volume is not accessible for host application write I/O at any time. The IBM Storwize V7000 allows read-only access to the auxiliary volume when it contains a consistent image. Storwize allows boot time operating system discovery to complete without an error, so that any hosts at the secondary site can be ready to start the applications with minimum delay, if required.

For example, many operating systems must read LBA zero to configure a logical unit. Although read access is allowed at the auxiliary in practice, the data on the auxiliary volumes cannot be read by a host because most operating systems write a “dirty bit” to the file system when it is mounted. Because this write operation is not allowed on the auxiliary volume, the volume cannot be mounted.

This access is provided only where consistency can be ensured. However, coherency cannot be maintained between reads that are performed at the auxiliary and later write I/Os that are performed at the master.

To enable access to the auxiliary volume for host operations, you must stop the Metro Mirror relationship by specifying the `-access` parameter. While access to the auxiliary volume for host operations is enabled, the host must be instructed to mount the volume before the application can be started, or instructed to perform a recovery process.

For example, the Metro Mirror requirement to enable the auxiliary copy for access differentiates it from third-party mirroring software on the host, which aims to emulate a single, reliable disk regardless of what system is accessing it. Metro Mirror retains the property that there are two volumes in existence but it suppresses one volume while the copy is being maintained.

The use of an auxiliary copy demands a conscious policy decision by the administrator that a failover is required, and that the tasks to be performed on the host that is involved in establishing the operation on the auxiliary copy are substantial. The goal is to make this copy rapid (much faster when compared to recovering from a backup copy) but not seamless.

The failover process can be automated through failover management software. The IBM Storwize V7000 provides SNMP traps and programming (or scripting) for the CLI to enable this automation.

## 11.9.9 Global Mirror Overview

In the following topics, we describe the Global Mirror copy service, which is an asynchronous remote copy service. It provides and maintains a consistent mirrored copy of a source volume to a target volume.

Global Mirror establishes a Global Mirror relationship between two volumes of equal size. The volumes in a Global Mirror relationship are referred to as the *master* (source) volume and the *auxiliary* (target) volume, which is the same as Metro Mirror. Consistency Groups can be used to maintain data integrity for dependent writes, which is similar to FlashCopy Consistency Groups.

Global Mirror writes data to the auxiliary volume asynchronously, which means that host writes to the master volume provide the host with confirmation that the write is complete before the I/O completes on the auxiliary volume.

Global Mirror has the following characteristics:

- Near-zero RPO
- Asynchronous
- Production application performance that is affected by I/O sequencing preparation time

### *Intracluster Global Mirror*

Although Global Mirror is available for intracluster, it has no functional value for production use. Intracluster Metro Mirror provides the same capability with less processor use. However, leaving this functionality in place simplifies testing and supports client experimentation and testing (for example, to validate server failover on a single test system). As with Intracluster Metro Mirror, you must consider the increase in the license requirement, because source and target exist on the same IBM Spectrum Virtualize system.

### *Intercluster Global Mirror*

Intercluster Global Mirror operations require a pair of IBM Spectrum Virtualize systems that are connected by several intercluster links. The two systems must be defined in a partnership to establish a fully functional Global Mirror relationship.

**Limit:** When a local fabric and a remote fabric are connected for Global Mirror purposes, the ISL hop count between a local node and a remote node must not exceed seven hops.

### 11.9.10 Asynchronous remote copy

Global Mirror is an asynchronous remote copy technique. In asynchronous remote copy, the write operations are completed on the primary site and the write acknowledgment is sent to the host before it is received at the secondary site. An update of this write operation is sent to the secondary site at a later stage, which provides the capability to perform remote copy over distances that exceed the limitations of synchronous remote copy.

The Global Mirror function provides the same function as Metro Mirror remote copy, but over long-distance links with higher latency without requiring the hosts to wait for the full round-trip delay of the long-distance link.

Figure 11-102 shows that a write operation to the master volume is acknowledged back to the host that is issuing the write before the write operation is mirrored to the cache for the auxiliary volume.

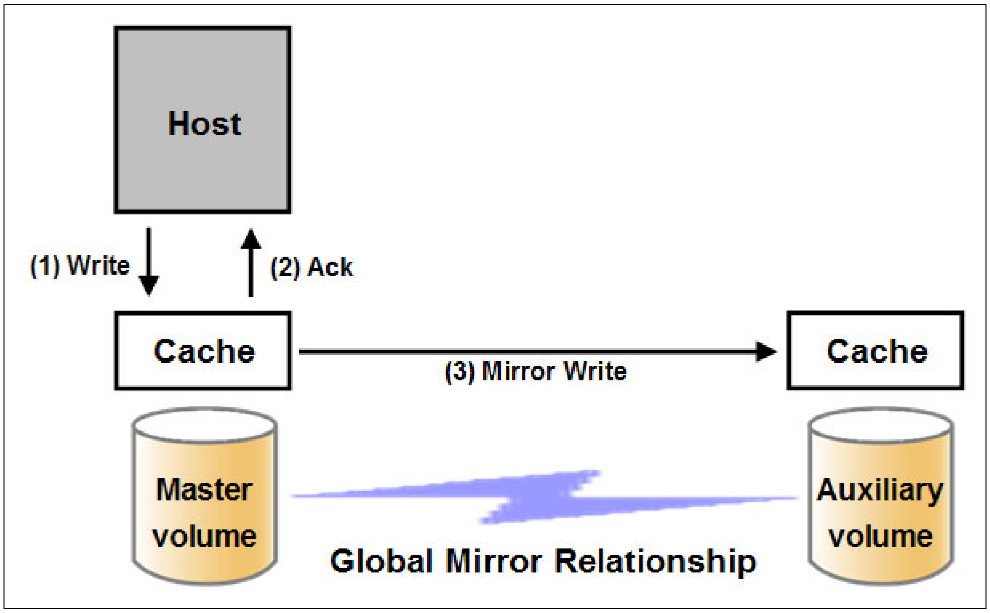


*Figure 11-102 Global Mirror write sequence*

The Global Mirror algorithms maintain a consistent image on the auxiliary always. They achieve this consistent image by identifying sets of I/Os that are active concurrently at the master, assigning an order to those sets, and applying those sets of I/Os in the assigned order at the secondary. As a result, Global Mirror maintains the features of Write Ordering and Read Stability.

The multiple I/Os within a single set are applied concurrently. The process that marshals the sequential sets of I/Os operates at the secondary system. Therefore, the process is not subject to the latency of the long-distance link. These two elements of the protocol ensure that the throughput of the total system can be grown by increasing system size while maintaining consistency across a growing data set.

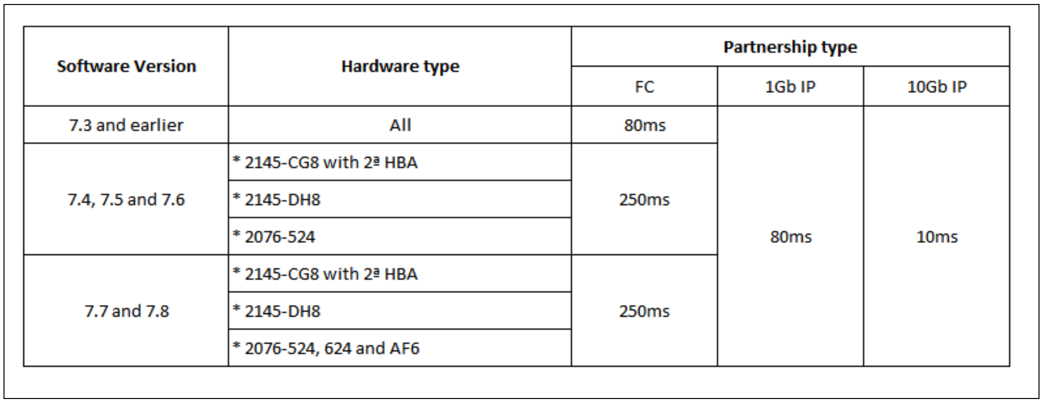
Global Mirror write I/O from production system to a secondary system requires serialization and sequence-tagging before being sent across the network to a remote site (to maintain a write-order consistent copy of data).

To avoid affecting the production site, IBM Spectrum Virtualize supports more parallelism in processing and managing Global Mirror writes on the secondary system by using the following methods:

- Secondary system nodes store replication writes in new redundant non-volatile cache
- Cache content details are shared between nodes
- Cache content details are batched together to make node-to-node latency less of an issue
- Nodes intelligently apply these batches in parallel as soon as possible
- Nodes internally manage and optimize Global Mirror secondary write I/O processing

In a failover scenario where the secondary site must become the master source of data, certain updates might be missing at the secondary site. Therefore, any applications that use this data must have an external mechanism for recovering the missing updates and reapplying them; for example, a transaction log replay.

Global Mirror is supported over FC, FC over IP (FCIP), FC over Ethernet (FCoE), and native IP connections. The maximum distance cannot exceed 80 ms round trip, which is about 4000 km (2485.48 miles) between mirrored systems. But starting with IBM Spectrum Virtualize V7.4, this distance was significantly increased for certain IBM Storwize Gen2 and IBM SAN Volume Controller configurations. Figure 11-103 shows the current supported distances for Global Mirror remote copy.



| Software Version | Hardware type | Partnership type | | |
|---|---|---|---|---|
| | | FC | 1Gb IP | 10Gb IP |
| 7.3 and earlier | All | 80ms | 80ms | 10ms |
| 7.4, 7.5 and 7.6 | * 2145-CG8 with 2ª HBA | 250ms | | |
| | * 2145-DH8 | | | |
| | * 2076-524 | | | |
| 7.7 and 7.8 | * 2145-CG8 with 2ª HBA | 250ms | | |
| | * 2145-DH8 | | | |
| | * 2076-524, 624 and AF6 | | | |

*Figure 11-103 Supported Global Mirror distances*

## 11.9.11 Global Mirror features

IBM Spectrum Virtualize Global Mirror supports the following features:

- Asynchronous remote copy of volumes that are dispersed over metropolitan-scale distances.
- The IBM Spectrum Virtualize implements the Global Mirror relationship between a volume pair, with each volume in the pair being managed by an IBM SAN Volume Controller or IBM Storwize V7000 running IBM Spectrum Virtualize.
- The IBM Storwize V7000 supports intracluster Global Mirror where both volumes belong to the same system (and I/O Group).
- The IBM Storwize V7000 intercluster Global Mirror in which each volume belongs to its separate IBM Storwize V7000 system. An IBM Storwize V7000 system can be configured

for partnership with 1 - 3 other systems. For more information about IP partnership restrictions, see IP partnership limitations.

- Intercluster and intracluster Global Mirror can be used concurrently, but not for the same volume.
- The IBM Storwize V7000 does not require a control network or fabric to be installed to manage Global Mirror. For intercluster Global Mirror, the IBM Storwize V7000 maintains a control link between the two systems. This control link is used to control the state and to coordinate the updates at either end. The control link is implemented on top of the same FC fabric connection that the IBM Storwize V7000 uses for Global Mirror I/O.
- The IBM Storwize V7000 implements a configuration model that maintains the Global Mirror configuration and state through major events, such as failover, recovery, and resynchronization, to minimize user configuration action through these events.
- The IBM Storwize V7000 implements flexible resynchronization support, enabling it to resynchronize volume pairs that experienced write I/Os to both disks, and to resynchronize only those regions that changed.
- An optional feature for Global Mirror is a delay simulation to be applied on writes that are sent to auxiliary volumes. It is useful in intracluster scenarios for testing purposes.

### Colliding writes

Before V4.3.1, the Global Mirror algorithm required that only a single write is active on any 512-byte logical block address (LBA) of a volume. If a further write is received from a host while the auxiliary write is still active (even though the master write might complete), the new host write is delayed until the auxiliary write is complete. This restriction is needed if a series of writes to the auxiliary must be tried again (which is called *reconstruction*). Conceptually, the data for reconstruction comes from the master volume.

If multiple writes are allowed to be applied to the master for a sector, only the most recent write gets the correct data during reconstruction. If reconstruction is interrupted for any reason, the intermediate state of the auxiliary is inconsistent. Applications that deliver such write activity do not achieve the performance that Global Mirror is intended to support. A volume statistic is maintained about the frequency of these collisions.

An attempt is made to allow multiple writes to a single location to be outstanding in the Global Mirror algorithm. There is still a need for master writes to be serialized, and the intermediate states of the master data must be kept in a non-volatile journal while the writes are outstanding to maintain the correct write ordering during reconstruction. Reconstruction must never overwrite data on the auxiliary with an earlier version. The volume statistic that is monitoring colliding writes is now limited to those writes that are not affected by this change.

Figure 11-104 shows a colliding write sequence example.

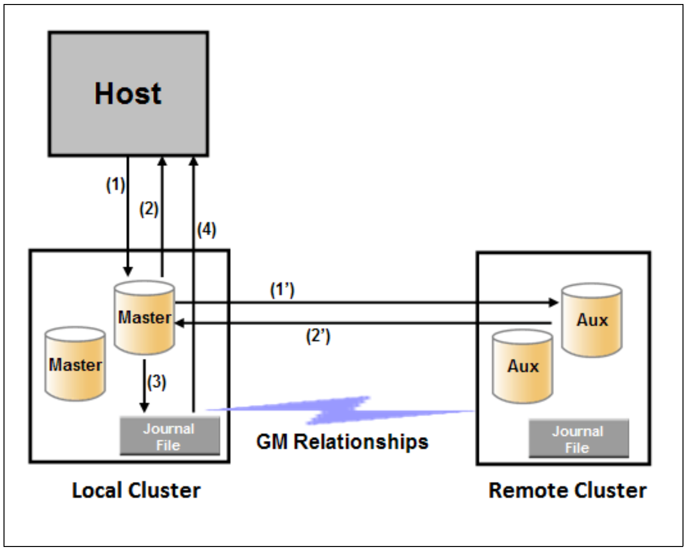


*Figure 11-104 Colliding writes example*

The following numbers correspond to the numbers that are shown in Figure 11-104:

- (1) The first write is performed from the host to LBA X.
- (2) The host is provided acknowledgment that the write completed even though the mirrored write to the auxiliary volume is not yet complete.
- (1') and (2') occur asynchronously with the first write.
- (3) The second write is performed from the host also to LBA X. If this write occurs before (2'), the write is written to the journal file.
- (4) The host is provided acknowledgment that the second write is complete.

### Delay simulation

An optional feature for Global Mirror enables a delay simulation to be applied on writes that are sent to auxiliary volumes. This feature enables you to perform testing that detects colliding writes. Therefore, you can use this feature to test an application before the full deployment of the feature. The feature can be enabled separately for each of the intracluster or intercluster Global Mirrors.

You specify the delay setting by using the **`chsystem`** command and view the delay by using the **`lssystem`** command. The `gm_intra_cluster_delay_simulation` field expresses the amount of time that intracluster auxiliary I/Os are delayed. The `gm_inter_cluster_delay_simulation` field expresses the amount of time that intercluster auxiliary I/Os are delayed. A value of zero disables the feature.

**Tip:** If you are experiencing repeated problems with the delay on your link, make sure that the delay simulator was properly disabled.

## 11.9.12 Using Change Volumes with Global Mirror

Global Mirror is designed to achieve an RPO as low as possible so that data is as up-to-date as possible. This design places several strict requirements on your infrastructure. In certain situations with low network link quality, congested hosts, or overloaded hosts, you might be affected by multiple `1920` congestion errors.

Congestion errors happen in the following primary situations:

- Congestion at the source site through the host or network
- Congestion in the network link or network path
- Congestion at the target site through the host or network

Global Mirror has functionality that is designed to address the following conditions, which might negatively affect certain Global Mirror implementations:

- The estimation of the bandwidth requirements tends to be complex.
- Ensuring the latency and bandwidth requirements can be met is often difficult.
- Congested hosts on the source or target site can cause disruption.
- Congested network links can cause disruption with only intermittent peaks.

To address these issues, *Change Volumes* were added as an option for Global Mirror relationships. Change Volumes use the FlashCopy functionality, but they cannot be manipulated as FlashCopy volumes because they are for a special purpose only. Change Volumes replicate point-in-time images on a cycling period. The default is 300 seconds.

Your change rate needs to include only the condition of the data at the point-in-time that the image was taken, rather than all the updates during the period. The use of this function can provide significant reductions in replication volume.

Global Mirror with Change Volumes has the following characteristics:

- Larger RPO
- Point-in-time copies
- Asynchronous
- Possible system performance resource requirements because point-in-time copies are created locally

Figure 11-105 shows a simple Global Mirror relationship without Change Volumes.

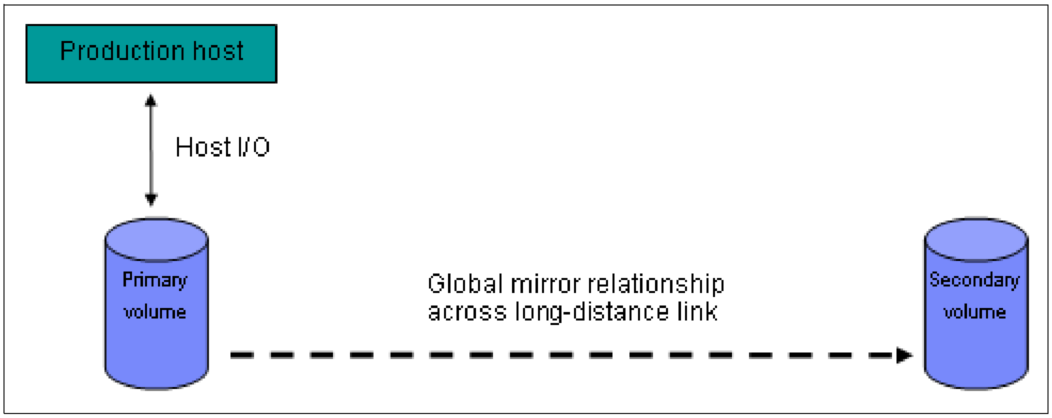


*Figure 11-105 Global Mirror without Change Volumes*

With Change Volumes, this environment looks as it is shown in Figure 11-106 on page 516.

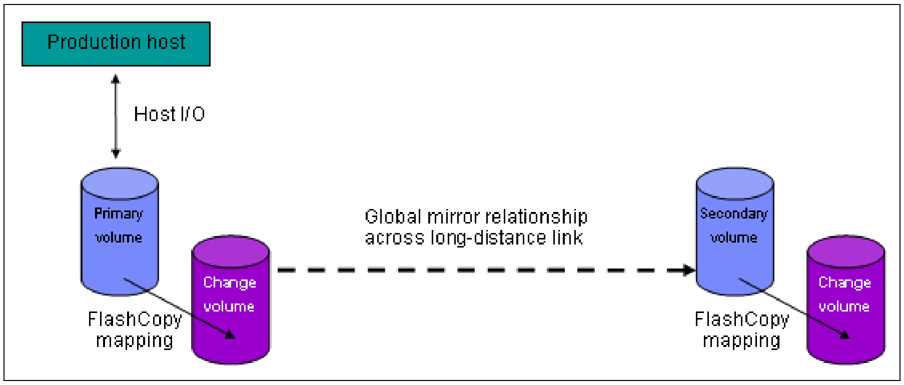


*Figure 11-106 Global Mirror with Change Volumes*

With Change Volumes, a FlashCopy mapping exists between the primary volume and the primary Change Volume. The mapping is updated on the cycling period (60 seconds to one day). The primary Change Volume is then replicated to the secondary Global Mirror volume at the target site, which is then captured in another Change Volume on the target site. This approach provides an always consistent image at the target site and protects your data from being inconsistent during resynchronization.

How Change Volumes might save you replication traffic is shown in Figure 11-107.

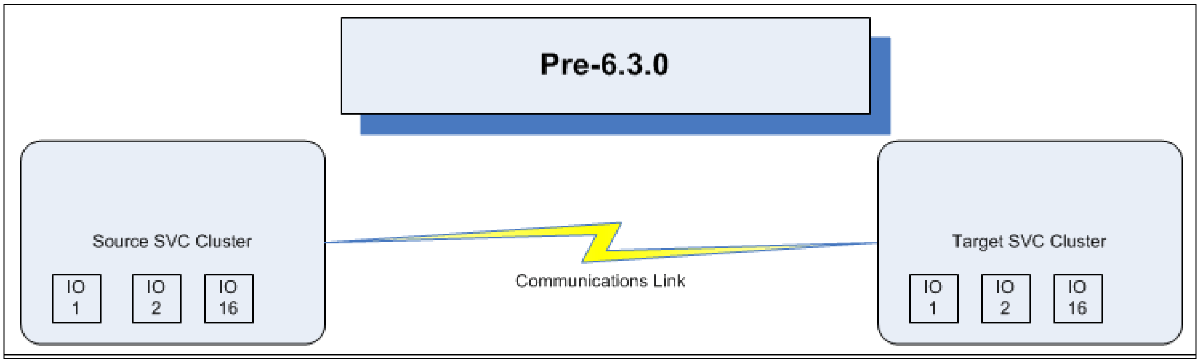


*Figure 11-107 Global Mirror I/O replication without Change Volumes*

In Figure 11-107, you can see several I/Os on the source and the same number on the target, and in the same order. Assuming that this data is the same set of data being updated repeatedly, this approach results in wasted network traffic. The I/O can be completed much more efficiently, as shown in Figure 11-108.

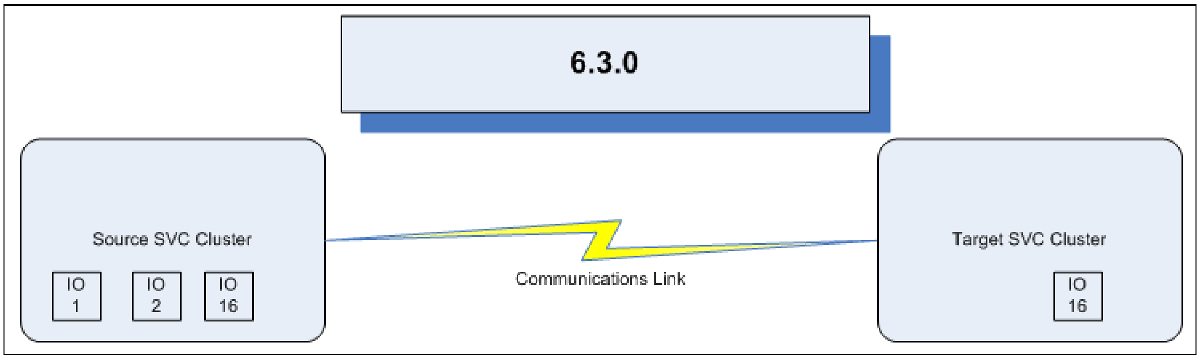


*Figure 11-108 Global Mirror I/O with Change Volumes V6.3.0 and beyond*

In Figure 11-108 on page 516, the same data is being updated repeatedly. Therefore, Change Volumes demonstrate significant I/O transmission savings by needing to send I/O number 16 only, which was the last I/O before the cycling period.

You can adjust the cycling period by using the **`chrcrelationship -cycleperiodseconds`** ***`<60 - 86400>`*** command from the CLI. If a copy does not complete in the cycle period, the next cycle does not start until the prior cycle completes. For this reason, the use of Change Volumes gives you the following possibilities for RPO:

- If your replication completes in the cycling period, your RPO is twice the cycling period.
- If your replication does not complete within the cycling period, RPO is twice the completion time. The next cycling period starts immediately after the prior cycling period is finished.

Carefully consider your business requirements versus the performance of Global Mirror with Change Volumes. Global Mirror with Change Volumes increases the intercluster traffic for more frequent cycling periods. Therefore, selecting the shortest cycle periods possible is not always the answer. In most cases, the default must meet requirements and perform well.

**Important:** When you create your Global Mirror volumes with Change Volumes, make sure that you remember to select the Change Volume on the auxiliary (target) site. Failure to do so leaves you exposed during a resynchronization operation.

## 11.9.13 Distribution of work among nodes

For the best performance, MM/GM volumes must have their preferred nodes evenly distributed among the nodes of the systems. Each volume within an I/O Group has a preferred node property that can be used to balance the I/O load between nodes in that group. MM/GM also uses this property to route I/O between systems.

If this preferred practice is not maintained, for example, source volumes are assigned to only one node in the I/O group, you can change the preferred node for each volume to distribute volumes evenly between the nodes. You can also change the preferred node for volumes that are in a remote copy relationship without affecting the host I/O to a particular volume.

The remote copy relationship type does not matter. (The remote copy relationship type can be MM, GM, or GM with Change Volumes.) You can change the preferred node both to the source and target volumes that are participating in the remote copy relationship.

## 11.9.14 Background copy performance

The background copy performance is subject to sufficient Redundant Array of Independent Disks (RAID) controller bandwidth. Performance is also subject to other potential bottlenecks, such as the intercluster fabric, and possible contention from host I/O for the IBM SAN Volume Controller or Storwize V7000 bandwidth resources.

Background copy I/O is scheduled to avoid bursts of activity that might have an adverse effect on system behavior. An entire grain of tracks on one volume is processed at around the same time but not as a single I/O. Double buffering is used to try to use sequential performance within a grain. However, the next grain within the volume might not be scheduled for some time. Multiple grains might be copied simultaneously, and might be enough to satisfy the requested rate, unless the available resources cannot sustain the requested rate.

Global Mirror paces the rate at which background copy is performed by the appropriate relationships. Background copy occurs on relationships that are in the `InconsistentCopying` state with a status of `Online`.

The quota of background copy (configured on the intercluster link) is divided evenly between all nodes that are performing background copy for one of the eligible relationships. This allocation is made irrespective of the number of disks for which the node is responsible. Each node in turn divides its allocation evenly between the multiple relationships that are performing a background copy.

The default value of the background copy is 25 megabytes per second (MBps), per volume.

> **Important:** The background copy value is a system-wide parameter that can be changed dynamically but only on a per-system basis and not on a per-relationship basis. Therefore, the copy rate of all relationships changes when this value is increased or decreased. In systems with many remote copy relationships, increasing this value might affect overall system or intercluster link performance. The background copy rate can be changed between 1 - 1000 MBps.

## 11.9.15 Thin-provisioned background copy

Metro Mirror and Global Mirror relationships preserve the space-efficiency of the master. Conceptually, the background copy process detects a deallocated region of the master and sends a special *zero buffer* to the auxiliary.

If the auxiliary volume is thin-provisioned and the region is deallocated, the special buffer prevents a write and, therefore, an allocation. If the auxiliary volume is not thin-provisioned or the region in question is an allocated region of a thin-provisioned volume, a buffer of “real” zeros is synthesized on the auxiliary and written as normal.

## 11.9.16 Methods of synchronization

This section describes two methods that can be used to establish a synchronized relationship.

### Full synchronization after creation

The full synchronization after creation method is the default method. It is the simplest method in that it requires no administrative activity apart from issuing the necessary commands. However, in certain environments, the available bandwidth can make this method unsuitable.

Use the following command sequence for a single relationship:

- Run **`mkrcrelationship`** without specifying the **`-sync`** option.
- Run **`startrcrelationship`** without specifying the **`-clean`** option.

### Synchronized before creation

In this method, the administrator must ensure that the master and auxiliary volumes contain identical data before creating the relationship by using the following technique:

- Both disks are created with the security delete feature to make all data zero.
- A complete tape image (or other method of moving data) is copied from one disk to the other disk.

With this technique, do not allow I/O on the master or auxiliary before the relationship is established.

Then, the administrator must run the following commands:

- Run **`mkrcrelationship`** with the **`-sync`** flag.

- Run **`startrcrelationship`** without the **`-clean`** flag.

> **Important:** Failure to perform these steps correctly can cause MM/GM to report the relationship as consistent when it is not, therefore creating a data loss or data integrity exposure for hosts accessing data on the auxiliary volume.

## 11.9.17 Practical use of Global Mirror

The practical use of Global Mirror is similar to the Metro Mirror described in Practical use of Metro Mirror. The main difference between the two remote copy modes is that Global Mirror and Global Mirror with Change Volumes are mostly used on much larger distances than Metro Mirror. Weak link quality or insufficient bandwidth between the primary and secondary sites can also be a reason to prefer asynchronous Global Mirror over synchronous Metro Mirror. Otherwise, the use cases for Metro Mirror and Global Mirror are the same.

## 11.9.18 Valid combinations of FlashCopy, Metro Mirror, and Global Mirror

Table 11-9 lists the combinations of FlashCopy and Metro Mirror or Global Mirror functions that are valid for a single volume.

*Table 11-9 Valid combination for a single volume*

| FlashCopy | Metro Mirror or Global Mirror source | Metro Mirror or Global Mirror target |
|---|---|---|
| FlashCopy Source | Supported | Supported |
| FlashCopy Target | Supported | Not supported |

## 11.9.19 Remote Copy configuration limits

Table 11-10 lists the Metro Mirror and Global Mirror configuration limits.

*Table 11-10 Metro Mirror configuration limits*

| Parameter | Value |
|---|---|
| Number of Metro Mirror or Global Mirror Consistency Groups per system | 256 |
| Number of Metro Mirror or Global Mirror relationships per system | 8192 |
| Number of Metro Mirror or Global Mirror relationships per Consistency Group | 8192 |
| Total volume size per I/O Group | There is a per I/O Group limit of 1024 terabyte (TB) on the quantity of master and auxiliary volume address spaces that can participate in Metro Mirror and Global Mirror relationships. This maximum configuration uses all 512 MiB of bitmap space for the I/O Group and allows 10 MiB of space for all remaining copy services features. |

## 11.9.20 Remote Copy states and events

In this section, we describe the various states of a MM/GM relationship and the conditions that cause them to change.

In Figure 11-109, the MM/GM relationship diagram shows an overview of the status that can apply to a MM/GM relationship in a connected state.

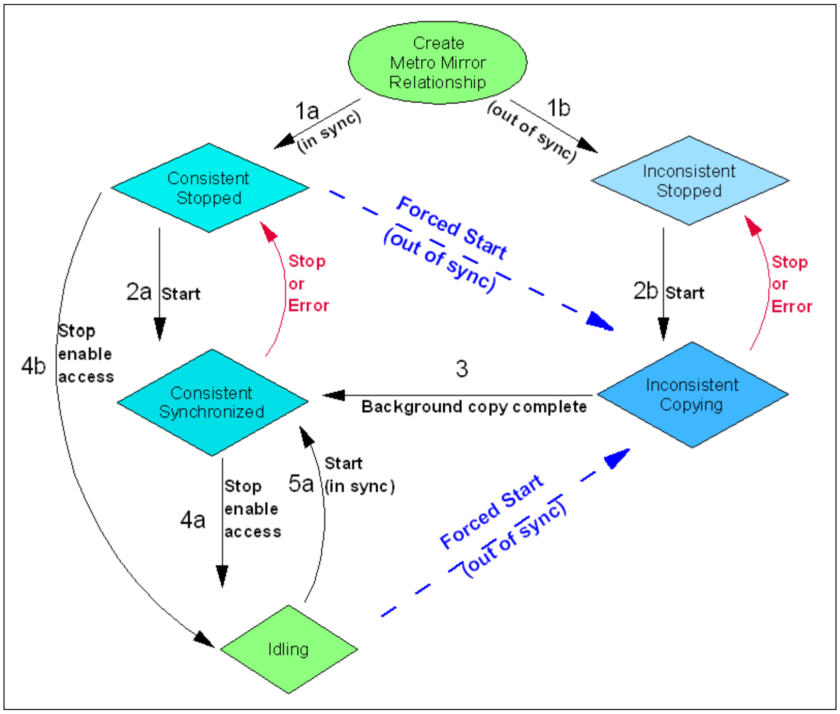


*Figure 11-109 Metro Mirror or Global Mirror mapping state diagram*

When the MM/GM relationship is created, you can specify whether the auxiliary volume is already in sync with the master volume, and the background copy process is then skipped. This capability is useful when MM/GM relationships are established for volumes that were created with the format option.

The following step identifiers are shown in Figure 11-109 on page 520:

- Step 1:
  a. The MM/GM relationship is created with the **-sync** option, and the MM/GM relationship enters the `ConsistentStopped` state.
  b. The MM/GM relationship is created without specifying that the master and auxiliary volumes are in sync, and the MM/GM relationship enters the `InconsistentStopped` state.
- Step 2:
  a. When a MM/GM relationship is started in the `ConsistentStopped` state, the MM/GM relationship enters the `ConsistentSynchronized` state. Therefore, no updates (write I/O) were performed on the master volume while in the `ConsistentStopped` state. Otherwise, the **-force** option must be specified, and the MM/GM relationship then enters the `InconsistentCopying` state while the background copy is started.

   b. When a MM/GM relationship is started in the `InconsistentStopped` state, the MM/GM relationship enters the `InconsistentCopying` state while the background copy is started.

- Step 3:

  When the background copy completes, the MM/GM relationship transitions from the `InconsistentCopying` state to the `ConsistentSynchronized` state.

- Step 4:

  a. When a MM/GM relationship is stopped in the `ConsistentSynchronized` state, the MM/GM relationship enters the `Idling` state when you specify the **`-access`** option, which enables write I/O on the auxiliary volume.

  b. When a MM/GM relationship is stopped in the `ConsistentSynchronized` state without an **`-access`** parameter, the auxiliary volumes remain read-only and the state of the relationship changes to `ConsistentStopped`.

  c. To enable write I/O on the auxiliary volume, when the MM/GM relationship is in the `ConsistentStopped` state, issue the **`svctask stoprcrelationship`** command, which specifies the **`-access`** option, and the MM/GM relationship enters the `Idling` state.

- Step 5:

  a. When a MM/GM relationship is started from the `Idling` state, you must specify the **`-primary`** argument to set the copy direction. If no write I/O was performed (to the master or auxiliary volume) while in the `Idling` state, the MM/GM relationship enters the `ConsistentSynchronized` state.

  b. If write I/O was performed to the master or auxiliary volume, the **`-force`** option must be specified and the MM/GM relationship then enters the `InconsistentCopying` state while the background copy is started. The background process copies only the data that changed on the primary volume while the relationship was stopped.

## Stop on Error

When a MM/GM relationship is stopped (intentionally, or because of an error), the state changes. For example, the MM/GM relationships in the `ConsistentSynchronized` state enter the `ConsistentStopped` state, and the MM/GM relationships in the `InconsistentCopying` state enter the `InconsistentStopped` state.

If the connection is broken between the two systems that are in a partnership, all (intercluster) MM/GM relationships enter a `Disconnected` state. For more information, see “Connected versus disconnected” on page 521.

> **Common states:** Stand-alone relationships and Consistency Groups share a common configuration and state model. All MM/GM relationships in a Consistency Group have the same state as the Consistency Group.

## State overview

In the following sections, we provide an overview of the various MM/GM states.

### *Connected versus disconnected*

Under certain error scenarios (for example, a power failure at one site that causes one complete system to disappear), communications between two systems in an MM/GM relationship can be lost. Alternatively, the fabric connection between the two systems might fail, which leaves the two systems running but they cannot communicate with each other.

When the two systems can communicate, the systems and the relationships that spans them are described as *connected*. When they cannot communicate, the systems and the relationships spanning them are described as *disconnected*.

In this state, both systems are left with fragmented relationships and are limited regarding the configuration commands that can be performed. The disconnected relationships are portrayed as having a changed state. The new states describe what is known about the relationship and the configuration commands that are permitted.

When the systems can communicate again, the relationships are reconnected. MM/GM automatically reconciles the two state fragments, considering any configuration or other event that occurred while the relationship was disconnected. As a result, the relationship can return to the state that it was in when it became disconnected, or it can enter a new state.

Relationships that are configured between volumes in the same IBM Storwize V7000 system (intracluster) are never described as being in a disconnected state.

### *Consistent versus inconsistent*

Relationships that contain volumes that are operating as secondaries can be described as being consistent or inconsistent. Consistency Groups that contain relationships can also be described as being consistent or inconsistent. The consistent or inconsistent property describes the relationship of the data on the auxiliary to the data on the master volume. It can be considered a property of the auxiliary volume.

An auxiliary volume is described as *consistent* if it contains data that might be read by a host system from the master if power failed at an imaginary point while I/O was in progress, and power was later restored. This imaginary point is defined as the *recovery point*.

The requirements for consistency are expressed regarding activity at the master up to the recovery point. The auxiliary volume contains the data from all of the writes to the master for which the host received successful completion and that data was not overwritten by a subsequent write (before the recovery point).

Consider writes for which the host did not receive a successful completion (that is, it received bad completion or no completion at all). If the host then performed a read from the master of that data that returned successful completion and no later write was sent (before the recovery point), the auxiliary contains the same data as the data that was returned by the read from the master.

From the point of view of an application, consistency means that an auxiliary volume contains the same data as the master volume at the recovery point (the time at which the imaginary power failure occurred). If an application is designed to cope with an unexpected power failure, this assurance of consistency means that the application can use the auxiliary and begin operation as though it was restarted after the hypothetical power failure. Again, maintaining the application write ordering is the key property of consistency.

For more information about dependent writes, see Consistency Groups.

If a relationship (or set of relationships) is inconsistent and an attempt is made to start an application by using the data in the secondaries, the following outcomes are possible:

- The application might decide that the data is corrupted and crash or exit with an event code.
- The application might fail to detect that the data is corrupted and return erroneous data.
- The application might work without a problem.

Because of the risk of data corruption, and in particular undetected data corruption, MM/GM strongly enforces the concept of consistency and prohibits access to inconsistent data.

Consistency as a concept can be applied to a single relationship or a set of relationships in a Consistency Group. Write ordering is a concept that an application can maintain across several disks that are accessed through multiple systems. Therefore, consistency must operate across all of those disks.

When you are deciding how to use Consistency Groups, the administrator must consider the scope of an application's data and consider all of the interdependent systems that communicate and exchange information.

If two programs or systems communicate and store details as a result of the information exchanged, either of the following actions might occur:

- All of the data that is accessed by the group of systems must be placed into a single Consistency Group.
- The systems must be recovered independently (each within its own Consistency Group). Then, each system must perform recovery with the other applications to become consistent with them.

### *Consistent versus synchronized*

A copy that is consistent and up-to-date is described as *synchronized*. In a synchronized relationship, the master and auxiliary volumes differ only in regions where writes are outstanding from the host.

Consistency does not mean that the data is up-to-date. A copy can be consistent and yet contain data that was frozen at a point in the past. Write I/O might continue to a master but not be copied to the auxiliary. This state arises when it becomes impossible to keep data up-to-date and maintain consistency. An example is a loss of communication between systems when you are writing to the auxiliary.

When communication is lost for an extended period, MM/GM tracks the changes that occurred on the master, but not the order or the details of such changes (write data). When communication is restored, it is impossible to synchronize the auxiliary without sending write data to the auxiliary out of order. Therefore, consistency is lost.

The following policies can be used to cope with this situation:

- Make a point-in-time copy of the consistent auxiliary before you allow the auxiliary to become inconsistent. If there is a disaster before consistency is achieved again, the point-in-time copy target provides a consistent (although out-of-date) image.
- Accept the loss of consistency and the loss of a useful auxiliary while synchronizing the auxiliary.

## Detailed states

In the following sections, we describe the states that are portrayed to the user for either Consistency Groups or relationships. We also describe information that is available in each state. The major states are designed to provide guidance about the available configuration commands.

### *InconsistentStopped*

`InconsistentStopped` is a connected state. In this state, the master is accessible for read and write I/O, but the auxiliary is not accessible for read or write I/O. A copy process must be started to make the auxiliary consistent. This state is entered when the relationship or

Consistency Group was `InconsistentCopying` and suffered a persistent error or received a **`stop`** command that caused the copy process to stop.

A **`start`** command causes the relationship or Consistency Group to move to the `InconsistentCopying` state. A **`stop`** command is accepted, but has no effect.

If the relationship or Consistency Group becomes disconnected, the auxiliary side transitions to `InconsistentDisconnected`. The master side transitions to `IdlingDisconnected`.

### *InconsistentCopying*

`InconsistentCopying` is a connected state. In this state, the master is accessible for read and write I/O, but the auxiliary is not accessible for read or write I/O. This state is entered after a **`start`** command is issued to an `InconsistentStopped` relationship or a Consistency Group.

It is also entered when a forced start is issued to an `Idling` or `ConsistentStopped` relationship or Consistency Group. In this state, a background copy process runs that copies data from the master to the auxiliary volume.

In the absence of errors, an `InconsistentCopying` relationship is active, and the copy progress increases until the copy process completes. In certain error situations, the copy progress might freeze or even regress.

A persistent error or **`stop`** command places the relationship or Consistency Group into an `InconsistentStopped` state. A **`start`** command is accepted but has no effect.

If the background copy process completes on a stand-alone relationship or on all relationships for a Consistency Group, the relationship or Consistency Group transitions to the `ConsistentSynchronized` state.

If the relationship or Consistency Group becomes disconnected, the auxiliary side transitions to `InconsistentDisconnected`. The master side transitions to `IdlingDisconnected`.

### *ConsistentStopped*

`ConsistentStopped` is a connected state. In this state, the auxiliary contains a consistent image, but it might be out-of-date in relation to the master. This state can arise when a relationship was in a `ConsistentSynchronized` state and experienced an error that forces a Consistency Freeze. It can also arise when a relationship is created with a `CreateConsistentFlag` set to `TRUE`.

Normally, write activity that follows an I/O error causes updates to the master, and the auxiliary is no longer synchronized. In this case, consistency must be given up for a period to reestablish synchronization. You must use a **`start`** command with the **`-force`** option to acknowledge this condition, and the relationship or Consistency Group transitions to `InconsistentCopying`. Enter this command only after all outstanding events are repaired.

In the unusual case where the master and the auxiliary are still synchronized (perhaps following a user stop, and no further write I/O was received), a **`start`** command takes the relationship to `ConsistentSynchronized`. No **`-force`** option is required. Also, in this case, you can enter a **`switch`** command that moves the relationship or Consistency Group to `ConsistentSynchronized` and reverses the roles of the master and the auxiliary.

If the relationship or Consistency Group becomes disconnected, the auxiliary transitions to `ConsistentDisconnected`. The master transitions to `IdlingDisconnected`.

An informational status log is generated whenever a relationship or Consistency Group enters the ConsistentStopped state with a status of Online. You can configure this event to generate an SNMP trap that can be used to trigger automation or manual intervention to issue a **start** command following a loss of synchronization.

### *ConsistentSynchronized*

ConsistentSynchronized is a connected state. In this state, the master volume is accessible for read and write I/O, and the auxiliary volume is accessible for read-only I/O. Writes that are sent to the master volume are also sent to the auxiliary volume. Either successful completion must be received for both writes, the write must be failed to the host, or a state must transition out of the ConsistentSynchronized state before a write is completed to the host.

A **stop** command takes the relationship to the ConsistentStopped state. A **stop** command with the **-access** parameter takes the relationship to the Idling state.

A **switch** command leaves the relationship in the ConsistentSynchronized state, but it reverses the master and auxiliary roles (it switches the direction of replicating data). A **start** command is accepted, but has no effect.

If the relationship or Consistency Group becomes disconnected, the same transitions are made as for ConsistentStopped.

### *Idling*

Idling is a connected state. Both master and auxiliary volumes operate in the master role. Therefore, both master and auxiliary volumes are accessible for write I/O.

In this state, the relationship or Consistency Group accepts a **start** command. MM/GM maintains a record of regions on each disk that received write I/O while they were idling. This record is used to determine what areas must be copied following a **start** command.

The **start** command must specify the new copy direction. A **start** command can cause a loss of consistency if either volume in any relationship received write I/O, which is indicated by the Synchronized status. If the **start** command leads to loss of consistency, you must specify the **-force** parameter.

Following a **start** command, the relationship or Consistency Group transitions to ConsistentSynchronized if there is no loss of consistency, or to InconsistentCopying if there is a loss of consistency.

Also, the relationship or Consistency Group accepts a **-clean** option on the **start** command while in this state. If the relationship or Consistency Group becomes disconnected, both sides change their state to IdlingDisconnected.

### *IdlingDisconnected*

IdlingDisconnected is a disconnected state. The target volumes in this half of the relationship or Consistency Group are all in the master role and accept read or write I/O.

The priority in this state is to recover the link to restore the relationship or consistency.

No configuration activity is possible (except for deletes or stops) until the relationship becomes connected again. At that point, the relationship transitions to a connected state. The exact connected state that is entered depends on the state of the other half of the relationship or Consistency Group, which depends on the following factors:

- The state when it became disconnected
- The write activity since it was disconnected
- The configuration activity since it was disconnected

If both halves are `IdlingDisconnected`, the relationship becomes `Idling` when it is reconnected.

While `IdlingDisconnected`, if a write I/O is received that causes the loss of synchronization (synchronized attribute transitions from `true` to `false`) and the relationship was not already stopped (either through a user stop or a persistent error), an event is raised to notify you of the condition. This same event also is raised when this condition occurs for the `ConsistentSynchronized` state.

### *InconsistentDisconnected*

`InconsistentDisconnected` is a disconnected state. The target volumes in this half of the relationship or Consistency Group are all in the auxiliary role, and do not accept read *or* write I/O. Except for deletes, no configuration activity is permitted until the relationship becomes connected again.

When the relationship or Consistency Group becomes connected again, the relationship becomes `InconsistentCopying` automatically unless either of the following conditions are true:

- The relationship was `InconsistentStopped` when it became disconnected.
- The user issued a **`stop`** command while disconnected.

In either case, the relationship or Consistency Group becomes `InconsistentStopped`.

### *ConsistentDisconnected*

`ConsistentDisconnected` is a disconnected state. The target volumes in this half of the relationship or Consistency Group are all in the auxiliary role, and accept read I/O but *not* write I/O.

This state is entered from `ConsistentSynchronized` or `ConsistentStopped` when the auxiliary side of a relationship becomes disconnected.

In this state, the relationship or Consistency Group displays an attribute of FreezeTime, which is the point when Consistency was frozen. When it is entered from `ConsistentStopped`, it retains the time that it had in that state. When it is entered from `ConsistentSynchronized`, the FreezeTime shows the last time at which the relationship or Consistency Group was known to be consistent. This time corresponds to the time of the last successful heartbeat to the other system.

A **`stop`** command with the **`-access`** flag set to `true` transitions the relationship or Consistency Group to the `IdlingDisconnected` state. This state allows write I/O to be performed to the auxiliary volume and is used as part of a DR scenario.

When the relationship or Consistency Group becomes connected again, the relationship or Consistency Group becomes `ConsistentSynchronized` only if this action does not lead to a loss of consistency. The following conditions must be true:

- The relationship was `ConsistentSynchronized` when it became disconnected.
- No writes received successful completion at the master while disconnected.

Otherwise, the relationship becomes `ConsistentStopped`. The FreezeTime setting is retained.

### *Empty*

This state applies only to Consistency Groups. It is the state of a Consistency Group that has no relationships and no other state information to show.

It is entered when a Consistency Group is first created. It is exited when the first relationship is added to the Consistency Group, at which point the state of the relationship becomes the state of the Consistency Group.

# 11.10 Remote Copy commands

In this section, we present commands that need to be issued to create and operate remote copy services.

## 11.10.1 Remote Copy process

The MM/GM process includes the following steps:

1. A system partnership is created between two IBM Storwize V7000 systems or IBM SAN Volume Controller (for intercluster MM/GM).
2. A MM/GM relationship is created between two volumes of the same size.
3. To manage multiple MM/GM relationships as one entity, the relationships can be made part of a MM/GM Consistency Group to ensure data consistency across multiple MM/GM relationships, or for ease of management.
4. The MM/GM relationship is started and when the background copy completes, the relationship is consistent and synchronized.
5. When synchronized, the auxiliary volume holds a copy of the production data at the master that can be used for disaster recovery.
6. To access the auxiliary volume, the MM/GM relationship must be stopped with the access option enabled before write I/O is submitted to the auxiliary.

Following these commands, the remote host server is mapped to the auxiliary volume and the disk is available for I/O.

For more information about MM/GM commands, see *IBM System Storage SAN Volume Controller and IBM Storwize V7000 Command-Line Interface User's Guide, GC27-2287.*

The command set for MM/GM contains the following broad groups:

- Commands to create, delete, and manipulate relationships and Consistency Groups
- Commands to cause state changes

If a configuration command affects more than one system, MM/GM performs the work to coordinate configuration activity between the systems. Certain configuration commands can be performed only when the systems are connected, and fail with no effect when they are disconnected.

Other configuration commands are permitted even though the systems are disconnected. The state is reconciled automatically by MM/GM when the systems become connected again.

For any command (with one exception) a single system receives the command from the administrator. This design is significant for defining the context for a CreateRelationship **mkrcrelationship** or CreateConsistencyGroup **mkrcconsistgrp** command, in which case the system that is receiving the command is called the *local system*.

The exception is a command that sets systems into a MM/GM partnership. The **mkfcpartnership** and **mkippartnership** commands must be issued on both, the local and remote systems.

The commands in this section are described as an abstract command set, and are implemented by either of the following methods:

- CLI can be used for scripting and automation.
- GUI can be used for one-off tasks.

## 11.10.2 Listing available system partners

Use the **lspartnershipcandidate** command to list the systems that are available for setting up a two-system partnership. This command is a prerequisite for creating MM/GM relationships.

> **Note:** This command is not supported on IP partnerships. Use **mkippartnership** for IP connections.

## 11.10.3 Changing the system parameters

When you want to change system parameters specific to any remote copy or Global Mirror only, use the **chsystem** command.

### The chsystem command

The **chsystem** command features the following parameters for MM/GM:

- **-relationshipbandwidthlimit** ***cluster_relationship_bandwidth_limit***

  This parameter controls the maximum rate at which any one remote copy relationship can synchronize. The default value for the relationship bandwidth limit is 25 MBps, but this value can now be specified 1 - 100,000 MBps. The partnership overall limit is controlled by the **chpartnership -linkbandwidthmbits** command, and must be set on each involved system.

  > **Important:** Do not set this value higher than the default without first establishing that the higher bandwidth can be sustained without affecting the host's performance. The limit must never be higher than the maximum that is supported by the infrastructure connecting the remote sites, regardless of the compression rates that you might achieve.

- **-gmlinktolerance** ***link_tolerance***

  This parameter specifies the maximum period that the system tolerates delay before stopping Global Mirror relationships. Specify values 60 - 86,400 seconds in increments of 10 seconds. The default value is 300. Do not change this value except under the direction of IBM Support.

- **-gmmaxhostdelay** ***max_host_delay***

  This parameter specifies the maximum time delay, in milliseconds, at which the Global Mirror link tolerance timer starts counting down. This threshold value determines the additional effect that Global Mirror operations can add to the response times of the Global Mirror source volumes. You can use this parameter to increase the threshold from the default value of 5 milliseconds.

- **-gminterdelaysimulation** ***link_tolerance***

  This parameter specifies the number of milliseconds that I/O activity (intercluster copying to an auxiliary volume) is delayed. This parameter enables you to test performance implications before Global Mirror is deployed and a long-distance link is obtained. Specify a value of 0 - 100 milliseconds in 1-millisecond increments. The default value is 0. Use this argument to test each intercluster Global Mirror relationship separately.

- **-gmintradelaysimulation** ***link_tolerance***

  This parameter specifies the number of milliseconds that I/O activity (intracluster copying to an auxiliary volume) is delayed. By using this parameter, you can test performance implications before Global Mirror is deployed and a long-distance link is obtained. Specify a value of 0 - 100 milliseconds in 1-millisecond increments. The default value is 0. Use this argument to test each intracluster Global Mirror relationship separately.

- **-maxreplicationdelay** ***max_replication_delay***

  This parameter sets a maximum replication delay in seconds. The value must be a number 1 - 360. This feature sets the maximum number of seconds to be tolerated to complete a single I/O. If I/O can't complete within the ***max_replication_delay*** the 1920 event is reported. This is the system-wide setting. When set to 0, the feature is disabled. This applies to Metro Mirror and Global Mirror relationships.

Use the **chsystem** command to adjust these values, as shown in the following example:

```
chsystem -gmlinktolerance 300
```

You can view all of these parameter values by using the **lssystem** ***<system_name>*** command.

### *gmlinktolerance*

We focus on the **gmlinktolerance** parameter in particular. If poor response extends past the specified tolerance, a 1920 event is logged and one or more GM relationships automatically stop to protect the application hosts at the primary site. During normal operations, application hosts experience a minimal effect from the response times because the GM feature uses asynchronous replication.

However, if GM operations experience degraded response times from the secondary system for an extended period, I/O operations begin to queue at the primary system. This queue results in an extended response time to application hosts. In this situation, the **gmlinktolerance** feature stops GM relationships, and the application host's response time returns to normal.

After a 1920 event occurs, the GM auxiliary volumes are no longer in the consistent_synchronized state until you fix the cause of the event and restart your GM relationships. For this reason, ensure that you monitor the system to track when these 1920 events occur.

You can disable the **gmlinktolerance** feature by setting the **gmlinktolerance** value to 0 (zero). However, the **gmlinktolerance** feature cannot protect applications from extended response times if it is disabled. It might be appropriate to disable the **gmlinktolerance** feature under the following circumstances:

- During SAN maintenance windows in which degraded performance is expected from SAN components, and application hosts can withstand extended response times from GM volumes.
- During periods when application hosts can tolerate extended response times and it is expected that the **gmlinktolerance** feature might stop the GM relationships. For example, if you test by using an I/O generator that is configured to stress the back-end storage, the **gmlinktolerance** feature might detect the high latency and stop the GM relationships. Disabling the **gmlinktolerance** feature prevents this result at the risk of exposing the test host to extended response times.

A `1920` event indicates that one or more of the SAN components cannot provide the performance that is required by the application hosts. This situation can be temporary (for example, a result of a maintenance activity) or permanent (for example, a result of a hardware failure or an unexpected host I/O workload).

If `1920` events are occurring, it can be necessary to use a performance monitoring and analysis tool, such as the IBM Virtual Storage Center, to help identify and resolve the problem.

## 11.10.4 System partnership

To create an IBM SAN Volume Controller or an IBM Storwize V7000 system partnership, use the **`mkfcpartnership`** command for traditional Fibre Channel (FC or FCoE) connections or **`mkippartnership`** for IP-based connections.

### The svctask mkfcpartnership command

Use the **`mkfcpartnership`** command to establish a one-way MM/GM partnership between the local system and a remote system. Alternatively, use **`mkippartnership`** to create IP-based partnership.

To establish a fully functional MM/GM partnership, you must issue this command on both systems. This step is a prerequisite for creating MM/GM relationships between volumes on the IBM Spectrum Virtualize systems.

When the partnership is created, you can specify the bandwidth to be used by the background copy process between the local and remote system. If it is not specified, the bandwidth defaults to 50 MBps. The bandwidth must be set to a value that is less than or equal to the bandwidth that can be sustained by the intercluster link.

#### *Background copy bandwidth effect on foreground I/O latency*

The background copy bandwidth determines the rate at which the background copy is attempted for MM/GM. The background copy bandwidth can affect foreground I/O latency in one of the following ways:

- The following result can occur if the background copy bandwidth is set too high compared to the MM/GM intercluster link capacity:
  - The background copy I/Os can back up on the MM/GM intercluster link.
  - There is a delay in the synchronous auxiliary writes of foreground I/Os.
  - The foreground I/O latency increases as perceived by applications.
- If the background copy bandwidth is set too high for the storage at the primary site, background copy read I/Os overload the primary storage and delay foreground I/Os.
- If the background copy bandwidth is set too high for the storage at the secondary site, background copy writes at the secondary site overload the auxiliary storage and again delay the synchronous secondary writes of foreground I/Os.

To set the background copy bandwidth optimally, ensure that you consider all three resources: primary storage, intercluster link bandwidth, and auxiliary storage. Provision the most restrictive of these three resources between the background copy bandwidth and the peak foreground I/O workload.

Perform this provisioning by calculation or by determining experimentally how much background copy can be allowed before the foreground I/O latency becomes unacceptable. Then, reduce the background copy to accommodate peaks in workload.

### The chpartnership command

To change the bandwidth that is available for background copy in the system partnership, use the **chpartnership -backgroundcopyrate** ***<percentage_of_link_bandwidth>*** command to specify the percentage of whole link capacity to be used by background copy process.

## 11.10.5 Creating a Metro Mirror/Global Mirror consistency group

Use the **mkrcconsistgrp** command to create an empty MM/GM Consistency Group.

The MM/GM consistency group name must be unique across all consistency groups that are known to the systems owning this consistency group. If the consistency group involves two systems, the systems must be in communication throughout the creation process.

The new consistency group does not contain any relationships and is in the Empty state. You can add MM/GM relationships to the group (upon creation or afterward) by using the **chrelationship** command.

## 11.10.6 Creating a Metro Mirror/Global Mirror relationship

Use the **mkrcrelationship** command to create a new MM/GM relationship. This relationship persists until it is deleted.

> **Optional parameter:** If you do not use the **-global** optional parameter, a Metro Mirror relationship is created rather than a Global Mirror relationship.

The auxiliary volume must be equal in size to the master volume or the command fails. If both volumes are in the same system, they must be in the same I/O Group. The master and auxiliary volume cannot be in an existing relationship, and they cannot be the target of a FlashCopy mapping. This command returns the new relationship (relationship_id) when successful.

When the MM/GM relationship is created, you can add it to an existing Consistency Group, or it can be a stand-alone MM/GM relationship if no Consistency Group is specified.

### The lsrcrelationshipcandidate command

Use the **lsrcrelationshipcandidate** command to list the volumes that are eligible to form an MM/GM relationship.

When the command is issued, you can specify the master volume name and auxiliary system to list the candidates that comply with the prerequisites to create a MM/GM relationship. If the command is issued with no parameters, all of the volumes that are not disallowed by another configuration state, such as being a FlashCopy target, are listed.

## 11.10.7 Changing Metro Mirror/Global Mirror relationship

Use the **chrcrelationship** command to modify the following properties of an MM/GM relationship:

- Change the name of an MM/GM relationship.
- Add a relationship to a group.
- Remove a relationship from a group using the **-force** flag.

**Adding an MM/GM relationship:** When an MM/GM relationship is added to a Consistency Group that is not empty, the relationship must have the same state and copy direction as the group to be added to it.

## 11.10.8 Changing Metro Mirror/Global Mirror consistency group

Use the **chrcconsistgrp** command to change the name of an MM/GM Consistency Group.

## 11.10.9 Starting Metro Mirror/Global Mirror relationship

Use the **startrcrelationship** command to start the copy process of an MM/GM relationship.

When the command is issued, you can set the copy direction if it is undefined, and, optionally, you can mark the auxiliary volume of the relationship as clean. The command fails if it is used as an attempt to start a relationship that is already a part of a consistency group.

You can issue this command only to a relationship that is connected. For a relationship that is idling, this command assigns a copy direction (master and auxiliary roles) and begins the copy process. Otherwise, this command restarts a previous copy process that was stopped by a **stop** command or by an I/O error.

If the resumption of the copy process leads to a period when the relationship is inconsistent, you must specify the **-force** parameter when the relationship is restarted. This situation can arise if, for example, the relationship was stopped and then further writes were performed on the original master of the relationship.

The use of the **-force** parameter here is a reminder that the data on the auxiliary becomes inconsistent while resynchronization (background copying) takes place. Therefore, this data is unusable for DR purposes before the background copy completes.

In the `Idling` state, you must specify the master volume to indicate the copy direction. In other connected states, you can provide the **-primary** argument, but it must match the existing setting.

## 11.10.10 Stopping Metro Mirror/Global Mirror relationship

Use the **stoprcrelationship** command to stop the copy process for a relationship. You can also use this command to enable write access to a consistent auxiliary volume by specifying the **-access** parameter.

This command applies to a stand-alone relationship. It is rejected if it is addressed to a relationship that is part of a Consistency Group. You can issue this command to stop a relationship that is copying from master to auxiliary.

If the relationship is in an inconsistent state, any copy operation stops and does not resume until you issue a **startrcrelationship** command. Write activity is no longer copied from the master to the auxiliary volume. For a relationship in the `ConsistentSynchronized` state, this command causes a Consistency Freeze.

When a relationship is in a consistent state (that is, in the `ConsistentStopped`, `ConsistentSynchronized`, or `ConsistentDisconnected` state), you can use the **-access** parameter with the **stoprcrelationship** command to enable write access to the auxiliary volume.

### 11.10.11 Starting Metro Mirror/Global Mirror consistency group

Use the **startrcconsistgrp** command to start an MM/GM consistency group. You can issue this command only to a consistency group that is connected.

For a consistency group that is idling, this command assigns a copy direction (master and auxiliary roles) and begins the copy process. Otherwise, this command restarts a previous copy process that was stopped by a **stop** command or by an I/O error.

### 11.10.12 Stopping Metro Mirror/Global Mirror consistency group

Use the **startrcconsistgrp** command to stop the copy process for an MM/GM consistency group. You can also use this command to enable write access to the auxiliary volumes in the group if the group is in a consistent state.

If the consistency group is in an inconsistent state, any copy operation stops and does not resume until you issue the **startrcconsistgrp** command. Write activity is no longer copied from the master to the auxiliary volumes that belong to the relationships in the group. For a consistency group in the `ConsistentSynchronized` state, this command causes a Consistency Freeze.

When a consistency group is in a consistent state (for example, in the `ConsistentStopped`, `ConsistentSynchronized`, or `ConsistentDisconnected` state), you can use the **-access** parameter with the **stoprcconsistgrp** command to enable write access to the auxiliary volumes within that group.

### 11.10.13 Deleting Metro Mirror/Global Mirror relationship

Use the **rmrcrelationship** command to delete the relationship that is specified. Deleting a relationship deletes only the logical relationship between the two volumes. It does not affect the volumes themselves.

If the relationship is disconnected at the time that the command is issued, the relationship is deleted only on the system on which the command is being run. When the systems reconnect, the relationship is automatically deleted on the other system.

Alternatively, if the systems are disconnected and you still want to remove the relationship on both systems, you can issue the **rmrcrelationship** command independently on both of the systems.

A relationship cannot be deleted if it is part of a consistency group. You must first remove the relationship from the consistency group.

If you delete an inconsistent relationship, the auxiliary volume becomes accessible even though it is still inconsistent. This situation is the one case in which MM/GM does not inhibit access to inconsistent data.

### 11.10.14 Deleting Metro Mirror/Global Mirror consistency group

Use the **rmrcconsistgrp** command to delete an MM/GM consistency group. This command deletes the specified consistency group. You can issue this command for any existing consistency group.

If the consistency group is disconnected at the time that the command is issued, the consistency group is deleted only on the system on which the command is being run. When the systems reconnect, the consistency group is automatically deleted on the other system.

Alternatively, if the systems are disconnected and you still want to remove the consistency group on both systems, you can issue the `rmrcconsistgrp` command separately on both of the systems.

If the consistency group is not empty, the relationships within it are removed from the consistency group before the group is deleted. These relationships then become stand-alone relationships. The state of these relationships is not changed by the action of removing them from the consistency group.

## 11.10.15 Reversing Metro Mirror/Global Mirror relationship

Use the `switchrcrelationship` command to reverse the roles of the master volume and the auxiliary volume when a stand-alone relationship is in a consistent state. When the command is issued, the wanted master must be specified.

## 11.10.16 Reversing Metro Mirror/Global Mirror consistency group

Use the `switchrcconsistgrp` command to reverse the roles of the master volume and the auxiliary volume when a consistency group is in a consistent state. This change is applied to all of the relationships in the consistency group. When the command is issued, the wanted master must be specified.

**Important:** Remember, by reversing the roles your current source volumes become targets, and target volumes become source volumes. Therefore, you lose write access to your current primary volumes.

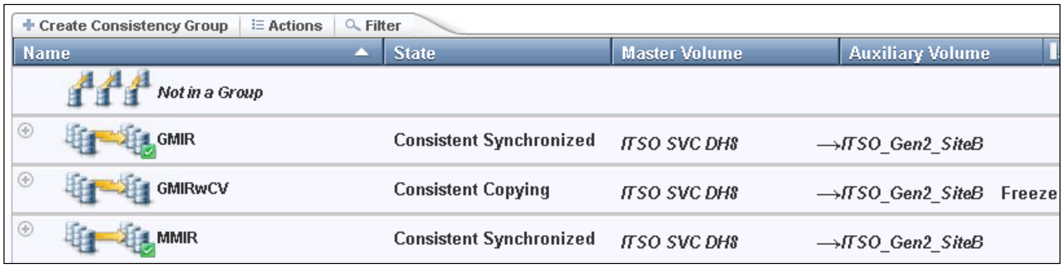
# 11.11 Managing Remote Copy using the GUI

It is often easier to control working with Metro Mirror or Global Mirror by using the GUI, if you have few mappings. When many mappings are used, run your commands by using the CLI. In this section, we describe the tasks that you can perform at a remote copy level.

The following panes are used to visualize and manage your remote copies:

- The Remote Copy pane, as shown in Figure 11-110.

  To access the Remote Copy pane, move the mouse pointer over the Copy Services selection and click **Remote Copy**.

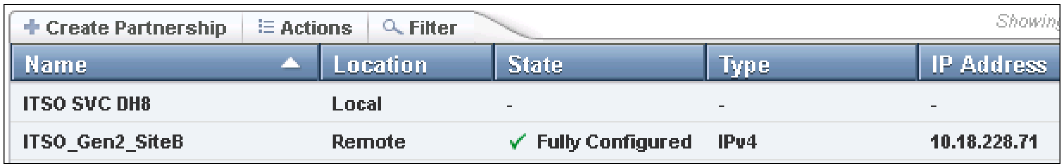


| Name | State | Master Volume | Auxiliary Volume | |
|---|---|---|---|---|
| *Not in a Group* | | | | |
| GMIR | Consistent Synchronized | ITSO SVC DH8 | →ITSO_Gen2_SiteB | |
| GMIRwCV | Consistent Copying | ITSO SVC DH8 | →ITSO_Gen2_SiteB | Freeze |
| MMIR | Consistent Synchronized | ITSO SVC DH8 | →ITSO_Gen2_SiteB | |

*Figure 11-110 Remote Copy pane*

- The Partnerships pane, as shown in Figure 11-111.

  To access the Partnerships pane, move the mouse pointer over the Copy Services selection and click **Partnerships**.



| Name | Location | State | Type | IP Address |
|---|---|---|---|---|
| ITSO SVC DH8 | Local | - | - | - |
| ITSO_Gen2_SiteB | Remote | ✓ Fully Configured | IPv4 | 10.18.228.71 |

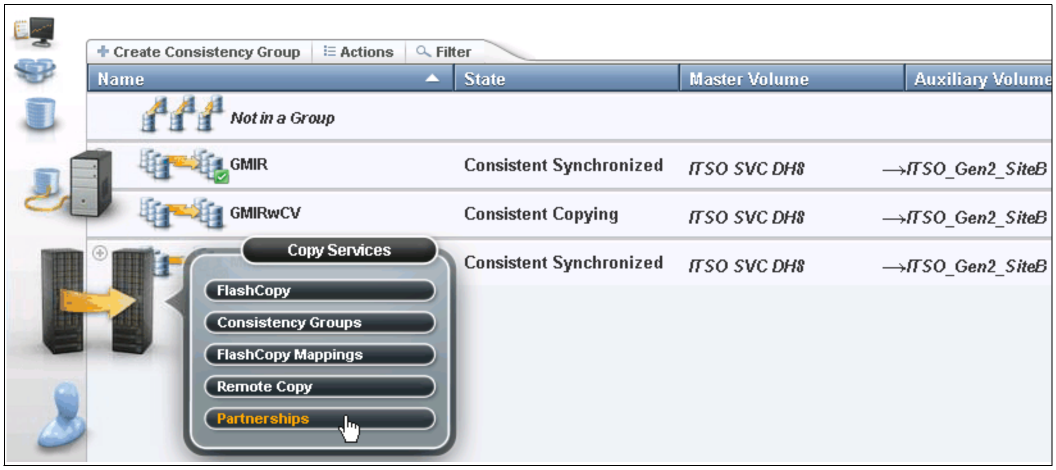
*Figure 11-111 Partnerships pane*

## 11.11.1 Creating Fibre Channel partnership

We perform this operation to create the partnership on both systems by using FC.

To create an FC partnership between the systems running IBM Spectrum Virtualize, use the GUI and complete the following steps:

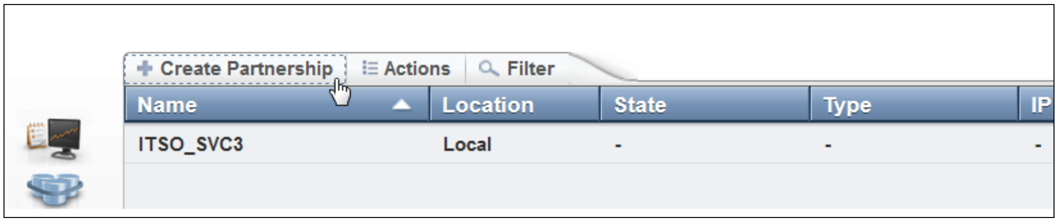
1. From the System pane, move the mouse pointer over Copy Services in the dynamic menu and click **Partnerships**. The Partnerships pane opens, as shown in Figure 11-112.



| Name | State | Master Volume | Auxiliary Volume |
|---|---|---|---|
| Not in a Group | | | |
| GMIR | Consistent Synchronized | ITSO SVC DH8 | →ITSO_Gen2_SiteB |
| GMIRwCV | Consistent Copying | ITSO SVC DH8 | →ITSO_Gen2_SiteB |
| | Consistent Synchronized | ITSO SVC DH8 | →ITSO_Gen2_SiteB |

*Figure 11-112 Selecting Partnerships window*

2. Click **Create Partnership** to create a partnership with another SVC or IBM Storwize system, as shown in Figure 11-113.



| Name | Location | State | Type | IP |
|---|---|---|---|---|
| ITSO_SVC3 | Local | - | - | - |

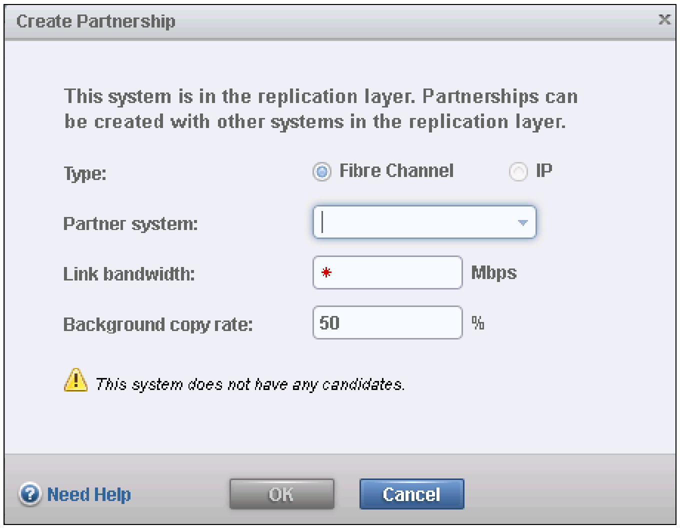
*Figure 11-113 Create a partnership*

3. In the Create Partnership window (Figure 11-114 on page 536), complete the following information:

- Select the partnership type, either Fibre Channel or IP. If you choose IP partnership, you must provide the IP address of the partner system and the partner system's CHAP key.
- If your partnership if based on Fibre Channel protocol, select an available partner system from the drop-down list. If no candidate is available, the following error message is displayed:

  ```
  This system does not have any candidates.
  ```

- Enter a link bandwidth in megabits per second (Mbps) that is used by the background copy process between the systems in the partnership.
- Enter the background copy rate.
- Click **OK** to confirm the partnership relationship, as shown in Figure 11-114 on page 536.



This system is in the replication layer. Partnerships can be created with other systems in the replication layer.

Type: ● Fibre Channel ○ IP

Partner system:

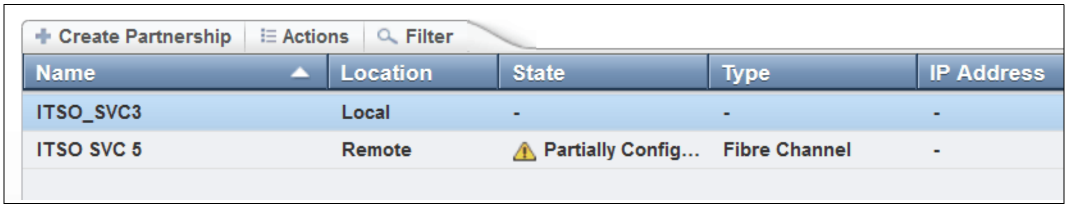
Link bandwidth: * Mbps

Background copy rate: 50 %

This system does not have any candidates.

*Figure 11-114 Select the type of partnership*

4. As shown in Figure 11-115, our partnership is in the Partially Configured state because this work was performed only on one side of the partnership so far.



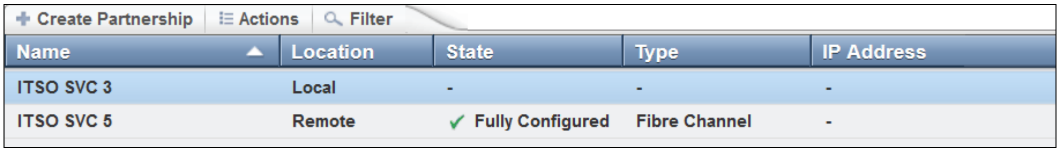
| Name | Location | State | Type | IP Address |
|---|---|---|---|---|
| ITSO_SVC3 | Local | - | - | - |
| ITSO SVC 5 | Remote | Partially Config... | Fibre Channel | - |

*Figure 11-115 Viewing system partnerships*

To fully configure the partnership between both systems, perform the same steps on the other system in the partnership. For simplicity and brevity, we show only the two most significant windows when the partnership is fully configured.

5. Starting the GUI at the partner system, select **ITSO SVC 3** for the system partnership. We specify the available bandwidth for the background copy (200 Mbps) and then click **OK**.

   Now that both sides of the system partnership are defined, the resulting windows are similar at the both systems as shown in Figure 11-116.



| Name | Location | State | Type | IP Address |
|---|---|---|---|---|
| ITSO SVC 3 | Local | - | - | - |
| ITSO SVC 5 | Remote | ✓ Fully Configured | Fibre Channel | - |

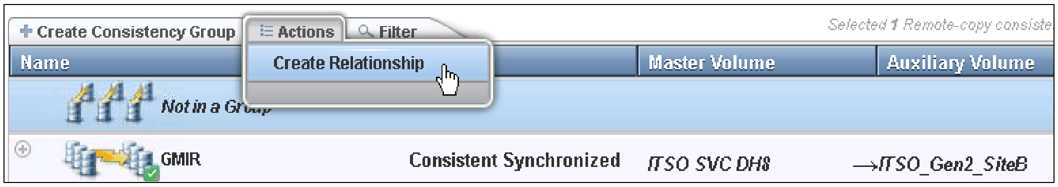
*Figure 11-116 System ITSO SVC 3: Fully configured remote partnership*

## 11.11.2 Creating stand-alone remote copy relationships

In this section, we create remote copy mappings for volumes with their respective remote targets. The source and target volumes were created before this operation was done on both systems. The target volume must have the same size as the source volume.

Complete the following steps to create stand-alone copy relationships:

1. From the System pane, select **Copy Services** → **Remote Copy** → **Actions**.
2. Click **Create Relationship**, as shown in Figure 11-117.

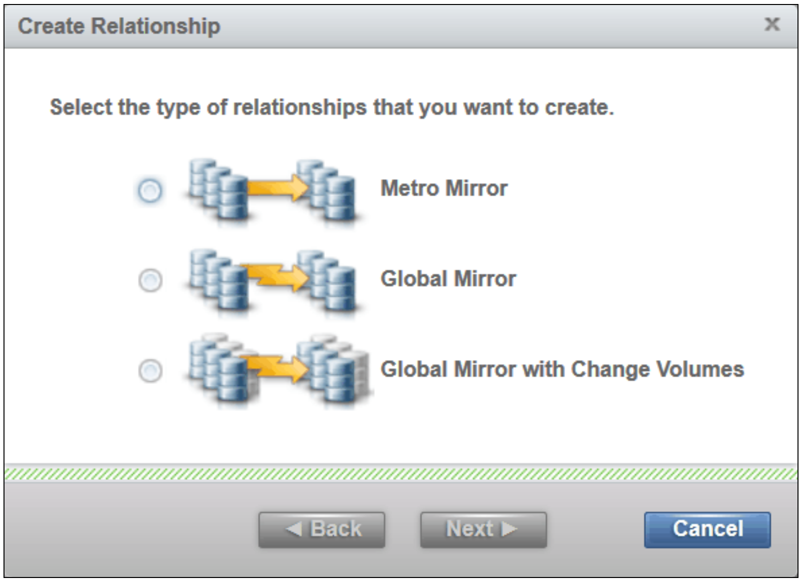


| Name | | Master Volume | Auxiliary Volume |
|---|---|---|---|
| Not in a Group | | | |
| GMIR | Consistent Synchronized | ITSO SVC DH8 | →ITSO_Gen2_SiteB |

*Figure 11-117 Create Relationship action*

3. In the Create Relationship window, select one of the following types of relationships that you want to create (as shown in Figure 11-118 on page 537):
   - – Metro Mirror
   - – Global Mirror
   - – Global Mirror with Change Volumes

   Select the relationship type, as shown in Figure 11-118.


Create Relationship

Select the type of relationships that you want to create.

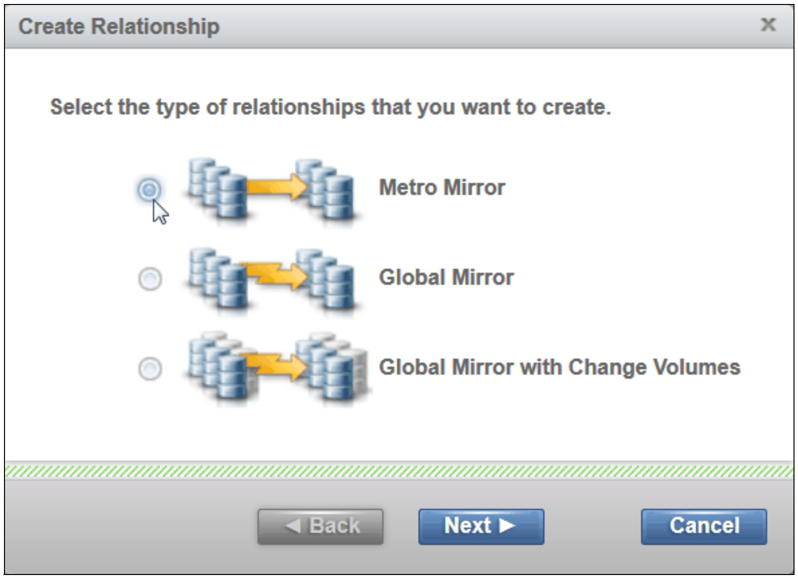
○ Metro Mirror

○ Global Mirror

○ Global Mirror with Change Volumes

*Figure 11-118 Select the type of relationship that you want to create*

4. We want to create a Metro Mirror relationship. See Figure 11-119.


Create Relationship

Select the type of relationships that you want to create.

● Metro Mirror

○ Global Mirror

○ Global Mirror with Change Volumes

*Figure 11-119 Selecting Metro Mirror as the type of relationship*

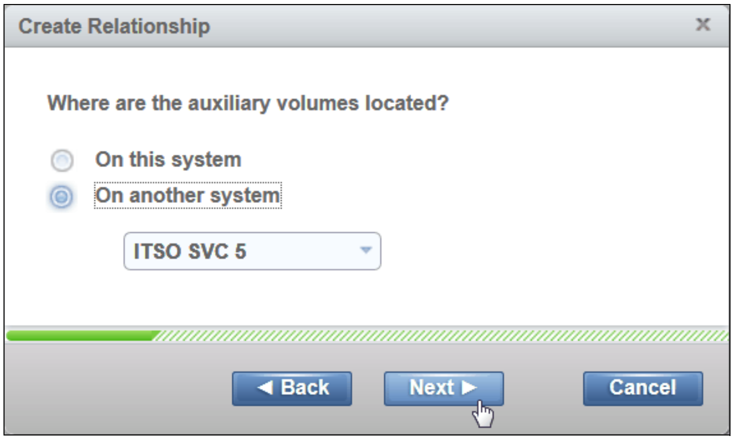
Click **Next**.

5. In the next window, select the location of the auxiliary volumes, as shown in Figure 11-120 on page 538:

   – On this system, which means that the volumes are local.

   – On another system, which means that you select the remote system from the drop-down list.

   After you make a selection, click **Next**.


Create Relationship

Where are the auxiliary volumes located?

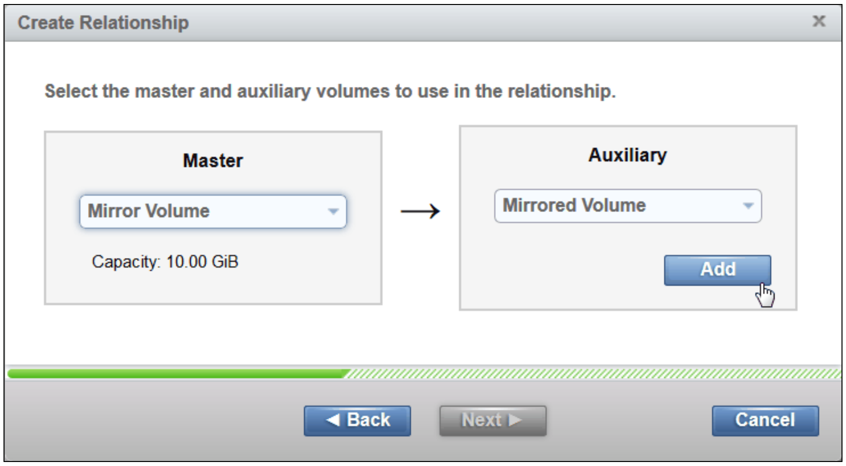
○ On this system

● On another system

ITSO SVC 5

*Figure 11-120 Specifying the location of the auxiliary volumes*

6. In the New Relationship window that is shown in Figure 11-121 on page 539, you can create relationships. Select a master volume in the Master drop-down list. Then, select an auxiliary volume in the Auxiliary drop-down list for this master and click **Add**. If needed, repeat this step to create other relationships.



Create Relationship

Select the master and auxiliary volumes to use in the relationship.

Master: Mirror Volume

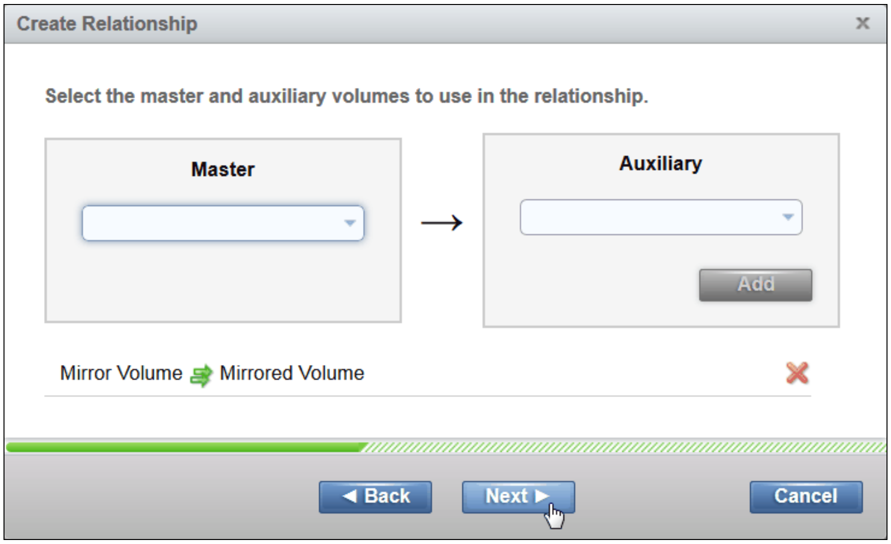
Capacity: 10.00 GiB

Auxiliary: Mirrored Volume

*Figure 11-121 Select a volume for mirroring*

> **Important:** The master and auxiliary volumes must be of equal size. Therefore, only the targets with the appropriate size are shown in the list for a specific source volume.

7. To remove a relationship that was created, click ✖, as shown in Figure 11-122.



Create Relationship

Select the master and auxiliary volumes to use in the relationship.

Master

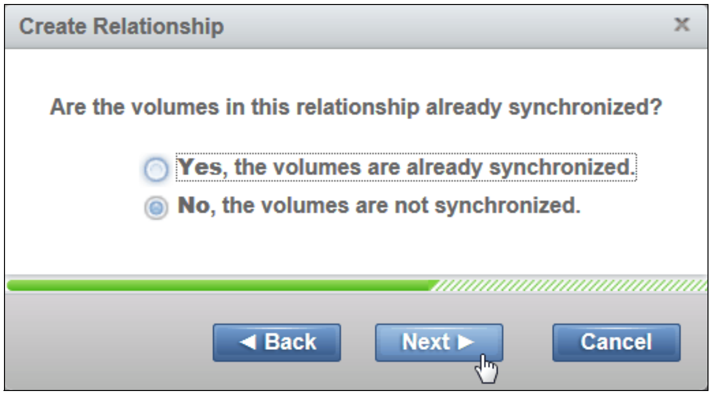
Auxiliary

Mirror Volume ➡ Mirrored Volume ✖

*Figure 11-122 Create the relationships between the master and auxiliary volumes*

After all of the relationships that you want to create are shown, click **Next**.

8. Specify whether the volumes are synchronized, as shown in Figure 11-123 on page 540. Then, click **Next**.

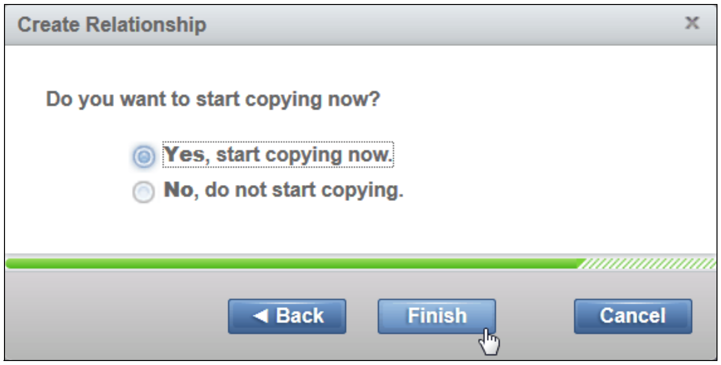


Are the volumes in this relationship already synchronized?

○ **Yes**, the volumes are already synchronized.
● **No**, the volumes are not synchronized.

*Figure 11-123 Volumes are already synchronized*

9. In the last window, select whether you want to start to copy the data, as shown in Figure 11-124. Click **Finish**.

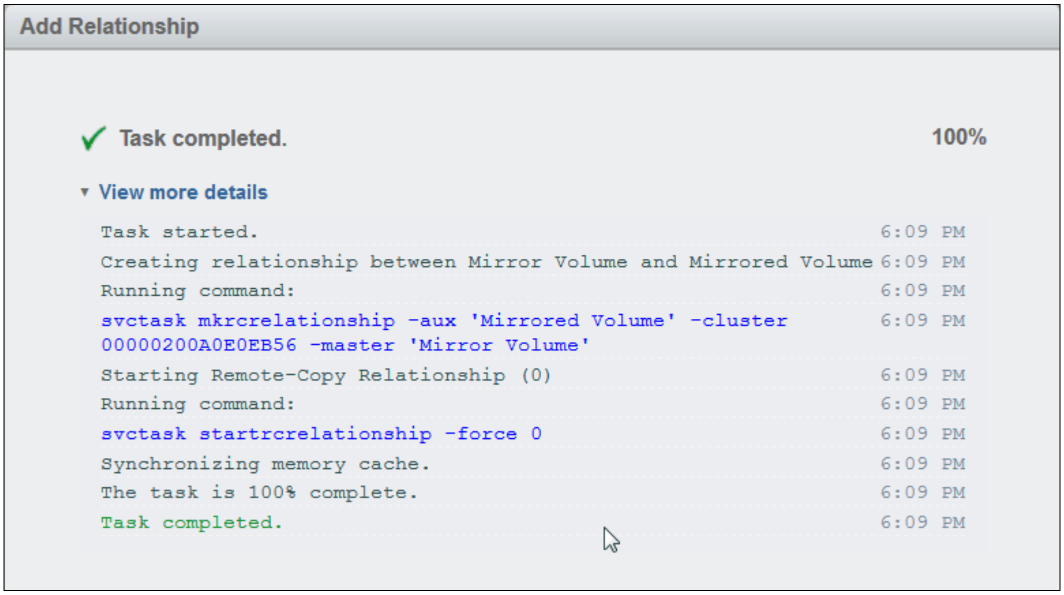


Do you want to start copying now?

● **Yes**, start copying now.
○ **No**, do not start copying.

*Figure 11-124 Synchronize now*

10. Figure 11-125 shows that the task to create the relationship is complete.



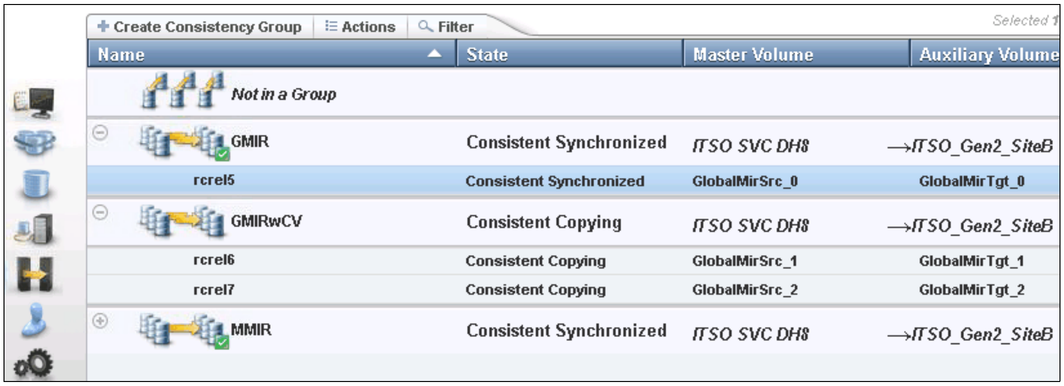
✓ Task completed. 100%

```
Task started.                                                6:09 PM
Creating relationship between Mirror Volume and Mirrored Volume 6:09 PM
Running command:                                             6:09 PM
svctask mkrcrelationship -aux 'Mirrored Volume' -cluster     6:09 PM
00000200A0E0EB56 -master 'Mirror Volume'
Starting Remote-Copy Relationship (0)                        6:09 PM
Running command:                                             6:09 PM
svctask startrcrelationship -force 0                         6:09 PM
Synchronizing memory cache.                                  6:09 PM
The task is 100% complete.                                   6:09 PM
Task completed.                                              6:09 PM
```

*Figure 11-125 Creation of Remote Copy relationship complete*

The relationships are visible in the Remote Copy pane. If you selected to copy the data, you can see that the status is Consistent Copying. You can check the copying progress in the Running Tasks status area.

After the copy is finished, the relationship status changes to Consistent synchronized. Figure 11-126 shows the Consistent Synchronized status.



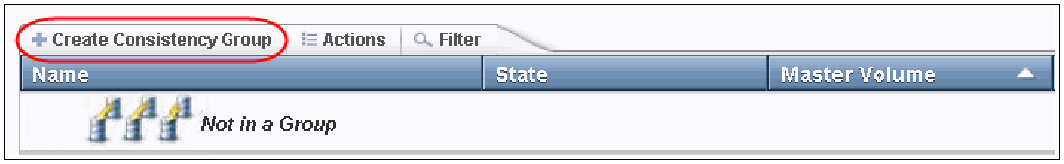
| Name | State | Master Volume | Auxiliary Volume |
|---|---|---|---|
| Not in a Group | | | |
| GMIR | Consistent Synchronized | ITSO SVC DH8 | →ITSO_Gen2_SiteB |
| rcrel5 | Consistent Synchronized | GlobalMirSrc_0 | GlobalMirTgt_0 |
| GMIRwCV | Consistent Copying | ITSO SVC DH8 | →ITSO_Gen2_SiteB |
| rcrel6 | Consistent Copying | GlobalMirSrc_1 | GlobalMirTgt_1 |
| rcrel7 | Consistent Copying | GlobalMirSrc_2 | GlobalMirTgt_2 |
| MMIR | Consistent Synchronized | ITSO SVC DH8 | →ITSO_Gen2_SiteB |

*Figure 11-126 Consistent copy of the mirrored volumes*

## 11.11.3 Creating Consistency Group

To create a Consistency Group, complete the following steps:

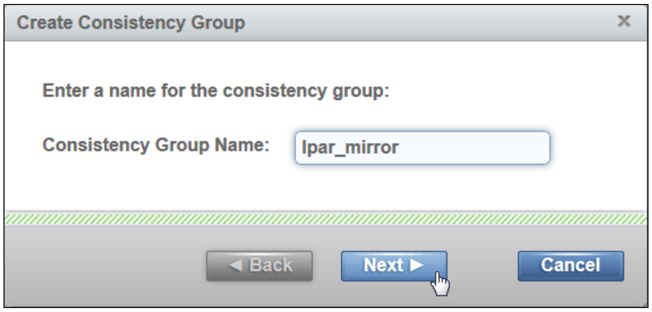
1. From the System pane, select **Copy Services** → **Remote Copy**.
2. Click **Create Consistency Group**, as shown in Figure 11-127.



| Name | State | Master Volume |
|---|---|---|
| Not in a Group | | |

*Figure 11-127 Selecting the Create Consistency Group option*

3. Enter a name for the Consistency Group, and then, click **Next**, as shown in Figure 11-128.


Create Consistency Group

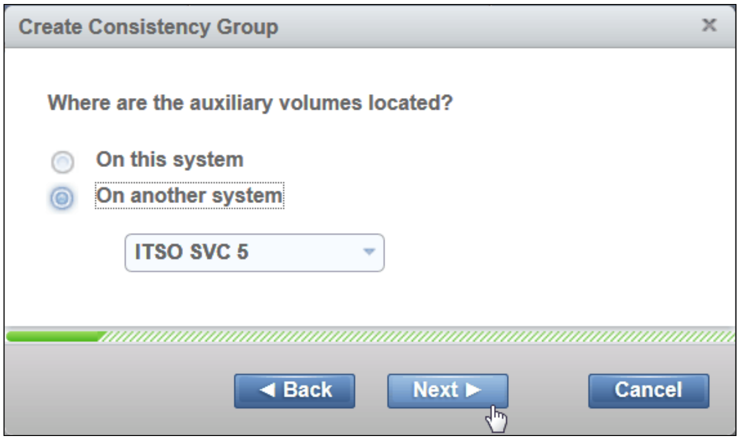
Enter a name for the consistency group:

Consistency Group Name: lpar_mirror

*Figure 11-128 Enter a Consistency Group name*

4. In the next window, select where the auxiliary volumes are located, as shown in Figure 11-129 on page 542:
   - On this system, which means that the volumes are local
   - On another system, which means that you select the remote system in the drop-down list

After you make a selection, click **Next**.



Create Consistency Group

Where are the auxiliary volumes located?

○ On this system
● On another system

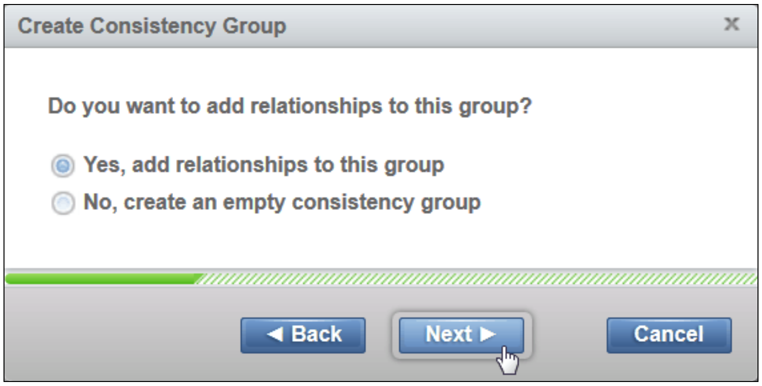
ITSO SVC 5

*Figure 11-129 Location of auxiliary volumes*

5. Select whether you want to add relationships to this group, as shown in Figure 11-130. The following options are available:

   – If you select Yes, click **Next** to continue the wizard and go to step 6.

   – If you select No, click **Finish** to create an empty Consistency Group that can be used later.



Create Consistency Group

Do you want to add relationships to this group?

● Yes, add relationships to this group
○ No, create an empty consistency group

*Figure 11-130 Add relationships to this group*

6. Select one of the following types of relationships to create, as shown in Figure 11-131 on page 543:

   – Metro Mirror

   – Global Mirror.

   – Global Mirror with Change Volumes

   Click **Next**.

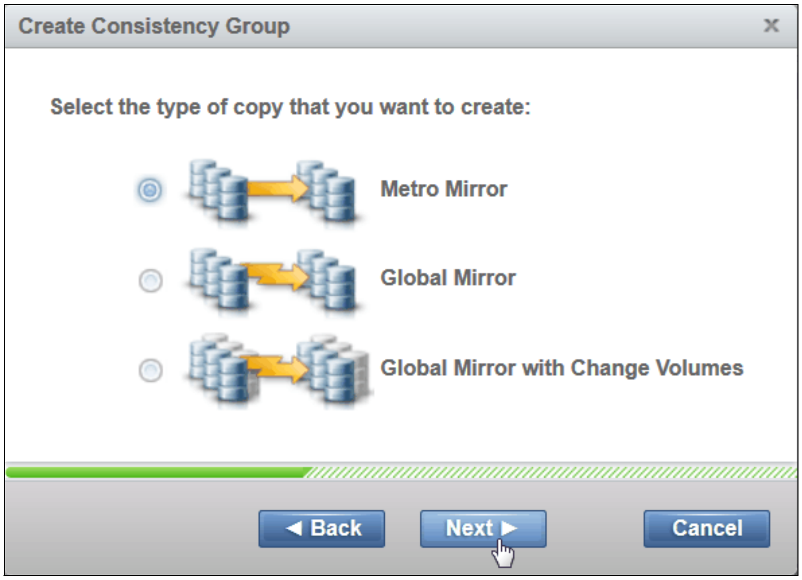
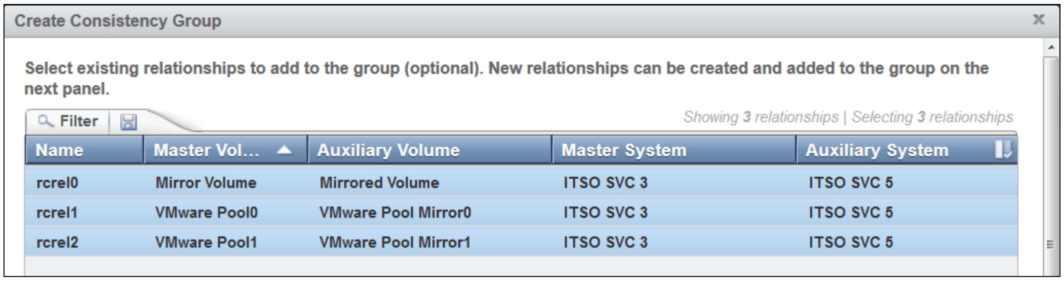

Create Consistency Group
Select the type of copy that you want to create:
● Metro Mirror
○ Global Mirror
○ Global Mirror with Change Volumes
◄ Back
Next ►
Cancel

*Figure 11-131 Select the type of relationship that you want to create*

7. As shown in Figure 11-132, you can optionally select existing relationships to add to the group. Click **Next**.

> **Note:** To select multiple relationships, hold down Ctrl and click the entries that you want to include.



Select existing relationships to add to the group (optional). New relationships can be created and added to the group on the next panel.

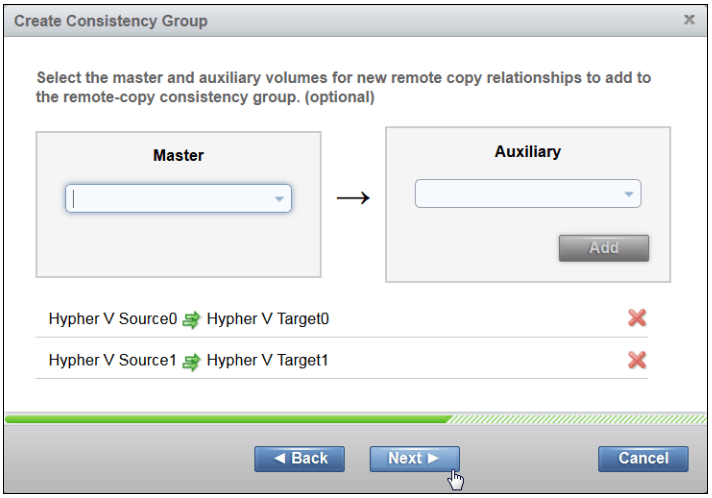
Showing 3 relationships | Selecting 3 relationships

| Name | Master Vol... | Auxiliary Volume | Master System | Auxiliary System |
|---|---|---|---|---|
| rcrel0 | Mirror Volume | Mirrored Volume | ITSO SVC 3 | ITSO SVC 5 |
| rcrel1 | VMware Pool0 | VMware Pool Mirror0 | ITSO SVC 3 | ITSO SVC 5 |
| rcrel2 | VMware Pool1 | VMware Pool Mirror1 | ITSO SVC 3 | ITSO SVC 5 |

*Figure 11-132 Select existing relationships to add to the group*

8. In the window that is shown in Figure 11-133 on page 544, you can create relationships. Select a volume in the Master drop-down list. Then, select a volume in the Auxiliary drop-down list for this master. Click **Add**, as shown in Figure 11-133 on page 544. Repeat this step to create other relationships, if needed.

   To remove a relationship that was created, click ✖ (Figure 11-133 on page 544). After all of the relationships that you want to create are displayed, click **Next**.


Create Consistency Group

Select the master and auxiliary volumes for new remote copy relationships to add to the remote-copy consistency group. (optional)

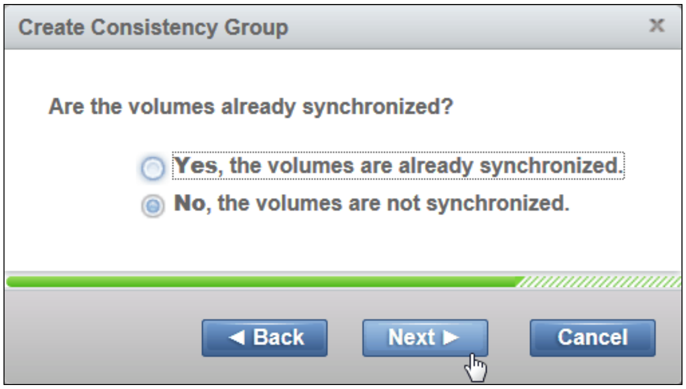
Master → Auxiliary

Hypher V Source0 → Hypher V Target0

Hypher V Source1 → Hypher V Target1

*Figure 11-133 Create relationships between the master and auxiliary volumes*

9. Specify whether the volumes are already synchronized. Then, click **Next** (Figure 11-134).

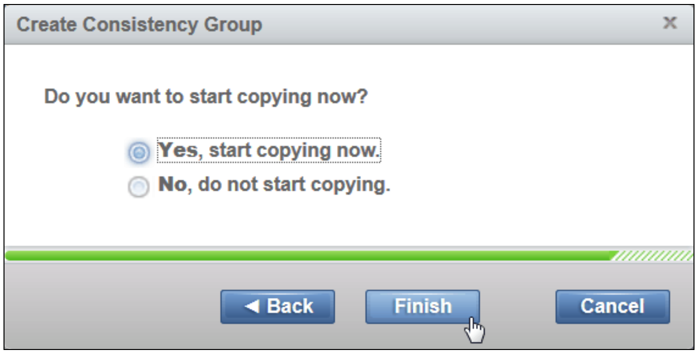

Create Consistency Group

Are the volumes already synchronized?

○ Yes, the volumes are already synchronized.
● No, the volumes are not synchronized.

*Figure 11-134 Volumes are already synchronized*

10. In the last window, select whether you want to start to copy the data. Then, click **Finish**, as shown in Figure 11-135.

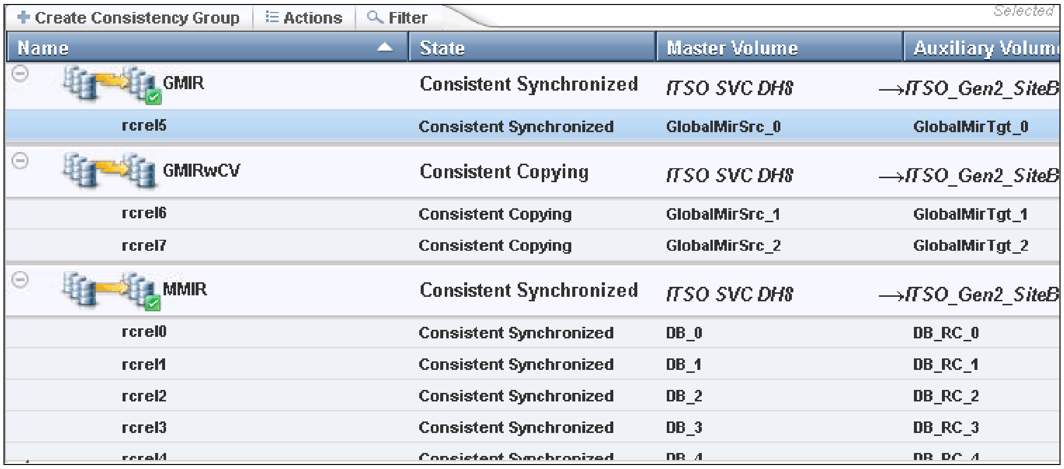

Create Consistency Group

Do you want to start copying now?

● Yes, start copying now.
○ No, do not start copying.

*Figure 11-135 Synchronize now*

11. The relationships are visible in the Remote Copy pane. If you selected to copy the data, you can see that the status of the relationships is Inconsistent copying. You can check the copying progress in the Running Tasks status area, as shown in Figure 11-136 on page 545.



| Name | State | Master Volume | Auxiliary Volum |
|---|---|---|---|
| GMIR | Consistent Synchronized | ITSO SVC DH8 | →ITSO_Gen2_SiteB |
| rcrel5 | Consistent Synchronized | GlobalMirSrc_0 | GlobalMirTgt_0 |
| GMIRwCV | Consistent Copying | ITSO SVC DH8 | →ITSO_Gen2_SiteB |
| rcrel6 | Consistent Copying | GlobalMirSrc_1 | GlobalMirTgt_1 |
| rcrel7 | Consistent Copying | GlobalMirSrc_2 | GlobalMirTgt_2 |
| MMIR | Consistent Synchronized | ITSO SVC DH8 | →ITSO_Gen2_SiteB |
| rcrel0 | Consistent Synchronized | DB_0 | DB_RC_0 |
| rcrel1 | Consistent Synchronized | DB_1 | DB_RC_1 |
| rcrel2 | Consistent Synchronized | DB_2 | DB_RC_2 |
| rcrel3 | Consistent Synchronized | DB_3 | DB_RC_3 |
| rcrel4 | Consistent Synchronized | DB_4 | DB_RC_4 |

*Figure 11-136 Consistency Group created with relationship in copying and synchronized status*

After the copies are completed, the relationships and the Consistency Group change to the Consistent Synchronized status.

## 11.11.4 Renaming Consistency Group

To rename a Consistency Group, complete the following steps:

1. From the System pane, select **Copy Services** → **Remote Copy**.
2. In the pane, select the Consistency Group that you want to rename. Then, select **Actions** → **Rename**, as shown in Figure 11-137.

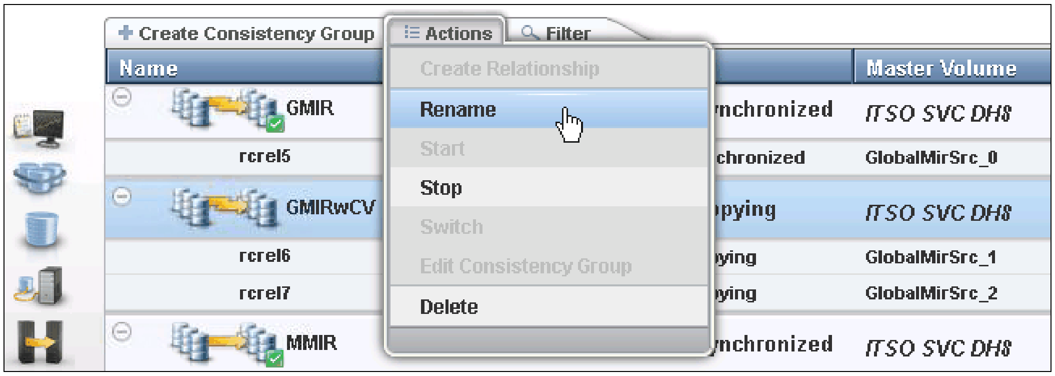
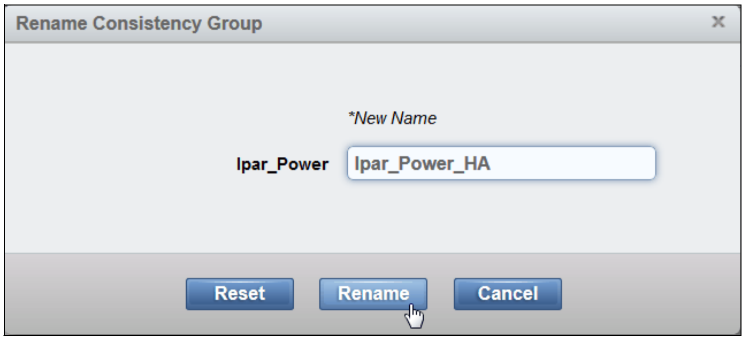


*Figure 11-137 Renaming a Consistency Group*

3. Enter the new name that you want to assign to the Consistency Group and click **Rename**, as shown in Figure 11-138 on page 546.



*New Name

lpar_Power: lpar_Power_HA

Figure 11-138   Changing the name for a Consistency Group

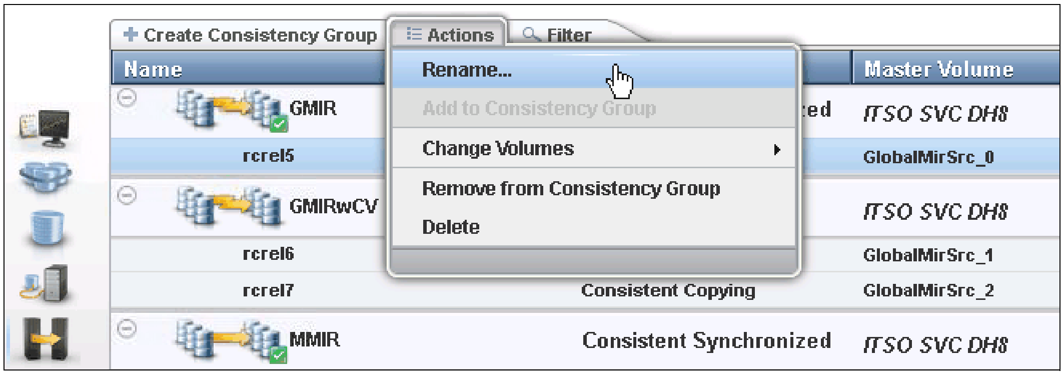
The new Consistency Group name is displayed on the Remote Copy pane.

## 11.11.5 Renaming remote copy relationship

Complete the following steps to rename a remote copy relationship:

1. From the System pane, select **Copy Services** → **Remote Copy**.
2. In the table, select the remote copy relationship mapping that you want to rename. Click **Actions** → **Rename**, as shown in Figure 11-139.

> **Tip:** You can also right-click a remote copy relationship and select **Rename**.



| Name | | Master Volume |
|---|---|---|
| GMIR | ...ted | ITSO SVC DH8 |
| rcrel5 | | GlobalMirSrc_0 |
| GMIRwCV | | ITSO SVC DH8 |
| rcrel6 | | GlobalMirSrc_1 |
| rcrel7 | Consistent Copying | GlobalMirSrc_2 |
| MMIR | Consistent Synchronized | ITSO SVC DH8 |

Figure 11-139   Rename remote copy relationship action

3. In the Rename Relationship window, enter the new name that you want to assign to the FlashCopy mapping and click **Rename**, as shown in Figure 11-140 on page 547.

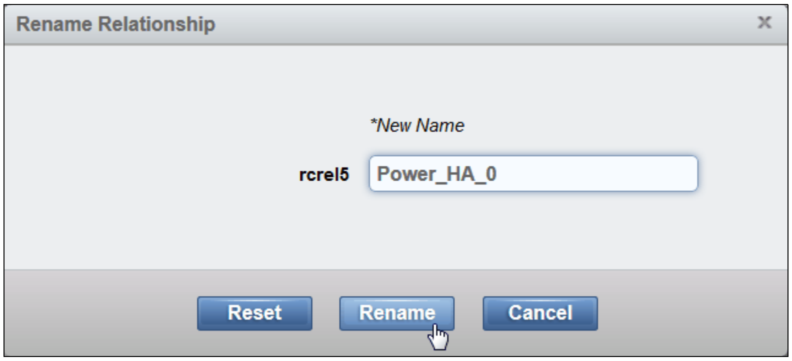


*Figure 11-140 Renaming a remote copy relationship*

> **Remote copy relationship name:** You can use the letters A - Z and a - z, the numbers 0 - 9, and the underscore (_) character. The remote copy name can be 1 - 15 characters. No blanks are allowed.

## 11.11.6 Moving stand-alone remote copy relationship to Consistency Group

Complete the following steps to move a remote copy relationship to a Consistency Group:

1. From the System pane, click **Copy Services** → **Remote Copy**.
2. Expand the **Not in a Group** column.
3. Select the relationship that you want to move to the Consistency Group**.**
4. Click **Actions** → **Add to Consistency Group**, as shown in Figure 11-141.

> **Tip:** You can also right-click a remote copy relationship and select **Add to Consistency Group**.

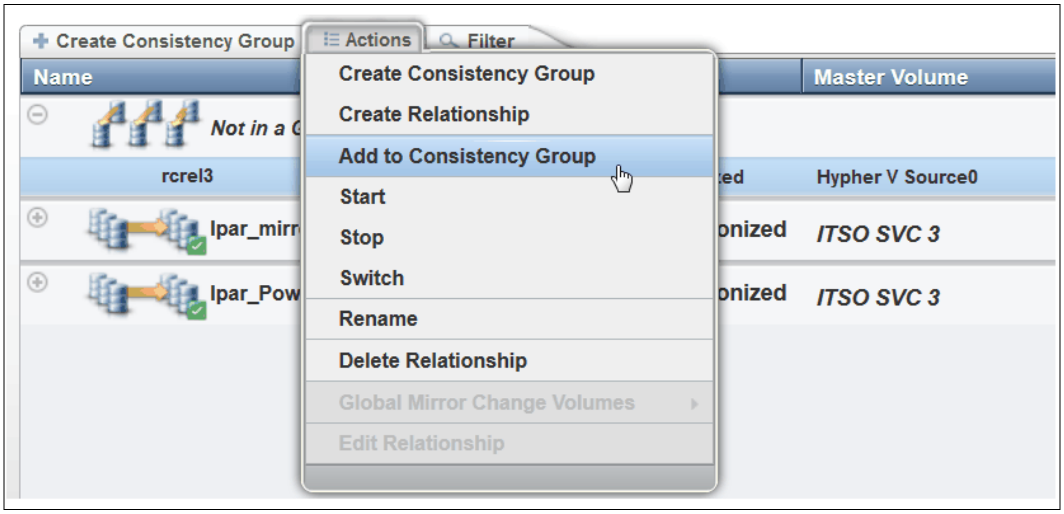
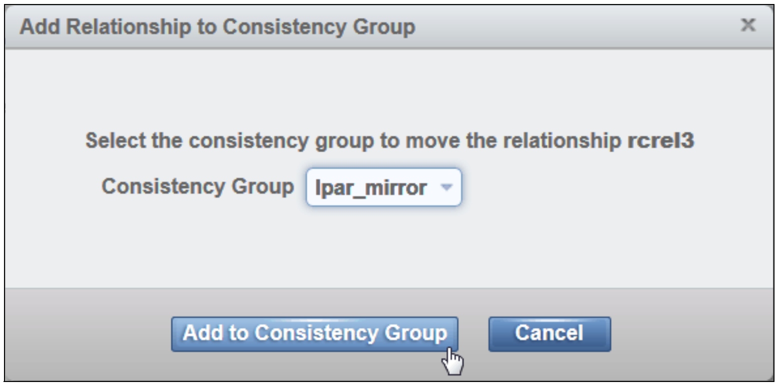


*Figure 11-141 Add to Consistency Group action*

5. In the Add Relationship to Consistency Group window, select the Consistency Group for this remote copy relationship by using the drop-down list, as shown in Figure 11-142 on page 548. Click **Add to Consistency Group** to confirm your changes.



Select the consistency group to move the relationship rcrel3

Consistency Group: lpar_mirror

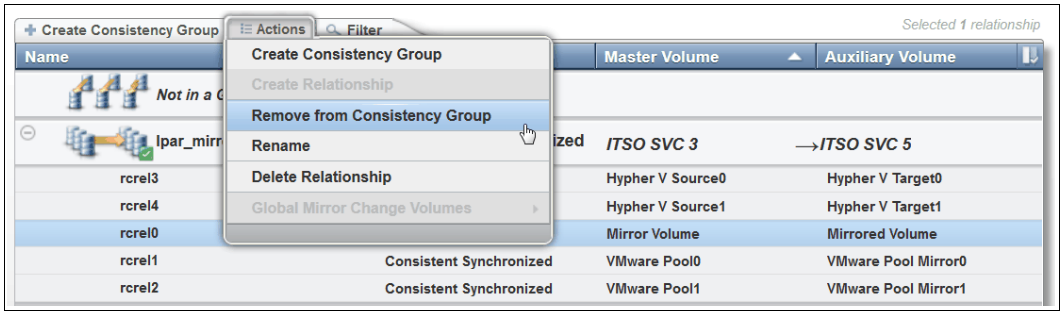
*Figure 11-142 Adding a relationship to a Consistency Group*

## 11.11.7 Removing remote copy relationship from Consistency Group

Complete the following steps to remove a remote copy relationship from a Consistency Group:

1. From the System pane, select **Copy Services** → **Remote Copy**.
2. Select a Consistency Group.
3. Select the remote copy relationship that you want to remove from the Consistency Group.
4. Click **Actions** → **Remove from Consistency Group**, as shown in Figure 11-143.

> **Tip:** You can also right-click a remote copy relationship and select **Remove from Consistency Group**.

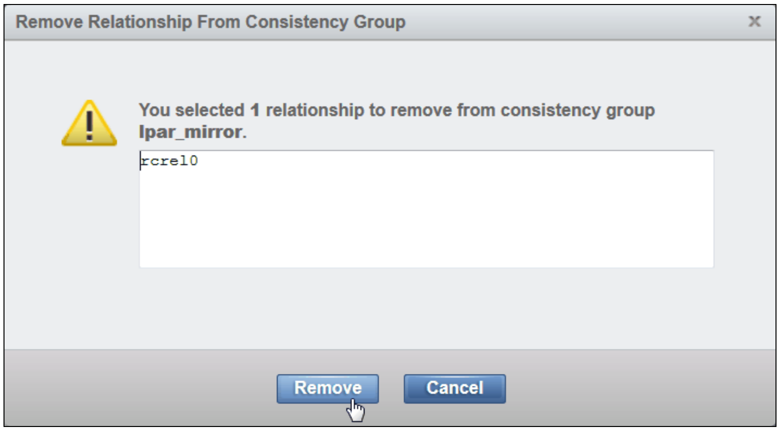


Selected 1 relationship

| Name | | Master Volume | Auxiliary Volume |
|---|---|---|---|
| Not in a C | | | |
| lpar_mirr | ized | ITSO SVC 3 | →ITSO SVC 5 |
| rcrel3 | | Hypher V Source0 | Hypher V Target0 |
| rcrel4 | | Hypher V Source1 | Hypher V Target1 |
| rcrel0 | | Mirror Volume | Mirrored Volume |
| rcrel1 | Consistent Synchronized | VMware Pool0 | VMware Pool Mirror0 |
| rcrel2 | Consistent Synchronized | VMware Pool1 | VMware Pool Mirror1 |

*Figure 11-143 Remove from Consistency Group action*

5. In the Remove Relationship From Consistency Group window, click **Remove**, as shown in Figure 11-144 on page 549.



Remove Relationship From Consistency Group

You selected 1 relationship to remove from consistency group **lpar_mirror**.

rcrel0

Figure 11-144 Remove a relationship from a Consistency Group

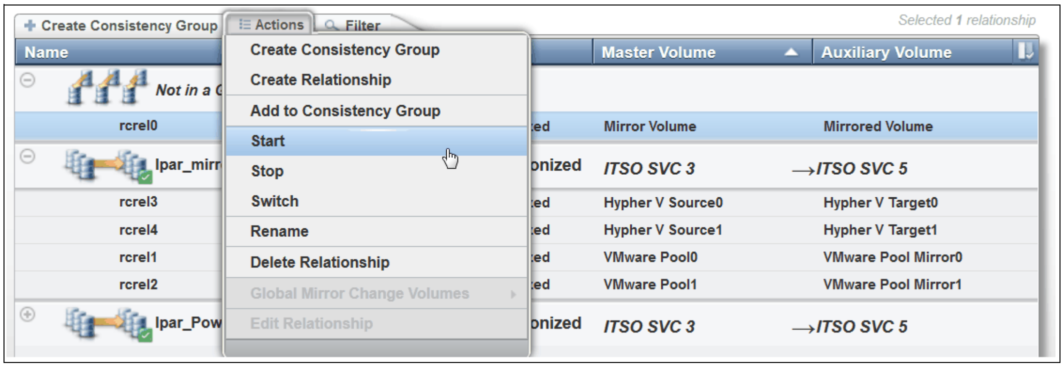
## 11.11.8 Starting remote copy relationship

When a remote copy relationship is created, the remote copy process can be started. Only relationships that are not members of a Consistency Group, or the only relationship in a Consistency Group, can be started individually.

Complete the following steps to start a remote copy relationship:

1. From the System pane, select **Copy Services** → **Remote Copy**.
2. Expand the **Not in a Group** column.
3. In the table, select the remote copy relationship that you want to start.
4. Click **Actions** → **Start** to start the remote copy process, as shown in Figure 11-145 on page 549.

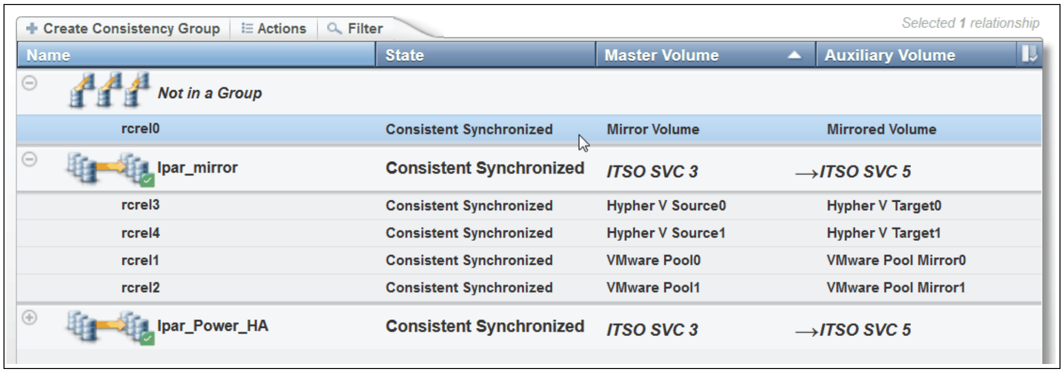
> **Tip:** You can also right-click a relationship and select **Start** from the list.



| Name | | Master Volume | Auxiliary Volume |
|---|---|---|---|
| Not in a G | | | |
| rcrel0 | ted | Mirror Volume | Mirrored Volume |
| lpar_mirr | onized | ITSO SVC 3 | →ITSO SVC 5 |
| rcrel3 | ted | Hypher V Source0 | Hypher V Target0 |
| rcrel4 | ted | Hypher V Source1 | Hypher V Target1 |
| rcrel1 | ted | VMware Pool0 | VMware Pool Mirror0 |
| rcrel2 | ted | VMware Pool1 | VMware Pool Mirror1 |
| lpar_Pow | onized | ITSO SVC 3 | →ITSO SVC 5 |

Figure 11-145 Starting the remote copy process

5. After the task is complete, the remote copy relationship status has a Consistent Synchronized state, as shown in Figure 11-146 on page 550.



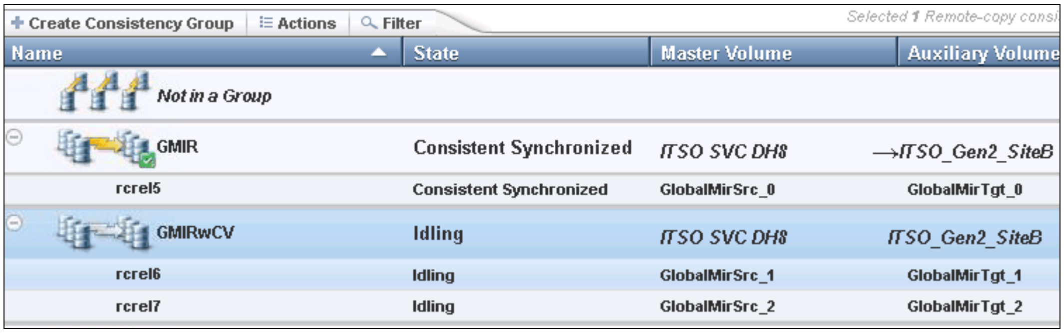
| Name | State | Master Volume | Auxiliary Volume |
|---|---|---|---|
| Not in a Group | | | |
| rcrel0 | Consistent Synchronized | Mirror Volume | Mirrored Volume |
| lpar_mirror | Consistent Synchronized | ITSO SVC 3 | →ITSO SVC 5 |
| rcrel3 | Consistent Synchronized | Hypher V Source0 | Hypher V Target0 |
| rcrel4 | Consistent Synchronized | Hypher V Source1 | Hypher V Target1 |
| rcrel1 | Consistent Synchronized | VMware Pool0 | VMware Pool Mirror0 |
| rcrel2 | Consistent Synchronized | VMware Pool1 | VMware Pool Mirror1 |
| lpar_Power_HA | Consistent Synchronized | ITSO SVC 3 | →ITSO SVC 5 |

*Figure 11-146 Consistent Synchronized remote copy relationship*

## 11.11.9 Starting remote copy Consistency Group

All of the mappings in a Consistency Group are brought to the same state. To start the remote copy Consistency Group, complete the following steps:

1. From the System pane, select **Copy Services** → **Remote Copy**.
2. Select the Consistency Group that you want to start, as shown in Figure 11-147.



| Name | State | Master Volume | Auxiliary Volume |
|---|---|---|---|
| Not in a Group | | | |
| GMIR | Consistent Synchronized | ITSO SVC DH8 | →ITSO_Gen2_SiteB |
| rcrel5 | Consistent Synchronized | GlobalMirSrc_0 | GlobalMirTgt_0 |
| GMIRwCV | Idling | ITSO SVC DH8 | ITSO_Gen2_SiteB |
| rcrel6 | Idling | GlobalMirSrc_1 | GlobalMirTgt_1 |
| rcrel7 | Idling | GlobalMirSrc_2 | GlobalMirTgt_2 |

*Figure 11-147 Remote Copy Consistency Groups view*

3. Click **Actions** → **Start** (Figure 11-148) to start the remote copy Consistency Group.

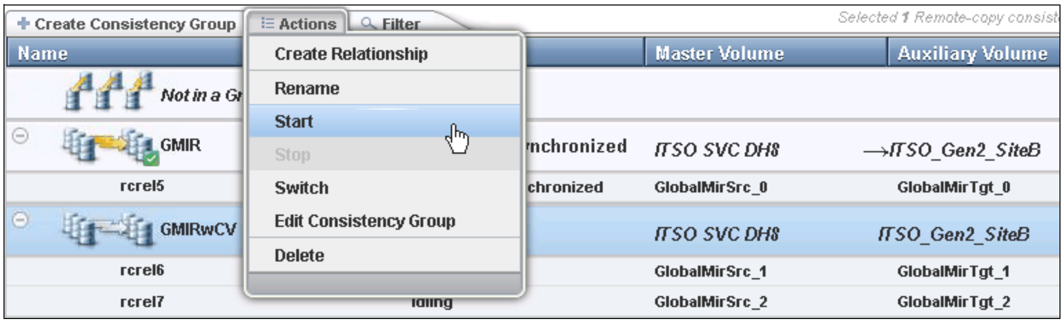
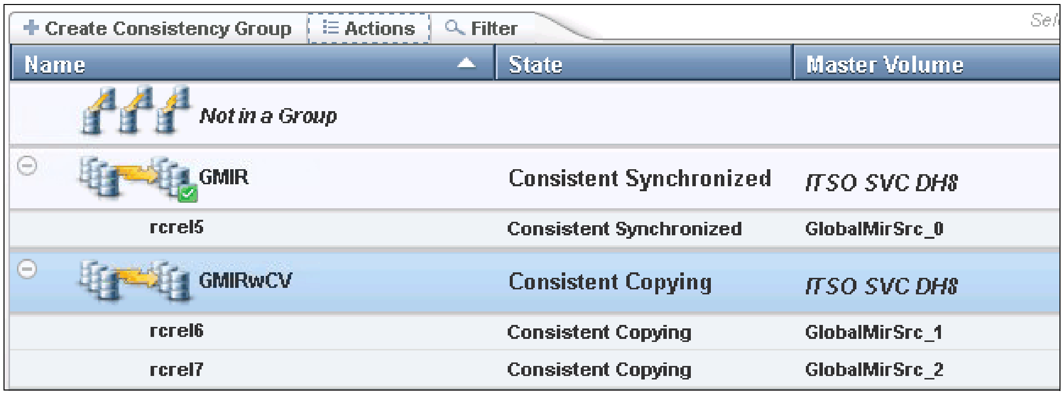


*Figure 11-148 Start action*

4. You can check the remote copy Consistency Group progress, as shown in Figure 11-149 on page 551.



| Name | State | Master Volume |
|---|---|---|
| Not in a Group | | |
| GMIR | Consistent Synchronized | ITSO SVC DH8 |
| rcrel5 | Consistent Synchronized | GlobalMirSrc_0 |
| GMIRwCV | Consistent Copying | ITSO SVC DH8 |
| rcrel6 | Consistent Copying | GlobalMirSrc_1 |
| rcrel7 | Consistent Copying | GlobalMirSrc_2 |

*Figure 11-149 Checking the remote copy Consistency Group progress*

5. After the task completes, the Consistency Group and all of its relationships becomes in a Consistent Synchronized state.

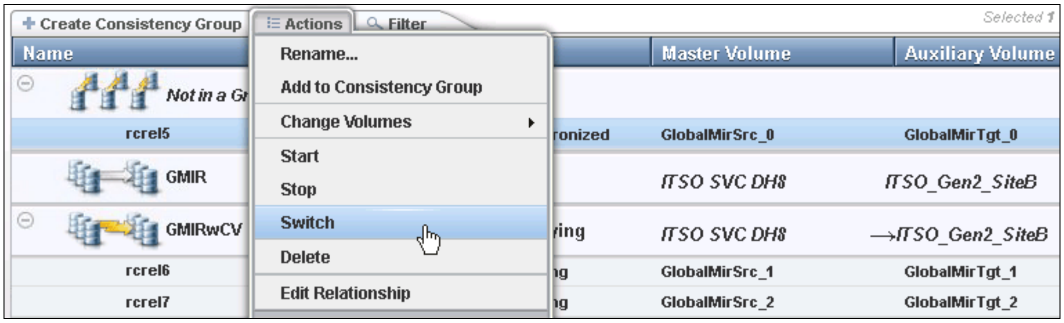
## 11.11.10 Switching copy direction

When a remote copy relationship is in the Consistent synchronized state, the copy direction for the relationship can be changed. Only relationships that are not a member of a Consistency Group (or the only relationship in a Consistency Group) can be switched individually. These relationships can be switched from master to auxiliary or from auxiliary to master, depending on the case.

Complete the following steps to switch a remote copy relationship:

1. From the System pane, select **Copy Services** → **Remote Copy**.
2. Expand the **Not in a Group** column.
3. In the table, select the remote copy relationship that you want to switch.
4. Click **Actions** → **Switch** (Figure 11-150) to start the remote copy process.

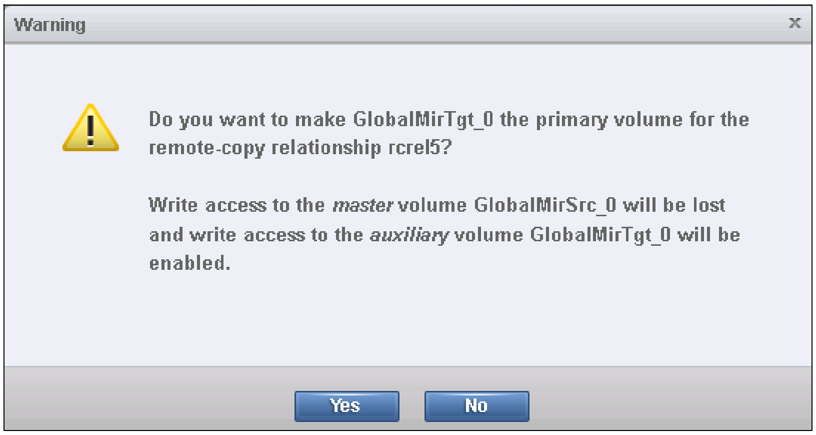
> **Tip:** You can also right-click a relationship and select **Switch**.



| Name | | Master Volume | Auxiliary Volume |
|---|---|---|---|
| Not in a Gr | | | |
| rcrel5 | ronized | GlobalMirSrc_0 | GlobalMirTgt_0 |
| GMIR | | ITSO SVC DH8 | ITSO_Gen2_SiteB |
| GMIRwCV | ying | ITSO SVC DH8 | →ITSO_Gen2_SiteB |
| rcrel6 | ng | GlobalMirSrc_1 | GlobalMirTgt_1 |
| rcrel7 | ng | GlobalMirSrc_2 | GlobalMirTgt_2 |

*Figure 11-150 Switch copy direction action*

5. The Warning window that is shown in Figure 11-151 on page 552 opens. A confirmation is needed to switch the remote copy relationship direction. The remote copy is switched from the master volume to the auxiliary volume. Click **Yes**.

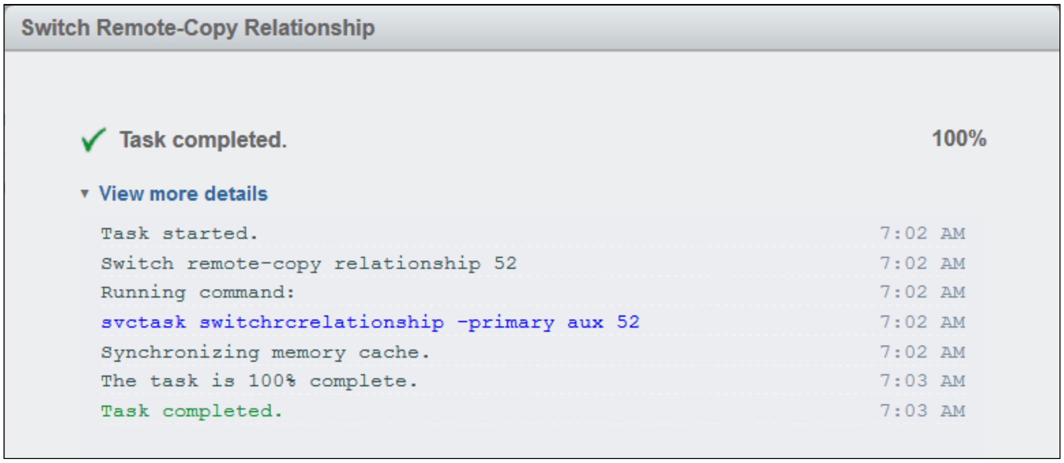


Do you want to make GlobalMirTgt_0 the primary volume for the remote-copy relationship rcrel5?

Write access to the *master* volume GlobalMirSrc_0 will be lost and write access to the *auxiliary* volume GlobalMirTgt_0 will be enabled.

*Figure 11-151 Warning window*

Figure 11-152 shows the command-line output about this task.



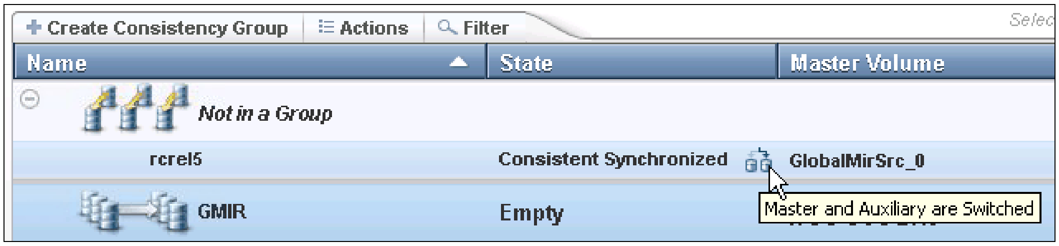
✓ Task completed. 100%

▾ View more details

```
Task started.                                          7:02 AM
Switch remote-copy relationship 52                     7:02 AM
Running command:                                       7:02 AM
svctask switchrcrelationship -primary aux 52           7:02 AM
Synchronizing memory cache.                            7:02 AM
The task is 100% complete.                             7:03 AM
Task completed.                                        7:03 AM
```

*Figure 11-152 Command-line output for switch relationship action*

The copy direction is now switched, as shown in Figure 11-153. The auxiliary volume is now accessible and shown as the primary volume. Also, the auxiliary volume is now synchronized to the master volume.



| Name | State | Master Volume |
|---|---|---|
| Not in a Group | | |
| rcrel5 | Consistent Synchronized | GlobalMirSrc_0 |
| GMIR | Empty | |

*Figure 11-153 Checking remote copy synchronization direction*

## 11.11.11 Switching the copy direction for a Consistency Group

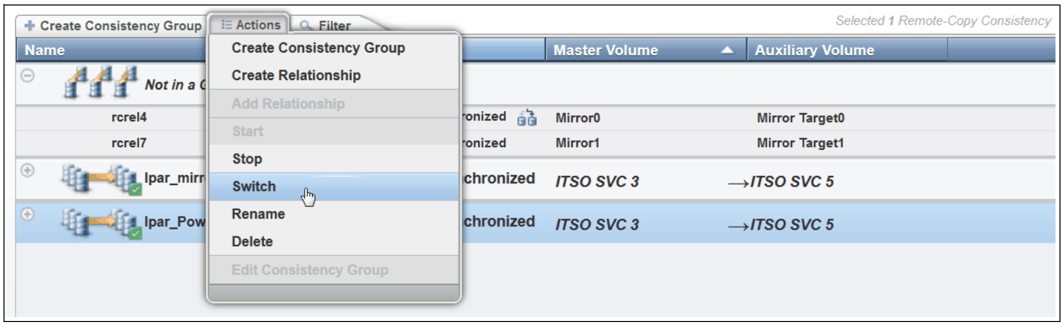
When a Consistency Group is in the Consistent Synchronized state, the copy direction for this Consistency Group can be changed.

> **Important:** When the copy direction is switched, it is crucial that no outstanding I/O exists to the volume that changes from primary to secondary because all of the I/O is inhibited to that volume when it becomes the secondary. Therefore, careful planning is required before you switch the copy direction for a Consistency Group.

Complete the following steps to switch a Consistency Group:

1. From the System pane, select **Copy Services** → **Remote Copy**.
2. Select the Consistency Group that you want to switch.
3. Click **Actions** → **Switch** (as shown in Figure 11-154 on page 553) to start the remote copy process.

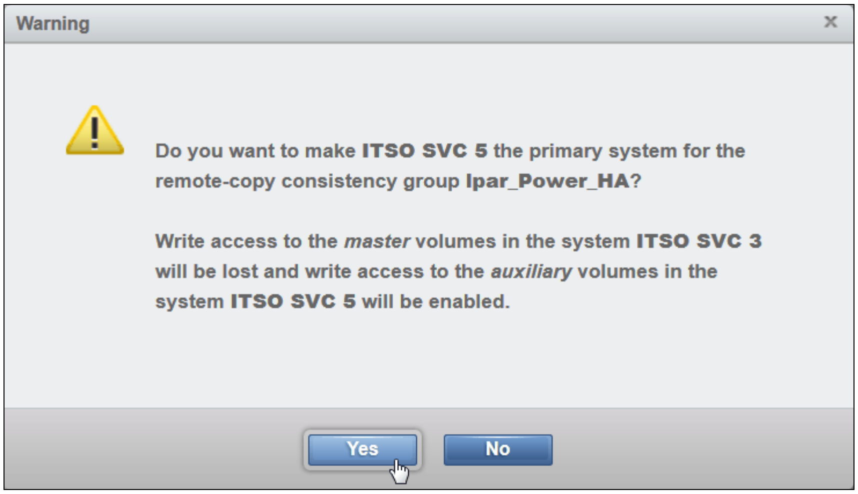
   > **Tip:** You can also right-click a relationship and select **Switch**.



| Name | | Master Volume | Auxiliary Volume |
|---|---|---|---|
| Not in a ( | | | |
| rcrel4 | ronized | Mirror0 | Mirror Target0 |
| rcrel7 | ronized | Mirror1 | Mirror Target1 |
| lpar_mirr | chronized | ITSO SVC 3 | →ITSO SVC 5 |
| lpar_Pow | chronized | ITSO SVC 3 | →ITSO SVC 5 |

*Figure 11-154 Switch action*

4. The warning window that is shown in Figure 11-155 opens. A confirmation is needed to switch the Consistency Group direction. In the example that is shown in Figure 11-155, the Consistency Group is switched from the master group to the auxiliary group. Click **Yes**.



Do you want to make **ITSO SVC 5** the primary system for the remote-copy consistency group **lpar_Power_HA**?

Write access to the *master* volumes in the system **ITSO SVC 3** will be lost and write access to the *auxiliary* volumes in the system **ITSO SVC 5** will be enabled.

*Figure 11-155 Warning window for ITSO SVC 3*

The remote copy direction is now switched. The auxiliary volume is now accessible and shown as a primary volume.

## 11.11.12 Stopping a remote copy relationship

After it is started, the remote copy process can be stopped, if needed. Only relationships that are not a member of a Consistency Group (or the only relationship in a Consistency Group) can be stopped individually. You can also use this command to enable write access to a consistent secondary volume.

Complete the following steps to stop a remote copy relationship:

1. From the System pane, select **Copy Services** → **Remote Copy**.
2. Expand the **Not in a Group** column.
3. In the table, select the remote copy relationship that you want to stop.
4. Click **Actions** → **Stop** (as shown in Figure 11-156) to stop the remote copy process.

> **Tip:** You can also right-click a relationship and select **Stop** from the list.

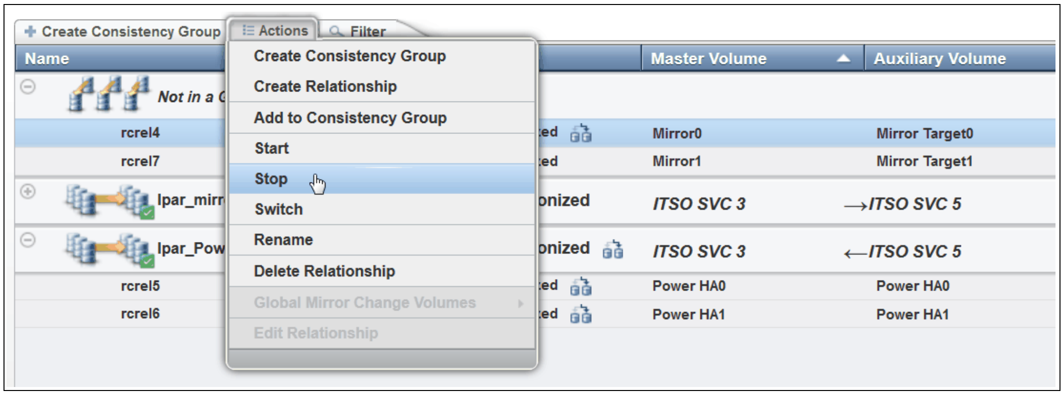
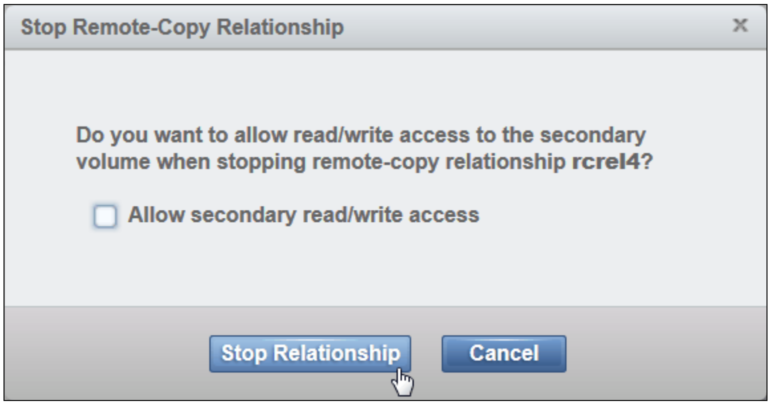


*Figure 11-156 Stop action*

5. The Stop Remote Copy Relationship window opens, as shown in Figure 11-157. To allow secondary read/write access, select **Allow secondary read/write access**. Then, click **Stop Relationship**.



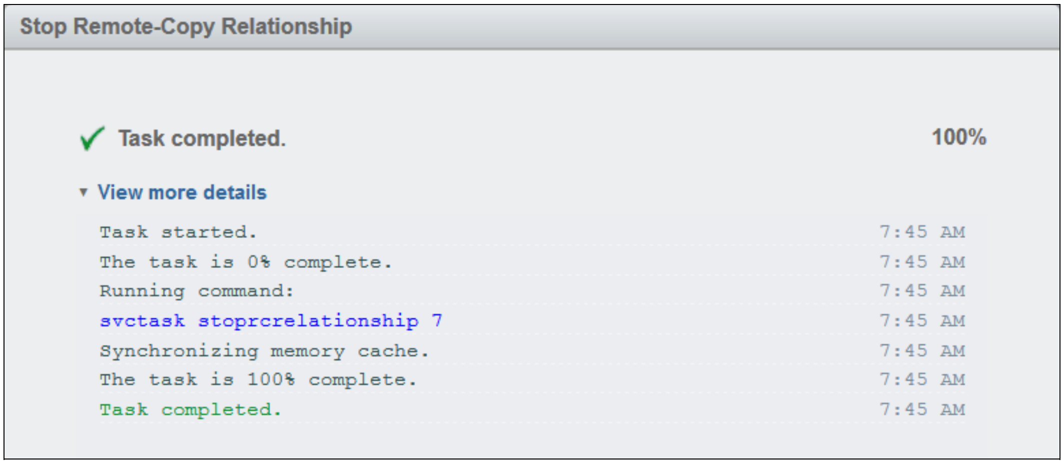
Stop Remote-Copy Relationship

Do you want to allow read/write access to the secondary volume when stopping remote-copy relationship rcrel4?

☐ Allow secondary read/write access

*Figure 11-157 Stop Remote Copy Relationship window*

6. Figure 11-158 shows the command-line output for the stop remote copy relationship.



Stop Remote-Copy Relationship

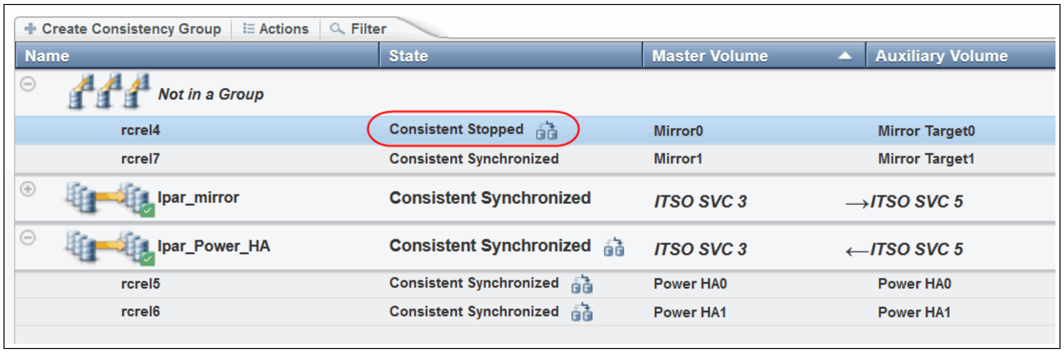
✓ Task completed. 100%

▾ View more details

```
Task started.                              7:45 AM
The task is 0% complete.                   7:45 AM
Running command:                           7:45 AM
svctask stoprcrelationship 7               7:45 AM
Synchronizing memory cache.                7:45 AM
The task is 100% complete.                 7:45 AM
Task completed.                            7:45 AM
```

*Figure 11-158 Stop remote copy relationship command-line output*

The new relationship status can be checked, as shown in Figure 11-159. The relationship is now Consistent Stopped.



| Name | State | Master Volume ▲ | Auxiliary Volume |
|---|---|---|---|
| *Not in a Group* | | | |
| rcrel4 | Consistent Stopped | Mirror0 | Mirror Target0 |
| rcrel7 | Consistent Synchronized | Mirror1 | Mirror Target1 |
| lpar_mirror | Consistent Synchronized | *ITSO SVC 3* | *→ITSO SVC 5* |
| lpar_Power_HA | Consistent Synchronized | *ITSO SVC 3* | *←ITSO SVC 5* |
| rcrel5 | Consistent Synchronized | Power HA0 | Power HA0 |
| rcrel6 | Consistent Synchronized | Power HA1 | Power HA1 |

*Figure 11-159 Checking remote copy synchronization status*

## 11.11.13 Stopping Consistency Group

After it is started, the Consistency Group can be stopped, if necessary. You can also use this task to enable write access to consistent secondary volumes.

Perform the following steps to stop a Consistency Group:

1. From the System pane, select **Copy Services** → **Remote Copy**.
2. In the table, select the Consistency Group that you want to stop.
3. Click **Actions** → **Stop** (as shown in Figure 11-160) to stop the remote copy Consistency Group.

> **Tip:** You can also right-click a relationship and select **Stop** from the list.

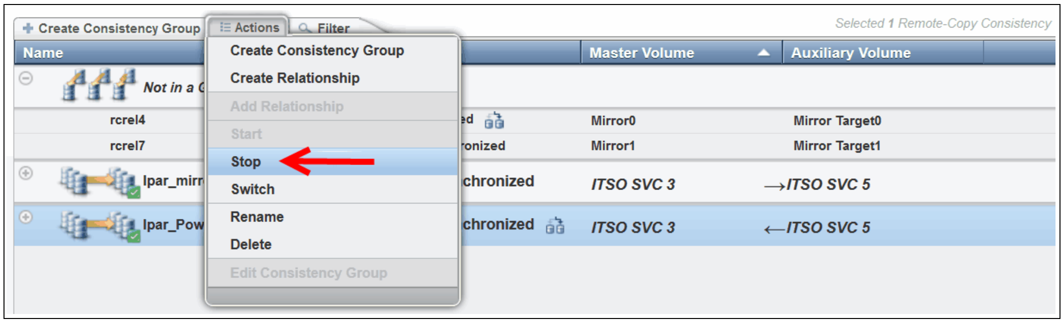
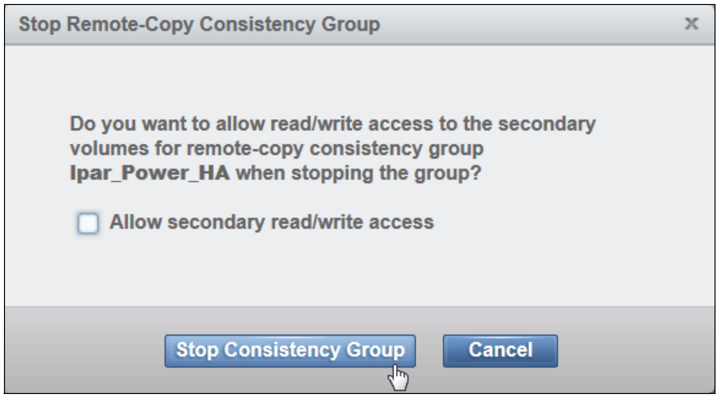


Figure 11-160 Selecting the Stop option

4. The Stop Remote Copy Consistency Group window opens, as shown in Figure 11-161. To allow secondary read/write access, select **Allow secondary read/write access**. Then, click **Stop Consistency Group**.

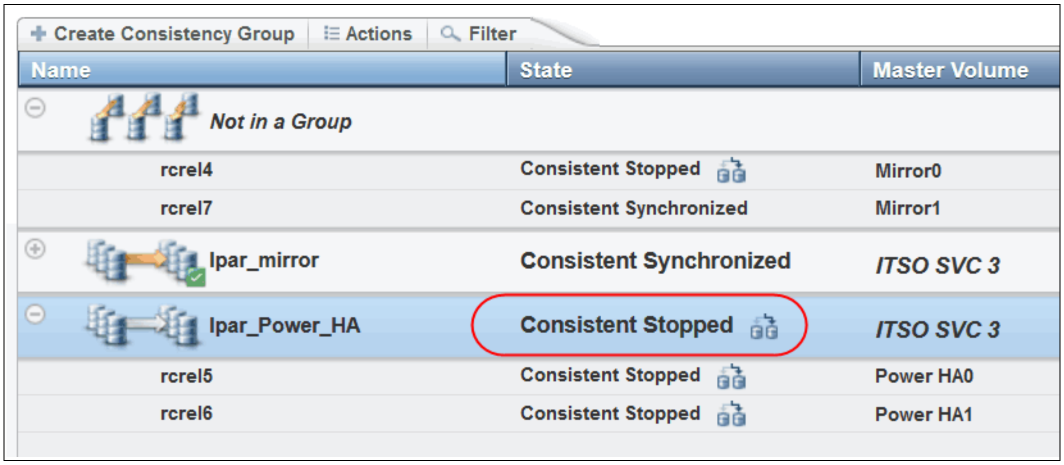

Stop Remote-Copy Consistency Group

Do you want to allow read/write access to the secondary volumes for remote-copy consistency group **lpar_Power_HA** when stopping the group?

□ Allow secondary read/write access

Figure 11-161 Stop Remote Copy Consistency Group window

The new relationship status can be checked, as shown in Figure 11-162 on page 556. The relationship is now Consistent Stopped.



| Name | State | Master Volume |
|---|---|---|
| *Not in a Group* | | |
| rcrel4 | Consistent Stopped | Mirror0 |
| rcrel7 | Consistent Synchronized | Mirror1 |
| lpar_mirror | Consistent Synchronized | *ITSO SVC 3* |
| lpar_Power_HA | Consistent Stopped | *ITSO SVC 3* |
| rcrel5 | Consistent Stopped | Power HA0 |
| rcrel6 | Consistent Stopped | Power HA1 |

Figure 11-162 Checking remote copy synchronization status

## 11.11.14 Deleting stand-alone remote copy relationships

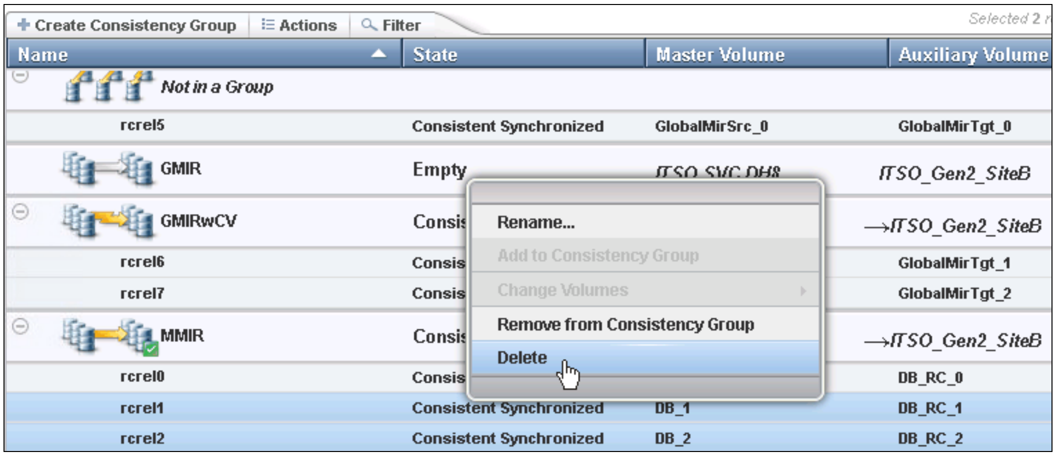
Complete the following steps to delete a stand-alone remote copy mapping:

1. From the System pane, select **Copy Services** → **Remote Copy**.
2. In the table, select the remote copy relationship that you want to delete.

> **Multiple remote copy mappings:** To select multiple remote copy mappings, hold down Ctrl and click the entries that you want.

3. Click **Actions** → **Delete**, as shown in Figure 11-163.

> **Tip:** You can also right-click a remote copy mapping and select **Delete**.

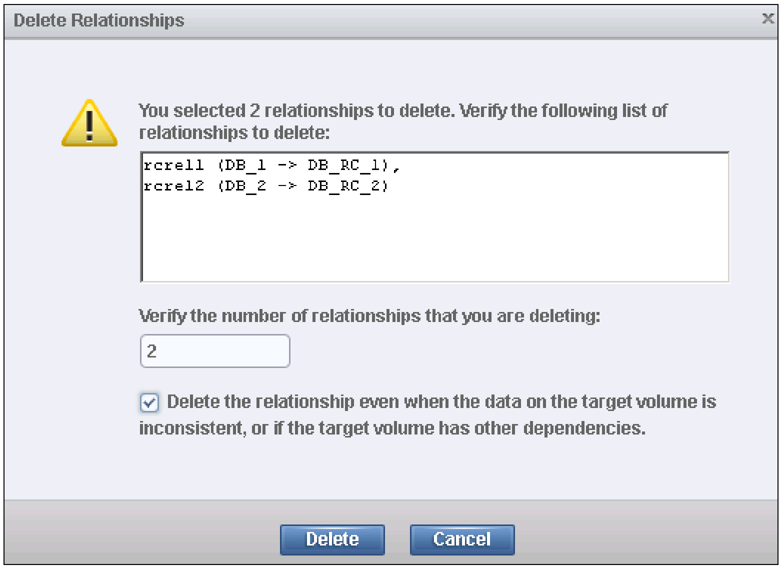


| Name | State | Master Volume | Auxiliary Volume |
|---|---|---|---|
| *Not in a Group* | | | |
| rcrel5 | Consistent Synchronized | GlobalMirSrc_0 | GlobalMirTgt_0 |
| GMIR | Empty | ITSO_SVC_DH8 | ITSO_Gen2_SiteB |
| GMIRwCV | Consis | | →ITSO_Gen2_SiteB |
| rcrel6 | Consis | | GlobalMirTgt_1 |
| rcrel7 | Consis | | GlobalMirTgt_2 |
| MMIR | Consis | | →ITSO_Gen2_SiteB |
| rcrel0 | Consis | | DB_RC_0 |
| rcrel1 | Consistent Synchronized | DB_1 | DB_RC_1 |
| rcrel2 | Consistent Synchronized | DB_2 | DB_RC_2 |

*Figure 11-163 Selecting the Delete Relationship option*

4. The Delete Relationship window opens (Figure 11-164 on page 557). In the “Verify the number of relationships that you are deleting” field, enter the number of volumes that you want to remove. This verification was added to help to avoid deleting the wrong relationships.

   Click **Delete**, as shown in Figure 11-164.



You selected 2 relationships to delete. Verify the following list of relationships to delete:

```
rcrel1 (DB_1 -> DB_RC_1),
rcrel2 (DB_2 -> DB_RC_2)
```

Verify the number of relationships that you are deleting:

2

☑ Delete the relationship even when the data on the target volume is inconsistent, or if the target volume has other dependencies.

*Figure 11-164 Delete remote copy relationship*

### 11.11.15 Deleting Consistency Group

> **Important:** Deleting a Consistency Group does not delete its remote copy mappings.

Complete the following steps to delete a Consistency Group:

1. From the System pane, select **Copy Services** → **Remote Copy**.
2. In the left column, select the Consistency Group that you want to delete.
3. Click **Actions** → **Delete**, as shown in Figure 11-165.

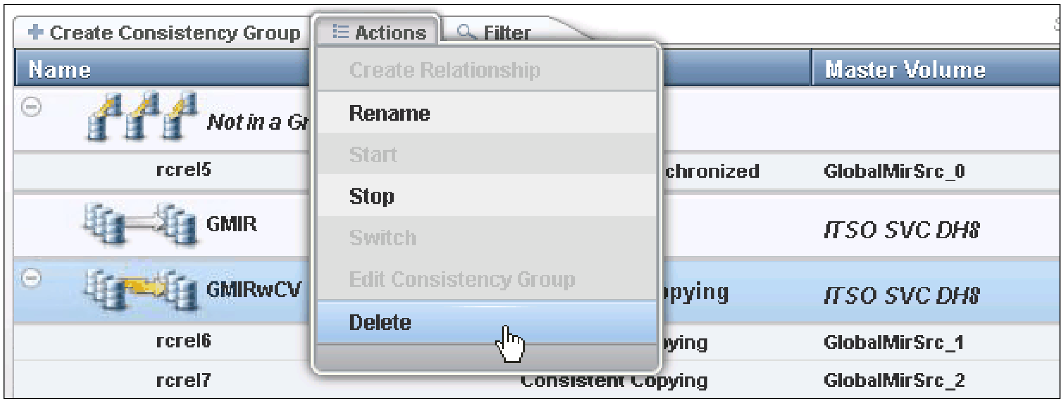
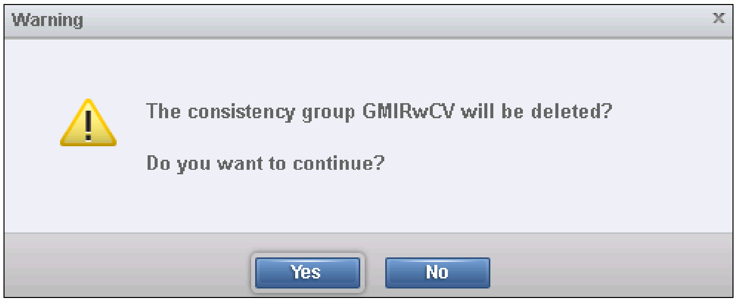


*Figure 11-165 Selecting the Delete Consistency Group option*

4. The warning window that is shown in Figure 11-166 opens. Click **Yes**.



The consistency group GMIRwCV will be deleted?

Do you want to continue?

*Figure 11-166 Confirmation message*

## 11.12 Troubleshooting remote copy

Remote copy (Metro Mirror and Global Mirror) has two primary error codes that are displayed: `1920` or `1720`. A `1920` is a congestion error. This error means that the source, the link between the source and target, or the target cannot keep up with the requested copy rate. A `1720` error is a heartbeat or system partnership communication error. This error often is more serious because failing communication between your system partners involves extended diagnostic time.

### 11.12.1 1920 error

A `1920` error (event ID `050010`) can have several triggers, including the following probable causes:

- Primary 2145 system or SAN fabric problem (10%)
- Primary 2145 system or SAN fabric configuration (10%)
- Secondary 2145 system or SAN fabric problem (15%)
- Secondary 2145 system or SAN fabric configuration (25%)
- Intercluster link problem (15%)
- Intercluster link configuration (25%)

In practice, the most often overlooked cause is latency. Global Mirror has a round-trip-time tolerance limit of 80 or 250 milliseconds, depending on the firmware version and the hardware model. See Figure 11-103 on page 512. A message that is sent from your source IBM Spectrum Virtualize system to your target system and the accompanying acknowledgment must have a total time of 80 or 250 milliseconds round trip. In other words, it must have up to 40 or 125 milliseconds latency each way.

The primary component of your round-trip time is the physical distance between sites. For every 1000 kilometers (621.4 miles), you observe a 5-millisecond delay each way. This delay does not include the time that is added by equipment in the path. Every device adds a varying amount of time depending on the device, but a good rule is 25 microseconds for pure hardware devices.

For software-based functions (such as compression that is implemented in applications), the added delay tends to be much higher (usually in the millisecond plus range.) Next, we describe an example of a physical delay.

Company A has a production site that is 1900 kilometers (1180.6 miles) away from its recovery site. The network service provider uses a total of five devices to connect the two sites. In addition to those devices, Company A employs a SAN FC router at each site to provide Fibre Channel over IP (FCIP) to encapsulate the FC traffic between sites.

Now, there are seven devices, and 1900 kilometers (1180.6 miles) of distance delay. All the devices are adding 200 microseconds of delay each way. The distance adds 9.5 milliseconds each way, for a total of 19 milliseconds. Combined with the device latency, the delay is 19.4 milliseconds of physical latency minimum, which is under the 80-millisecond limit of Global Mirror until you realize that this number is the best case number.

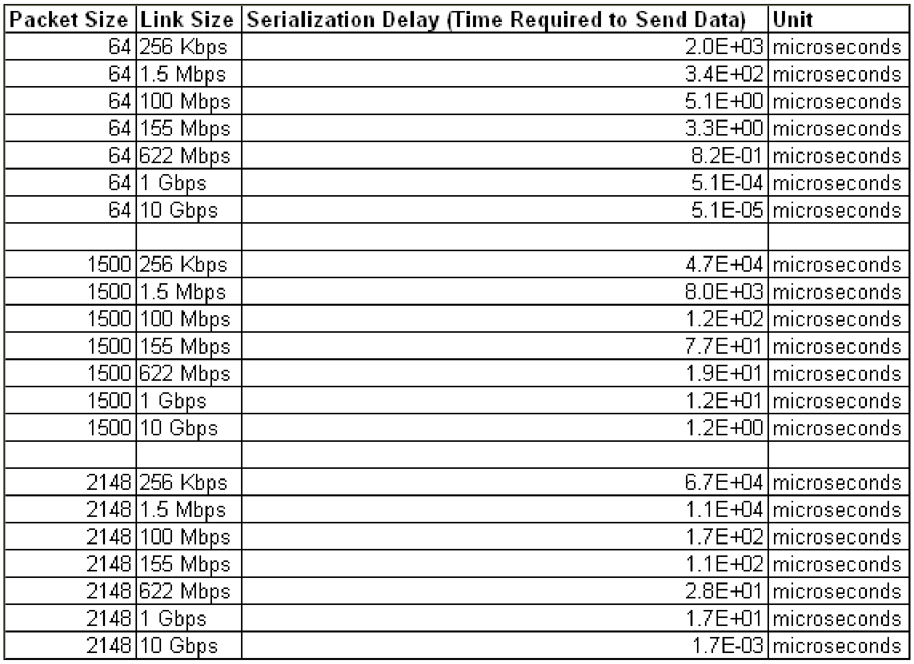
The link quality and bandwidth play a large role. Your network provider likely ensures a latency maximum on your network link. Therefore, be sure to stay as far beneath the Global Mirror round-trip-time (RTT) limit as possible. You can easily double or triple the expected physical latency with a lower quality or lower bandwidth network link. Then, you are within the range of exceeding the limit if high I/O occurs that exceeds the existing bandwidth capacity.

When you get a `1920` event, always check the latency first. The FCIP routing layer can introduce latency if it is not properly configured. If your network provider reports a much lower latency, you might have a problem at your FCIP routing layer. Most FCIP routing devices have built-in tools to enable you to check the RTT. When you are checking latency, remember that TCP/IP routing devices (including FCIP routers) report RTT using standard 64-byte ping packets.

In Figure 11-167, you can see why the effective transit time must be measured only by using packets that are large enough to hold an FC frame, or 2148 bytes (2112 bytes of payload and 36 bytes of header). Allow estimated resource requirements to be a safe amount, because various switch vendors have optional features that might increase this size. After you verify your latency by using the proper packet size, proceed with normal hardware troubleshooting.

Before we proceed, we look at the second largest component of your RTT, which is *serialization delay*. Serialization delay is the amount of time that is required to move a packet of data of a specific size across a network link of a certain bandwidth. The required time to move a specific amount of data decreases as the data transmission rate increases.

Figure 11-167 shows the orders of magnitude of difference between the link bandwidths. It is easy to see how 1920 errors can arise when your bandwidth is insufficient. Never use a TCP/IP ping to measure RTT for FCIP traffic.



| Packet Size | Link Size | Serialization Delay (Time Required to Send Data) | Unit |
|---|---|---|---|
| 64 | 256 Kbps | 2.0E+03 | microseconds |
| 64 | 1.5 Mbps | 3.4E+02 | microseconds |
| 64 | 100 Mbps | 5.1E+00 | microseconds |
| 64 | 155 Mbps | 3.3E+00 | microseconds |
| 64 | 622 Mbps | 8.2E-01 | microseconds |
| 64 | 1 Gbps | 5.1E-04 | microseconds |
| 64 | 10 Gbps | 5.1E-05 | microseconds |
| | | | |
| 1500 | 256 Kbps | 4.7E+04 | microseconds |
| 1500 | 1.5 Mbps | 8.0E+03 | microseconds |
| 1500 | 100 Mbps | 1.2E+02 | microseconds |
| 1500 | 155 Mbps | 7.7E+01 | microseconds |
| 1500 | 622 Mbps | 1.9E+01 | microseconds |
| 1500 | 1 Gbps | 1.2E+01 | microseconds |
| 1500 | 10 Gbps | 1.2E+00 | microseconds |
| | | | |
| 2148 | 256 Kbps | 6.7E+04 | microseconds |
| 2148 | 1.5 Mbps | 1.1E+04 | microseconds |
| 2148 | 100 Mbps | 1.7E+02 | microseconds |
| 2148 | 155 Mbps | 1.1E+02 | microseconds |
| 2148 | 622 Mbps | 2.8E+01 | microseconds |
| 2148 | 1 Gbps | 1.7E+01 | microseconds |
| 2148 | 10 Gbps | 1.7E-03 | microseconds |

*Figure 11-167 Effect of packet size (in bytes) versus the link size*

In Figure 11-167, the amount of time in microseconds that is required to transmit a packet across network links of varying bandwidth capacity is compared. The following packet sizes are used:

- 64 bytes: The size of the common ping packet
- 1500 bytes: The size of the standard TCP/IP packet
- 2148 bytes: The size of an FC frame

Finally, your path maximum transmission unit (MTU) affects the delay that is incurred to get a packet from one location to another location. An MTU might cause fragmentation or be too large and cause too many retransmits when a packet is lost.

## 11.12.2 1720 error

The 1720 error (event ID 050020) is the other problem remote copy might encounter. The amount of bandwidth that is needed for system-to-system communications varies based on the number of nodes. It is important that it is not zero. When a partner on either side stops communication, you see a 1720 appear in your error log. According to the product documentation, there are no likely field-replaceable unit breakages or other causes.

The source of this error is most often a fabric problem or a problem in the network path between your partners. When you receive this error, check your fabric configuration for zoning of more than one host bus adapter (HBA) port for each node per I/O Group if your fabric has more than 64 HBA ports zoned. One port for each node per I/O Group per fabric that is associated with the host is the suggested zoning configuration for fabrics.

For those fabrics with 64 or more host ports, this recommendation becomes a rule. Therefore, you see four paths to each volume discovered on the host because each host needs to have at least two FC ports from separate HBA cards, each in a separate fabric. On each fabric, each host FC port is zoned to two of node ports where each port comes from one node canister. This gives four paths per host volume. More than four paths per volume are supported but not recommended.

Improper zoning can lead to SAN congestion, which can inhibit remote link communication intermittently. Checking the zero buffer credit timer from IBM Virtual Storage Center and comparing against your sample interval reveals potential SAN congestion. If a zero buffer credit timer is above 2% of the total time of the sample interval, it might cause problems.

Next, always ask your network provider to check the status of the link. If the link is acceptable, watch for repeats of this error. It is possible in a normal and functional network setup to have occasional `1720` errors, but multiple occurrences could indicate a larger problem.

If you receive multiple `1720` errors, recheck your network connection and then check the system partnership information to verify its status and settings. Then, proceed to perform diagnostics for every piece of equipment in the path between the two Storwize systems. It often helps to have a diagram that shows the path of your replication from both logical and physical configuration viewpoints.

If your investigations fail to resolve your remote copy problems, contact your IBM Support representative for a more complete analysis.

# 12

# Encryption

Encryption protects against the potential exposure of sensitive user data that is stored on discarded, lost, or stolen storage devices. IBM Storwize V7000 Gen2/Gen2+ support optional encryption of data at rest.

Specifically, this chapter provides information about the following topics:

- Introducing encryption
- Defining encryption of data at rest
- Activating encryption
- Enabling encryption
- Using encryption
- Rekeying an encryption-enabled system
- Disabling encryption
- Restrictions

## 12.1 Introducing encryption

First, we define the encryption types:

- Hardware encryption: data is encrypted using Serial-attached SCSI (SAS) hardware for internal storage. After encryption is enabled, all internal storage array objects are created as hardware encrypted by default.
- Software encryption: data is encrypted using AES_NI CPU instruction set and engines on external, attached storage at the pool level.

**Note:** Software encryption is available in IBM Spectrum Virtualize code V7.6 and later.

Both methods of encryption protect against the potential exposure of sensitive user data and user metadata that are stored on discarded, lost, or stolen storage devices. Both methods of encryption can also facilitate the warranty return or disposal of hardware.

Both methods of encryption use the same encryption algorithm, the same key management. And, they use the same license.

Selection of the encryption method is fully transparent. The code uses hardware encryption for internal SAS-attached disks. External disks are encrypted by software. The encryption method cannot not be selected manually.

**Note:** The design for encryption is based on the concept that a system must either be encrypted or not encrypted. By implementing encryption, we chose not to encourage a solution that contained both encrypted volumes and unencrypted volumes. However, environments can exist where encryption is enabled with pre-existing unencrypted volumes. In this situation there are options to migrate unencrypted volumes to become encrypted volumes. Migration is required as there is no direct method of encrypting data on pre-existing unencrypted volumes.

## 12.2 Defining encryption of data at rest

*Encryption* is the process of encoding data so that only authorized parties can read it. Secret keys are used to encode the data according to well-known algorithms.

Encryption of data at rest is defined by the following characteristics:

- *"Data at rest"* means that the data is encrypted on the end device (drives).
- The algorithm that is used is the Advanced Encryption Standard (AES) US government standard from 2001.
- Encryption of data at rest complies with the Federal Information Processing Standard 140 (FIPS-140) standard, but it not certified.
- Ciphertext stealing (XTS)-AES 256 is used for data keys.
- AES 256 is used for master keys.
- The algorithm is public. The only secrets are the keys.
- The symmetric key algorithm is used. (The same key used to encrypt and decrypt data.)

The encryption of system data and the encryption of system metadata are not required, so system data and metadata are not encrypted.

Encryption is enabled at a system level and all of the following prerequisites must be met *before* you can use encryption:

- You must purchase an encryption license before you activate the function.

  If you did not purchase a license, contact an IBM marketing representative or IBM Business Partner to purchase an encryption license.
- At least three USB flash drives are required before encryption enablement. They are available as a feature code from IBM.
- You must activate the license that you purchased.
- Encryption must be enabled.

**Note:** Only data at rest is encrypted host to storage communication and Remote Mirroring are not encrypted.

Figure 12-1 shows an encryption example. Encrypted disks and encrypted data paths are marked in blue. Unencrypted disks and data paths are marked in red. The server sends unencrypted data to a Storwize V7000 Gen2 system, which stores hardware-encrypted data on internal disks. The data is mirrored to a remote Storwize V7000 Gen1 system by using Remote Copy. The data is not encrypted during replication. Because the Storwize V7000 Gen1 is unable to perform any encryption activities, data on the Storwize V7000 Gen1 is not encrypted.

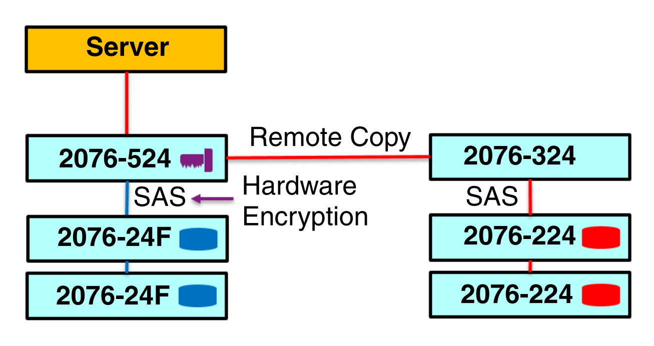


*Figure 12-1 Encryption on single site*

If both data copies must be encrypted, the Storwize V7000 Gen1 must be replaced by an encryption capable IBM Spectrum Virtualize system, as shown in Figure 12-2. Although both copies are encrypted, the Remote Copy communication between both sites is still not encrypted.

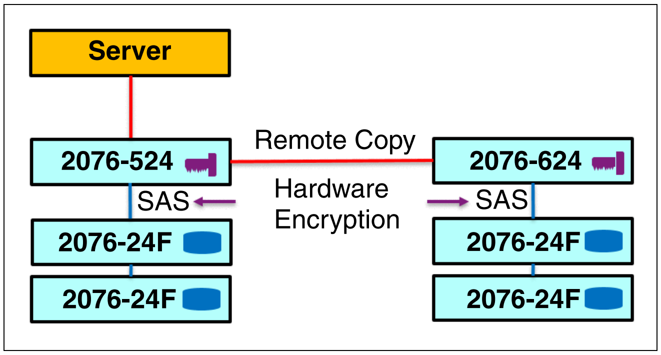


*Figure 12-2 Encryption on both sites*

Figure 12-3 shows an example for software and hardware encryption. Software encryption is used to encrypt an external virtualized storage system. Hardware encryption is used for internal, SAS-attached disk drives.

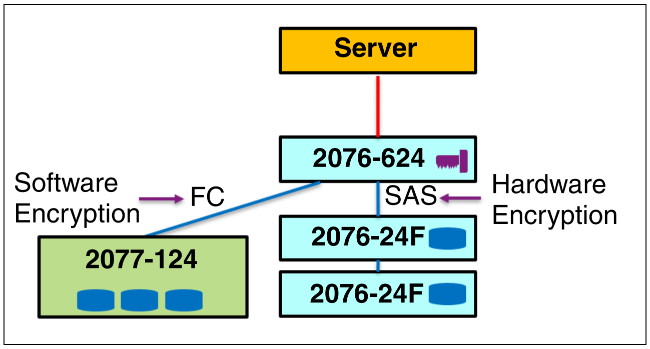


*Figure 12-3 Example of software encryption and hardware encryption*

Hardware encryption and software encryption in the Storwize code stack are shown in Figure 12-4. The functions that are implemented in software are shown in blue. The external storage system is shown in yellow. The hardware encryption on the SAS chip is marked in pink. Compression is performed “before” encryption. Therefore, both compression and encryption can be applied to the same data.

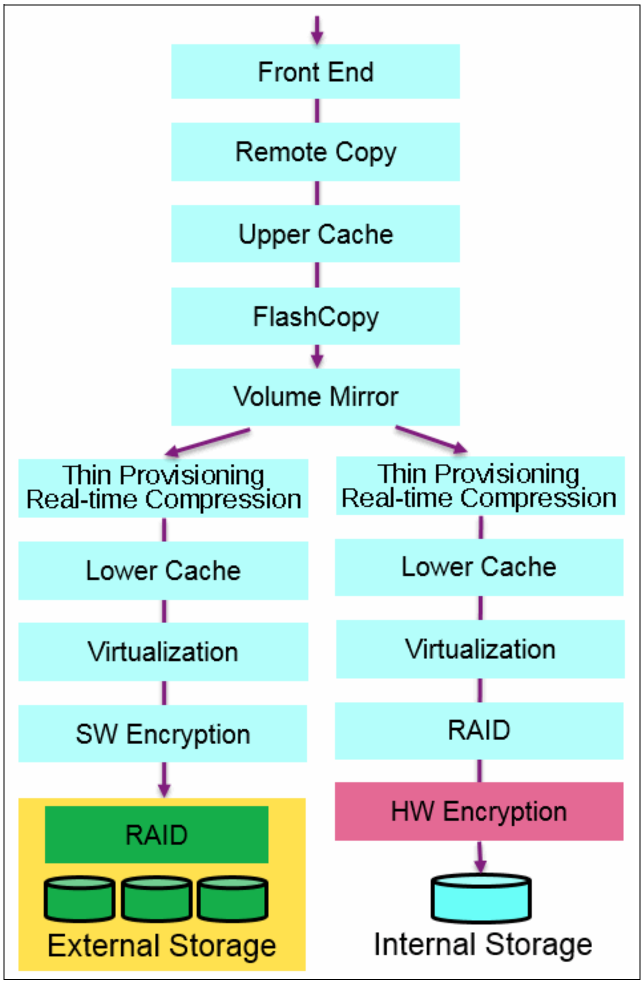


*Figure 12-4 Encryption in the software stack*

Each volume copy can use different encryption types (hardware, software, or no encryption). The encryption type depends on the pool that is used for the specific copy. You can migrate data between different encryption types by using volume migration or volume mirroring.

### 12.2.1 Encryption keys

Hardware and software encryption use the same encryption key technology. The only difference is the object that is encrypted by using the keys. The following objects can be encrypted:

- Arrays (hardware encryption)
- Pools (software encryption)
- Child pools (software encryption)
- Managed disks (software encryption)

Encryption keys can be described as follows:

- Keys are unique for each object, and they are created when the object is created. The creation of multiple keys is possible.
- Two types of keys are available:
  - Master access key (one for each system):
    - The master access key is created when encryption is enabled.
    - The master access key is stored on USB flash drives or a key server when encryption is enabled.
    - It can be copied or backed up as necessary.
    - It is *not* stored in the Cluster Systems Management (CSM) memory or any other non-volatile storage on the cluster.
  - Data encryption key (one for each encrypted object):
    - The data encryption key is used to encrypt data. It is created automatically when an encrypted object, such as an array, a pool, or a child pool, is created.
    - Not self-encrypting Managed disks (MDisk) are automatically encrypted using the data encryption key of the pool or child pool they belong to.
    - The data encryption key is stored in secure memory.
    - The data encryption key is encrypted with the master access key during cluster internal communication.
    - No way exists to view the data encryption key.
    - The data encryption key cannot be changed.
    - The data encryption key is discarded when the object is deleted (*secure erase*).

**Important:** If all master access key copies are lost and the system must cold reboot, all encrypted data is gone. No method exists, even for IBM, to decrypt the data without the keys. If encryption is enabled and the system cannot access the master access key, all SAS hardware is offline, including unencrypted arrays.

### 12.2.2 Encryption licenses

Encryption is a licensed feature that uses key-based licensing. A license must be present for each Storwize V7000 Gen2/Gen2+ control enclosure in the system before you enable encryption.

Attempts to add a control enclosure can fail if the correct license for the control enclosure that is being added does not exist. You can add licenses to the system for control enclosures that are not part of the system.

No trial licenses for encryption exist on the basis that when the trial runs out, the access to the data would be lost. Therefore, you must *purchase* an encryption license before you activate encryption. Licenses are generated by *IBM Data storage feature activation* (DSFA) based on the serial number (S/N) and the machine type and model number (MTM) of the control enclosures.

During the initial system setup, you are prompted to activate the license on the system. Or, after the initial setup, you can activate the license in the running environment. License activation is not enough to enable encryption on the system. Additional steps and three USB flash drives are required to enable encryption.

Contact your IBM marketing representative or IBM Business Partner to purchase an encryption license.

# 12.3 Activating encryption

The first step in order to use encryption is to activate your license.

Activation of the license can be performed in one of two ways, either *automatically* or *manually*. Both methods for activating the license can be started during the *initial system setup* or while the *system is running*.

## 12.3.1 Obtaining an encryption license

*You must purchase an encryption license before you activate encryption.* If you did not purchase a license, contact an IBM marketing representative or IBM Business Partner to purchase an Encryption license.

When you purchase a license you should receive a function authorization document with an *authorization code* printed on it. This code is enough to carry on with the automatic activation process.

If the automatic activation process fails or if you prefer using the manual activation process, use this page to retrieve your license keys:

https://www.ibm.com/storage/dsfa/storwize/selectMachine.wss

Ensure that the following information is available:

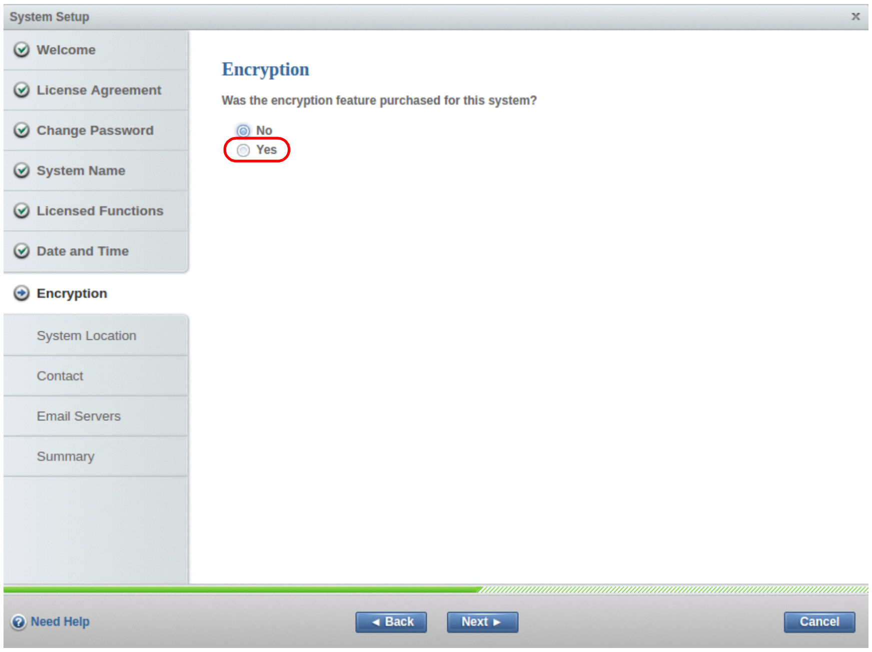
- Machine type (MT)
- Serial number (S/N)
- Machine signature
- Authorization code

See “Activate the license manually” on page 575 for instructions on how to retrieve the *machine signature* of a control enclosure.

## 12.3.2 Start activation process during initial system setup

One of the steps in the initial setup relates to the encryption license activation. The system asks “`Was the encryption feature purchased for this system?`”. You reply with either **No** or **Yes**. To activate encryption at this stage follow these steps:

1. Select **Yes**, as shown in Figure 12-5.



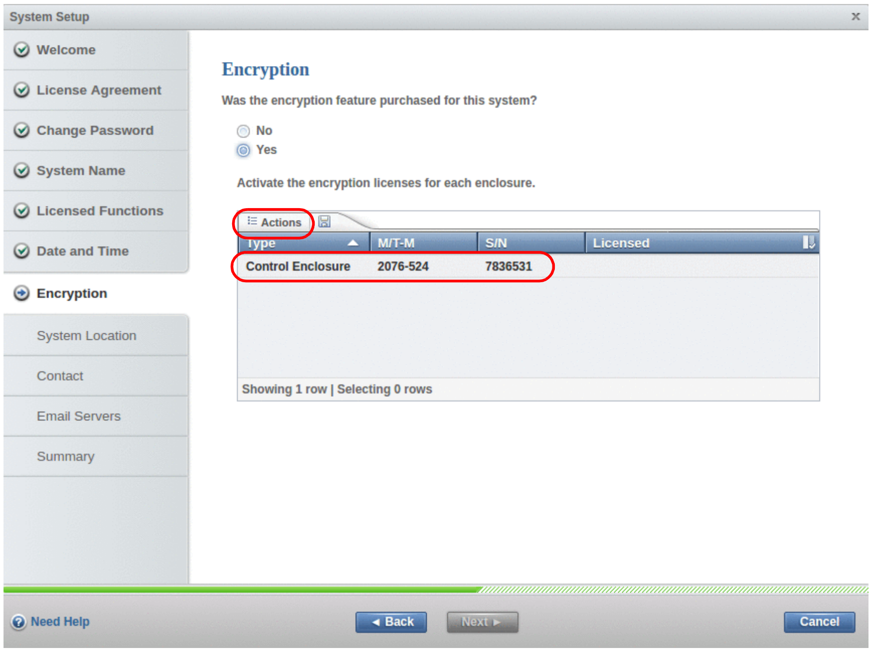
Encryption

Was the encryption feature purchased for this system?

● No
○ Yes

*Figure 12-5 Encryption activation during initial system setup*

2. The Encryption panel displays information about your local machine. Select the `Control Enclosure` and click on **Actions**, as shown in Figure 12-6.



Encryption

Was the encryption feature purchased for this system?

○ No
● Yes

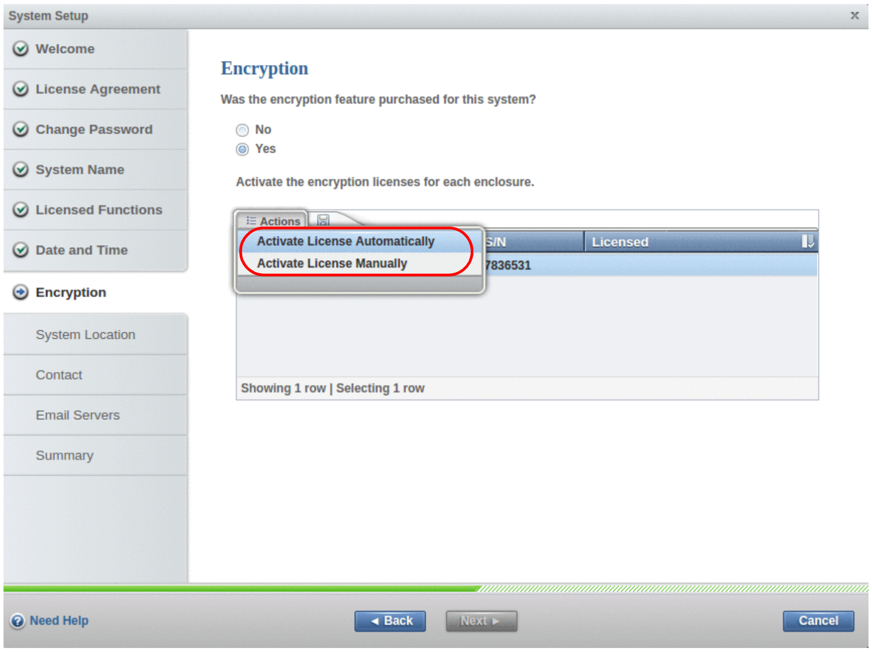
Activate the encryption licenses for each enclosure.

| Type | M/T-M | S/N | Licensed |
|---|---|---|---|
| Control Enclosure | 2076-524 | 7836531 | |

Showing 1 row | Selecting 0 rows

*Figure 12-6 Information panel of the local machine during initial system setup*

3. The Actions menu tab contains two activation options (`Activate license Automatically` and `Activate License Manually`), as shown in Figure 12-7. Use either options to activate encryption. See “Activate the license automatically” on page 572 for instructions on how to complete an automatic activation process. See “Activate the license manually” on page 575 for instructions on how to complete a manual activation process.



Encryption

Was the encryption feature purchased for this system?

○ No
● Yes

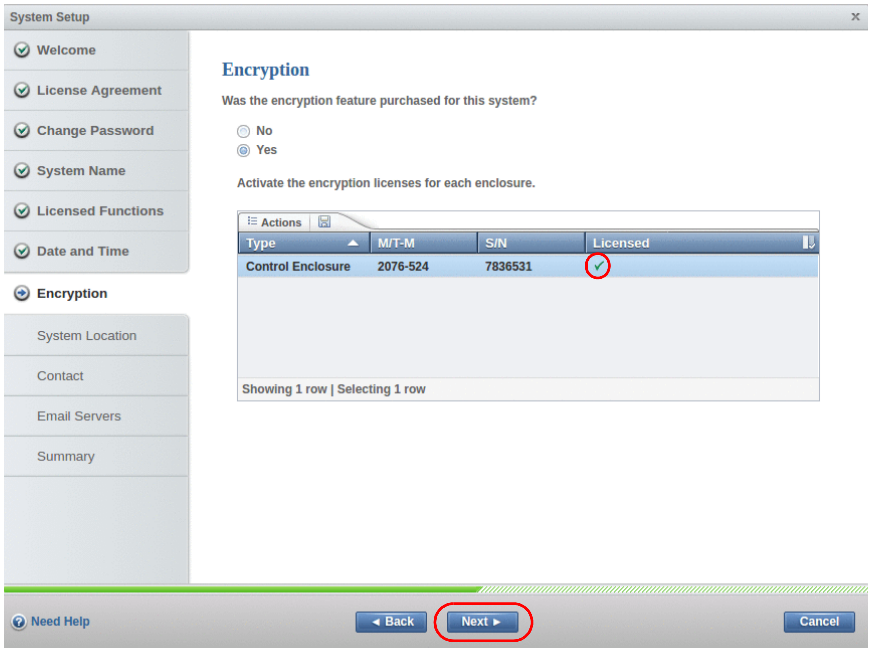
Activate the encryption licenses for each enclosure.

| S/N | Licensed |
|---|---|
| 7836531 | |

Showing 1 row | Selecting 1 row

*Figure 12-7 Selecting activation process during initial system setup*

4. Once either activation process is complete, you can see a green check mark under the heading that is labeled `Licensed` and you can proceed with the initial system setup using **Next**, as shown in Figure 12-8.



Encryption

Was the encryption feature purchased for this system?

○ No
● Yes

Activate the encryption licenses for each enclosure.

| Type | M/T-M | S/N | Licensed |
|---|---|---|---|
| Control Enclosure | 2076-524 | 7836531 | ✓ |

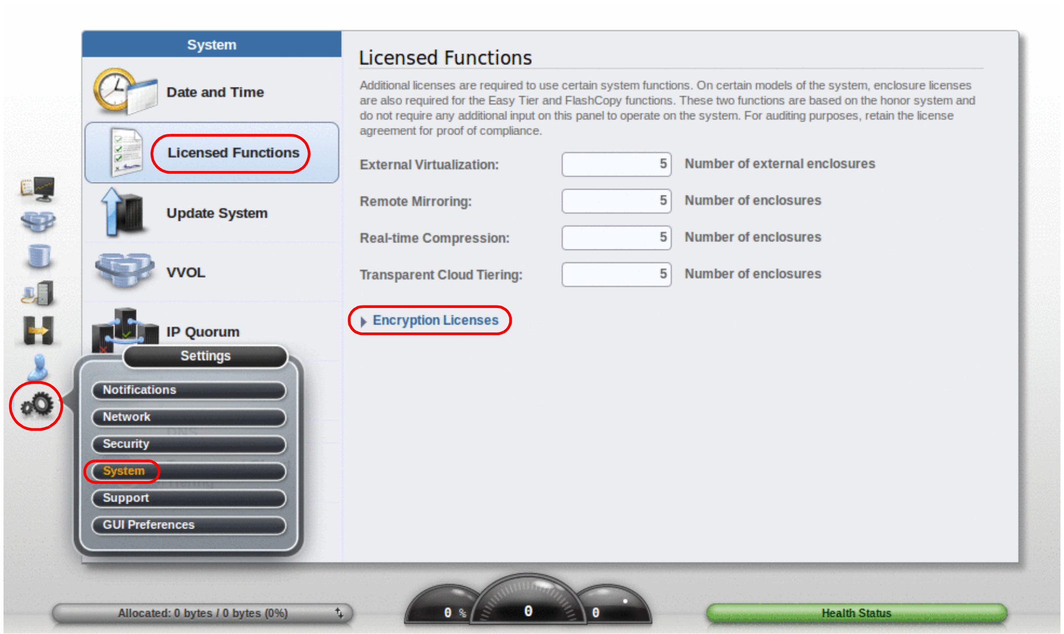
Showing 1 row | Selecting 1 row

*Figure 12-8 Successful encryption license activation during initial system setup*

## 12.3.3 Start activation process on a running system

On a running system, encryption license activation is available among the system setting menus. To activate encryption at this stage follow these steps:

1. Navigate to **Settings** → **System** → **Licensed Functions** and click on **Encryption Licenses**, as shown in Figure 12-9.



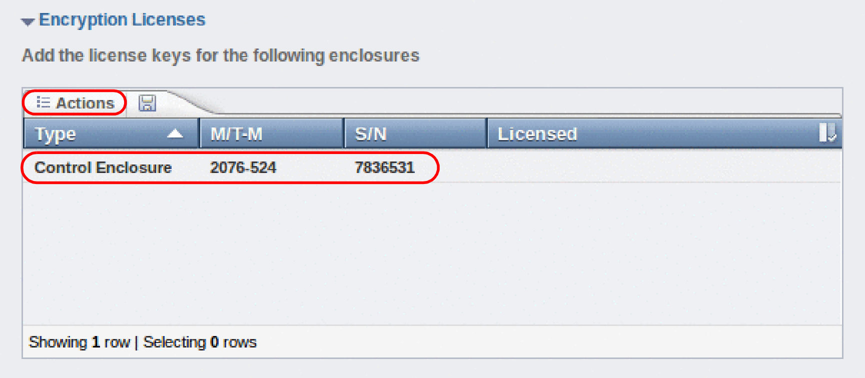
Licensed Functions

Additional licenses are required to use certain system functions. On certain models of the system, enclosure licenses are also required for the Easy Tier and FlashCopy functions. These two functions are based on the honor system and do not require any additional input on this panel to operate on the system. For auditing purposes, retain the license agreement for proof of compliance.

| | | |
|---|---|---|
| External Virtualization: | 5 | Number of external enclosures |
| Remote Mirroring: | 5 | Number of enclosures |
| Real-time Compression: | 5 | Number of enclosures |
| Transparent Cloud Tiering: | 5 | Number of enclosures |

*Figure 12-9 Selecting Encryption Licenses on the System Licensed Functions page*

2. The Encryption Licenses box displays information about your control enclosures. Select the `Control Enclosure` you want to install an encryption license for and click on **Actions**, as shown in Figure 12-10.



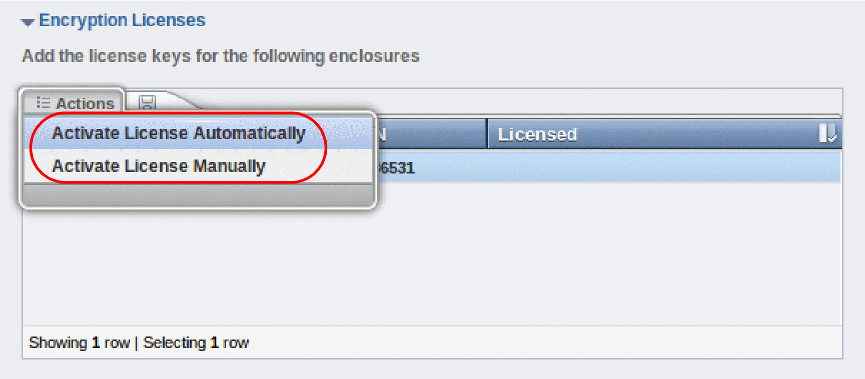
Add the license keys for the following enclosures

| Type | M/T-M | S/N | Licensed |
|---|---|---|---|
| Control Enclosure | 2076-524 | 7836531 | |

Showing **1** row | Selecting **0** rows

*Figure 12-10 Select the enclosure where you want to enable the encryption*

3. The Actions menu tab contains two activation options (`Activate license Automatically` and `Activate License Manually`), as shown in Figure 12-11. Use either options to activate encryption. See “Activate the license automatically” on page 572 for instructions on how to complete an automatic activation process. See “Activate the license manually” on page 575 for instructions on how to complete a manual activation process.

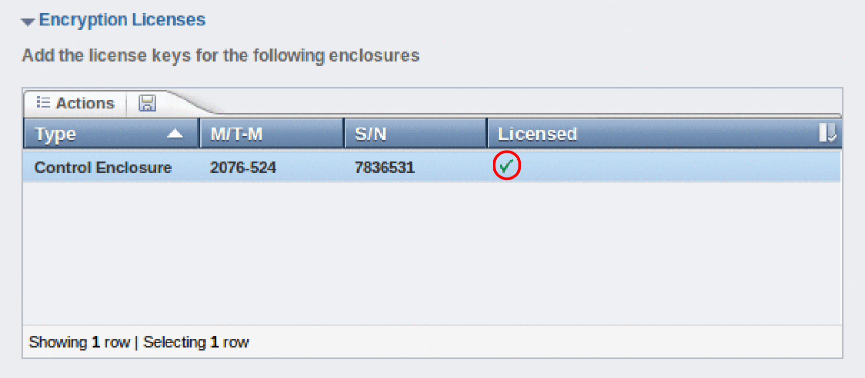


Add the license keys for the following enclosures

Showing 1 row | Selecting 1 row

*Figure 12-11 Activate the encryption license manually*

4. Once either activation process is complete, you can see a green check mark under the heading that is labeled `Licensed` for the Control Enclosure, as shown in Figure 12-12.



Add the license keys for the following enclosures

| Type | M/T-M | S/N | Licensed |
|---|---|---|---|
| Control Enclosure | 2076-524 | 7836531 | ✓ |

Showing 1 row | Selecting 1 row

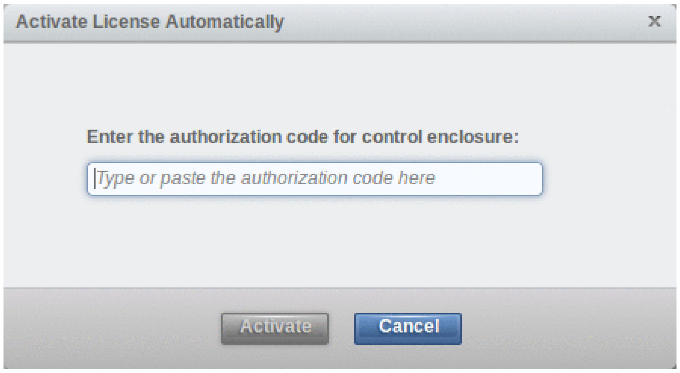
*Figure 12-12 Successful encryption license activation on a running system*

## 12.3.4 Activate the license automatically

> **Important:** To perform this operation, the personal computer that is used to connect to the GUI and activate the license must connect to an external network.

To activate the encryption license for a control enclosure automatically follow this procedure:

1. If you selected `Activate License Automatically` during initial system setup or on a running system the **Activate License Automatically** window opens, as shown in Figure 12-13.



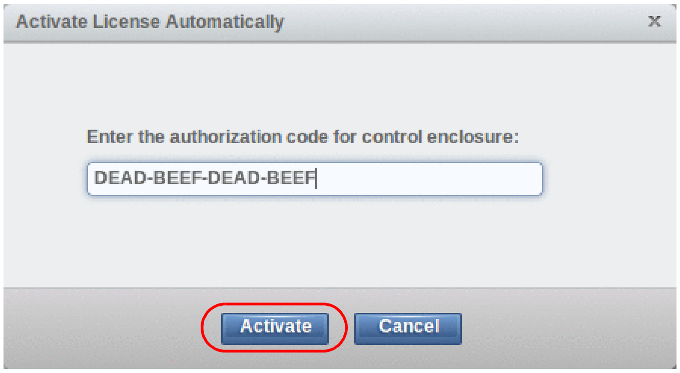
Activate License Automatically

Enter the authorization code for control enclosure:

Type or paste the authorization code here

*Figure 12-13 Encryption license activate license automatically panel*

2. Enter the authorization code that is specific to the enclosure that you selected, as shown in Figure 12-14. You can now click **Activate**.



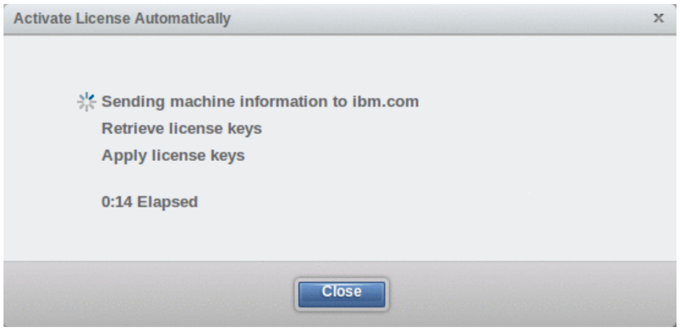
Activate License Automatically

Enter the authorization code for control enclosure:

DEAD-BEEF-DEAD-BEEF

*Figure 12-14 Entering an authorization code*

3. The system connects to IBM to verify the authorization code and retrieve the license key. Figure 12-15 shows this connection. If everything works correctly, the procedure takes less than a minute.



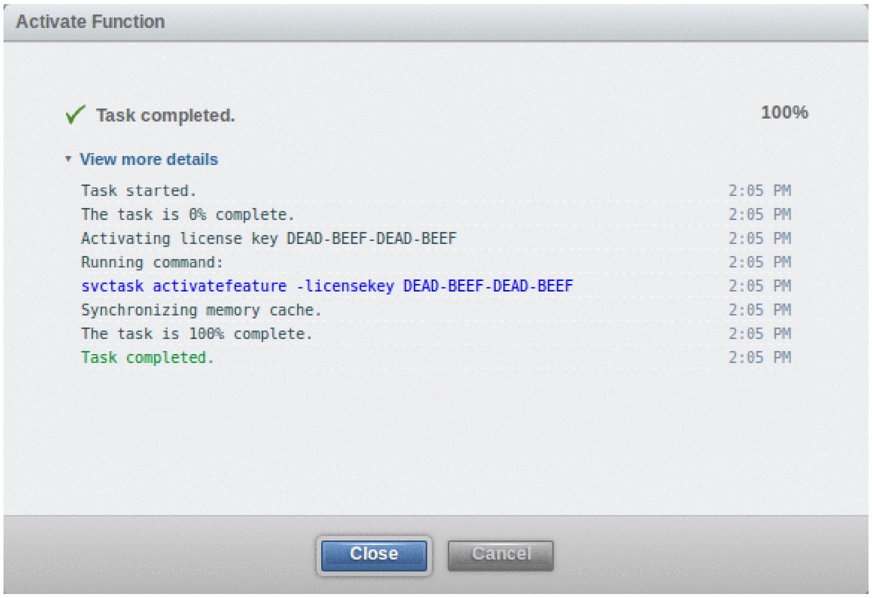
Activate License Automatically

Sending machine information to ibm.com
Retrieve license keys
Apply license keys

0:14 Elapsed

*Figure 12-15 Activating encryption*

4. After the license key has been retrieved it is automatically applied, as shown in Figure 12-16. The window closes automatically if it has been successful.



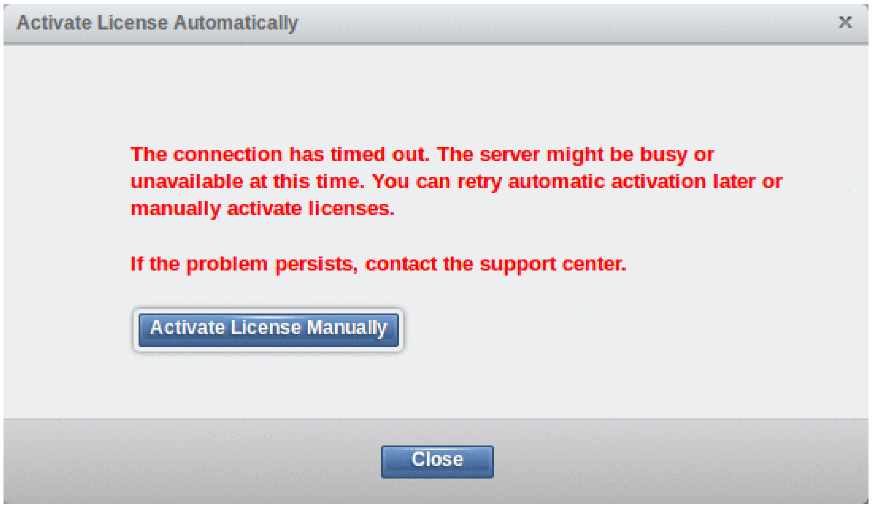
```
Task started.                                      2:05 PM
The task is 0% complete.                           2:05 PM
Activating license key DEAD-BEEF-DEAD-BEEF         2:05 PM
Running command:                                   2:05 PM
svctask activatefeature -licensekey DEAD-BEEF-DEAD-BEEF   2:05 PM
Synchronizing memory cache.                        2:05 PM
The task is 100% complete.                         2:05 PM
Task completed.                                    2:05 PM
```

*Figure 12-16 Successful encryption license automatic activation*

## Problems with automatic license activation

If connections problems occur with the automatic license activation procedure, the system times out after 3 minutes with an error, as shown in Figure 12-17.

Check whether the personal computer that is used to connect to the GUI and activate the license can access the internet. If this activation procedure still does not work, try to use the manual activation procedure described at “Activate the license manually” on page 575.

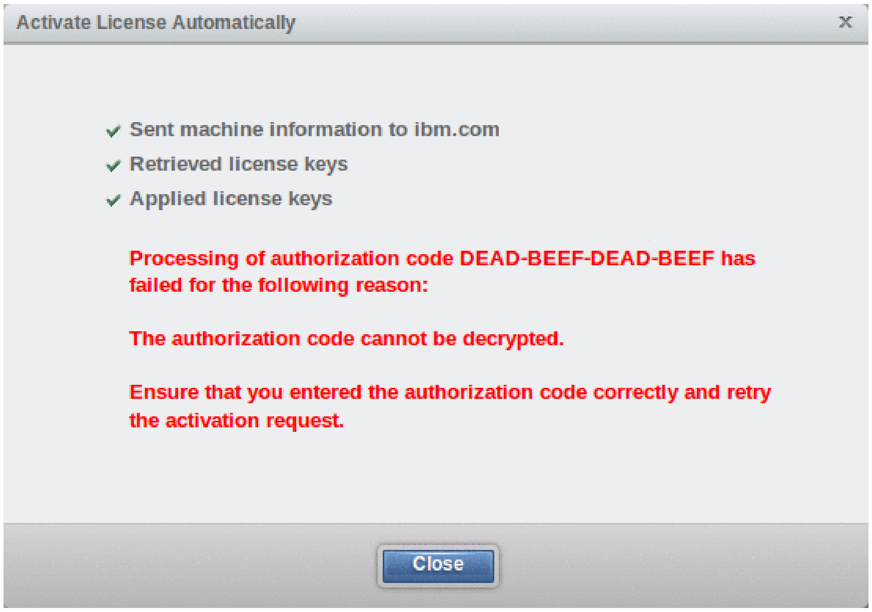


The connection has timed out. The server might be busy or unavailable at this time. You can retry automatic activation later or manually activate licenses.

If the problem persists, contact the support center.

*Figure 12-17 Automatic license activation time out*

Although authorization codes and encryption license keys use the same 4 x 4 hexadecimal digit format, you can only use each in the appropriate activation process. If you use a license key when the system expects an authorization code, the error message that is shown in Figure 12-18 appears.



Activate License Automatically

✓ Sent machine information to ibm.com

✓ Retrieved license keys

✓ Applied license keys

Processing of authorization code DEAD-BEEF-DEAD-BEEF has failed for the following reason:

The authorization code cannot be decrypted.

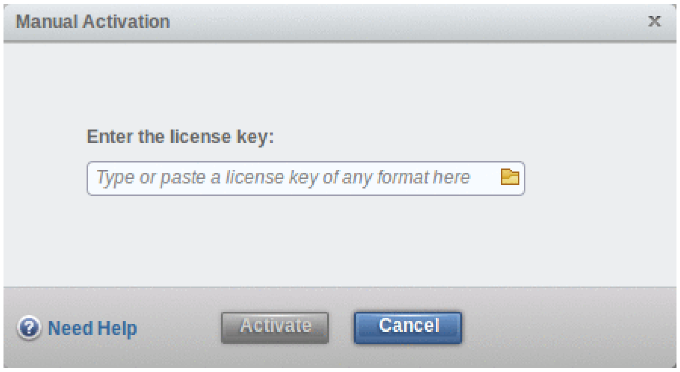
Ensure that you entered the authorization code correctly and retry the activation request.

*Figure 12-18 Authorization code failure*

## 12.3.5 Activate the license manually

To activate the encryption license for a control enclosure manually follow this procedure:

1. If you selected `Activate License Manually` during initial system setup or on a running system the **Manual Activation** window opens, as shown in Figure 12-19.



Manual Activation

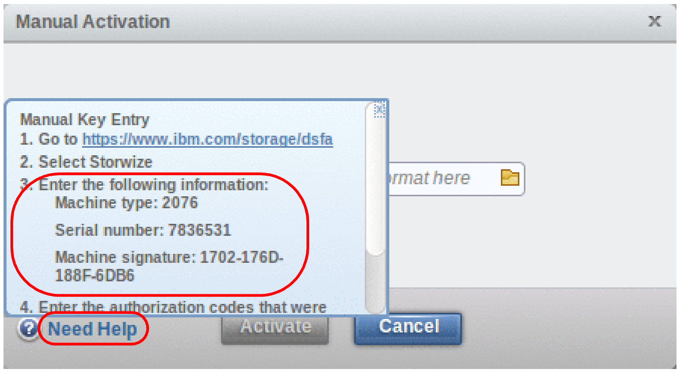
Enter the license key:

*Type or paste a license key of any format here*

*Figure 12-19 Encryption license manual activation panel*

2. If you have not done so already, you need to obtain the encryption license for the control enclosure you are trying to activate. Click **Need Help** to show the information you will need to follow the instructions in “Obtaining an encryption license” on page 568, as shown in Figure 12-20:

   – Machine type (MT)
   – Serial number (S/N)
   – Machine signature

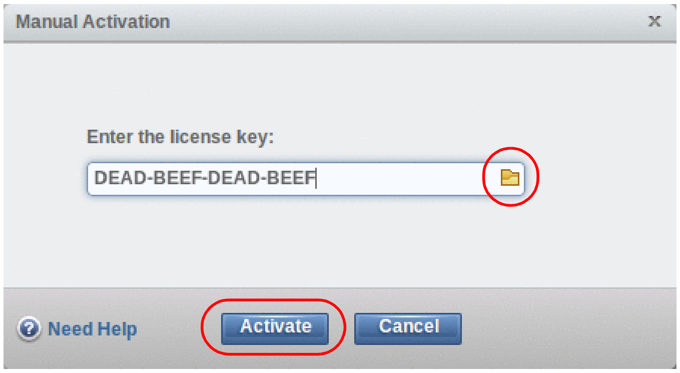


Manual Key Entry
1. Go to https://www.ibm.com/storage/dsfa
2. Select Storwize
3. Enter the following information:
Machine type: 2076
Serial number: 7836531
Machine signature: 1702-176D-188F-6DB6
4. Enter the authorization codes that were

*Figure 12-20 Information to obtain an encryption license*

3. You can enter the license key either by typing it, by using cut or copy and paste, or by selecting a license key file downloaded from DSFA after clicking on the folder icon. In Figure 12-21, the sample key is already entered. You can now click **Activate**.

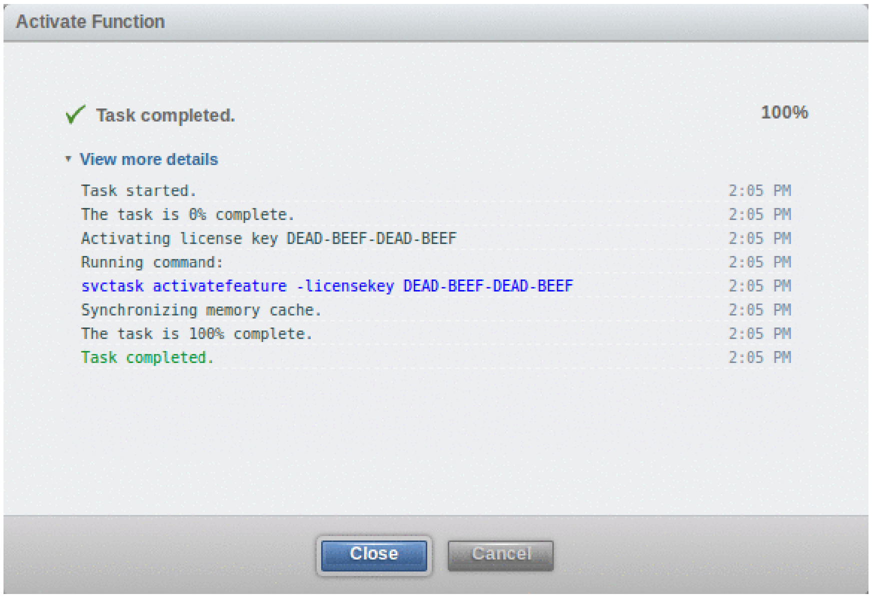
> **Note:** If using a license key file, the file can only contain the license key for a single control enclosure.



Enter the license key: DEAD-BEEF-DEAD-BEEF

*Figure 12-21 Entering an encryption license key*

4. Figure 12-22 shows the completion of the manual activation process.

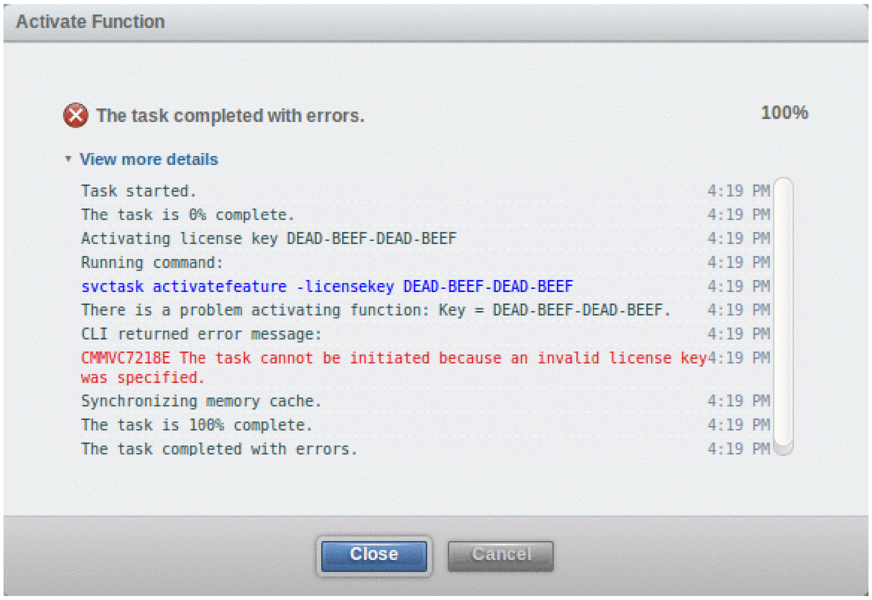


```
Task started.                                              2:05 PM
The task is 0% complete.                                   2:05 PM
Activating license key DEAD-BEEF-DEAD-BEEF                 2:05 PM
Running command:                                           2:05 PM
svctask activatefeature -licensekey DEAD-BEEF-DEAD-BEEF    2:05 PM
Synchronizing memory cache.                                2:05 PM
The task is 100% complete.                                 2:05 PM
Task completed.                                            2:05 PM
```

*Figure 12-22 Successful encryption license manual activation*

### Problems with manual license activation

Although authorization codes and encryption license keys use the same 4 x 4 hexadecimal digit format, you can only use each in the appropriate activation process. If you use an authorization code when the system expects a license key, the error message that is shown in Figure 12-23 appears.



```
Task started.                                                          4:19 PM
The task is 0% complete.                                               4:19 PM
Activating license key DEAD-BEEF-DEAD-BEEF                             4:19 PM
Running command:                                                       4:19 PM
svctask activatefeature -licensekey DEAD-BEEF-DEAD-BEEF                4:19 PM
There is a problem activating function: Key = DEAD-BEEF-DEAD-BEEF.     4:19 PM
CLI returned error message:                                            4:19 PM
CMMVC7218E The task cannot be initiated because an invalid license key4:19 PM
was specified.
Synchronizing memory cache.                                            4:19 PM
The task is 100% complete.                                             4:19 PM
The task completed with errors.                                        4:19 PM
```

*Figure 12-23 License key failure*

## 12.4 Enabling encryption

In this section, we describe the process to create and store system master access key copies also referred to as encryption keys. These keys can be stored on USB flash drives or a key server.

**Note:** Key server support is available in IBM Spectrum Virtualize code V7.8 and later.

**Important:** The encryption keys cannot be stored on both USB flash drives and a key server. Only one method can be used. To move from one method to the other all encrypted volumes must be first unencrypted. Therefore, it is crucial to make the correct choice when planning to use encryption.

This might change in future releases, please check the IBM Knowledge center for your current level of code at:

http://www.ibm.com/support/knowledgecenter/ST3FR7

The following list of key server and USB flash drive characteristics might help you to choose the type of encryption enablement that you want to use.

Key servers can have the following characteristics:

- Physical access to the system is not required to process a rekeying operation.
- Support for businesses that have security requirements not to use USB ports.
- Strong key generation.
- Key self-replication and automatic backups.
- Implementations follow an open standard that aids in interoperability.
- Audit detail.
- Ability to administer access to data separately from storage devices.

USB flash drives have the following characteristics:

- Physical access to the system may be required to process a rekeying operation.
- No moving parts with almost no read operations or write operations to the USB flash drive.
- Inexpensive to maintain and use.
- Convenient and easy to have multiple identical USB flash drives available as backups.

## 12.4.1 Starting the encryption enablement process

If the license activation passes all of the required steps, you can now enable encryption. Encryption enablement is performed outside of the initial system setup by using the **Enable Encryption** wizard, which is available in the GUI. The same steps can be performed by using the CLI.

The wizard can be accessed clicking **Enable Encryption** on the **Suggested task** window which appears at log in after the successful activation of encryption as shown in Figure 12-24.

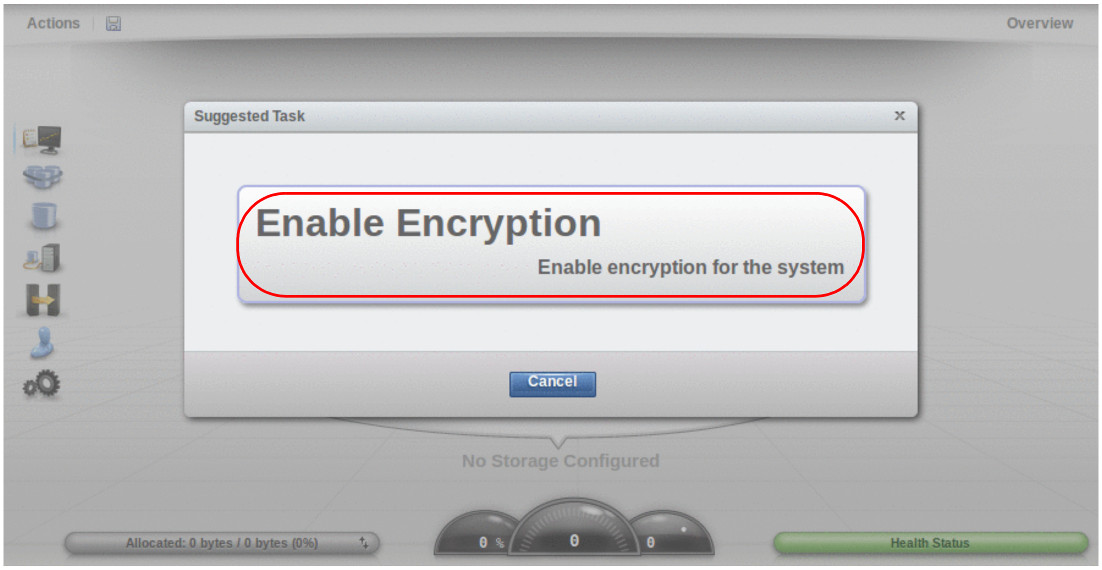


*Figure 12-24 Enable Encryption from the suggested task window*

You can also navigate to **Settings** → **Security** → **Encryption** and click **Enable Encryption**, as shown in Figure 12-25.

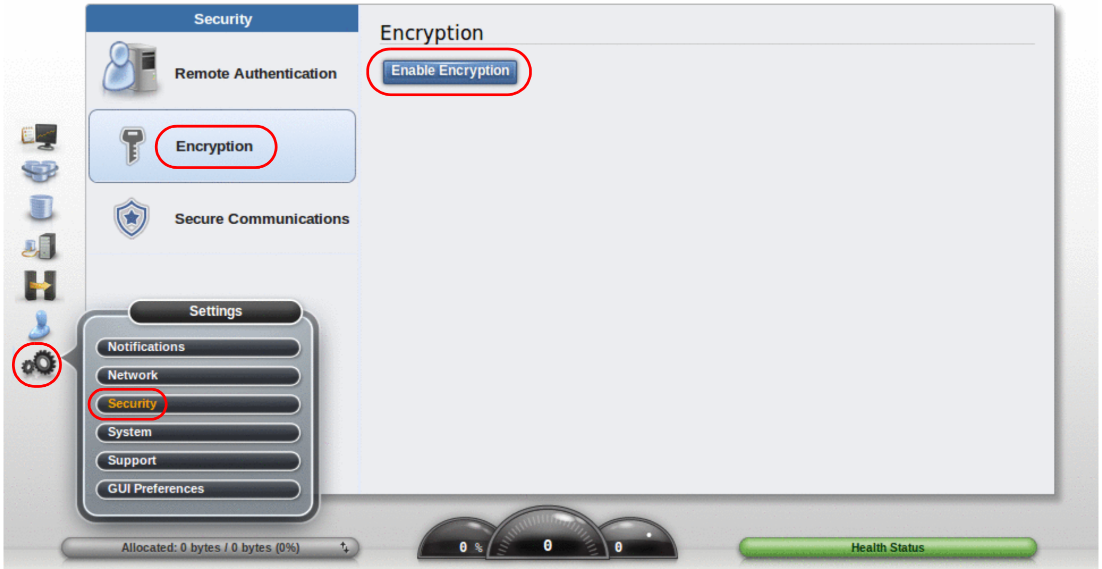
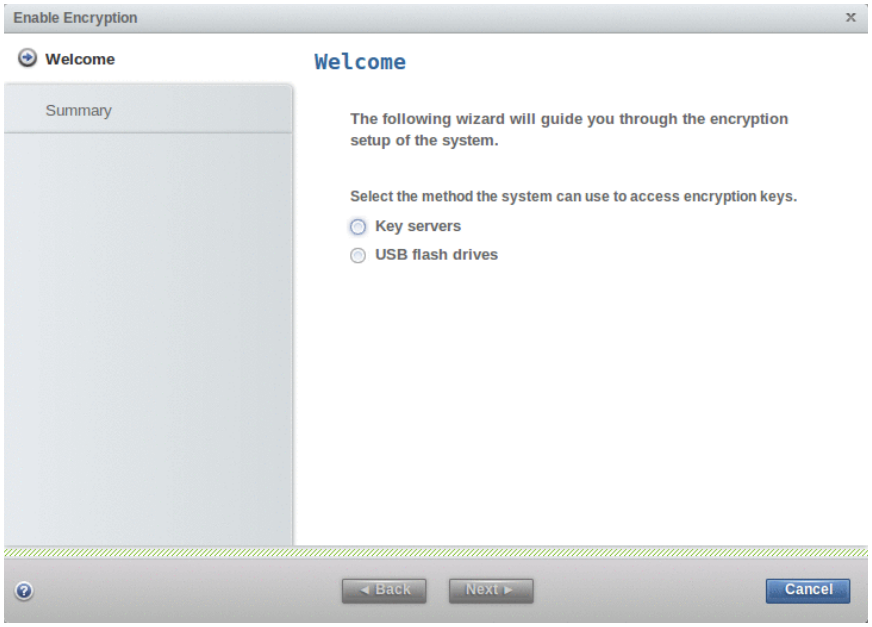


*Figure 12-25 Enable Encryption from the Security panel*

The **Enable Encryption** wizard will start by asking which method to use for storing the encryption keys, as shown in Figure 12-26.



Welcome

The following wizard will guide you through the encryption setup of the system.

Select the method the system can use to access encryption keys.

○ Key servers

○ USB flash drives

*Figure 12-26 Enable Encryption welcome message*

See “Enabling encryption using a key server” on page 580 for instructions on how to enable encryption by storing encryption keys on a key server. See “Enabling encryption using USB flash drives” on page 585 for instructions on how to enable encryption by storing encryption keys on USB flash drives.

## 12.4.2 Enabling encryption using a key server

A key server is a centralized server or application that receives and then distributes encryption keys to the Spectrum Virtualize system. The Spectrum Virtualize system must be on the same network as the key server.

Spectrum Virtualize supports enabling encryption using an IBM Security Key Lifecycle Manager (SKLM) key server. Before you can create the key server object, the key server must be configured. SKLM supports Key Management Interoperability Protocol (KMIP), which is a standard for encryption of stored data and management of cryptographic keys.

> **Important:** At the time of writing Spectrum Virtualize only supports a single key server object. This might change in future releases, please check the IBM Knowledge center for your current level of code at:
>
> http://www.ibm.com/support/knowledgecenter/ST3FR7

Ensure that you complete the following tasks on the SKLM server before you enable encryption:

- Configure the SKLM server to use Transport Layer Security version 2 (TLSv2). The default setting is TLSv1, but Spectrum Virtualize supports only version 2.
- Ensure that the database service is started automatically on startup.
- Ensure there is at least one Secure Sockets Layer (SSL) certificate for browser access, if not create one.

- Create a SPECTRUM_VIRT device group for Spectrum Virtualize systems. A device group allows for restricted management of subsets of devices within a larger pool.

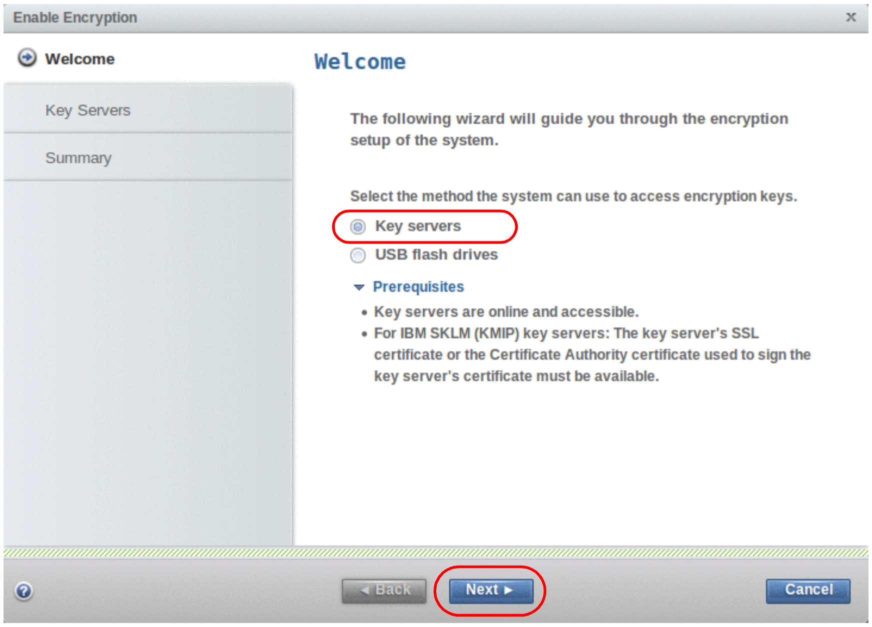
For more information about completing these tasks, see the SKLM Knowledge Center at:

https://www.ibm.com/support/knowledgecenter/SSWPVP

*Access to the key server storing the correct master access key is required to enable encryption for the cluster after a system restart* such as a system wide reboot or power loss. Access to the key server is not required during a warm reboot, such as a node exiting service mode or a single node reboot. The data center power-on procedure needs to include instructions to ensure connectivity between the system and the key server.

To enable encryption using a key server follow these steps:

1. In the **Enable Encryption** wizard **Welcome** tab select **Key servers** and click **Next**, as shown in Figure 12-27. For the **Key servers** option to be selectable you need to have *service IPs* configured on all your nodes.



Welcome

The following wizard will guide you through the encryption setup of the system.

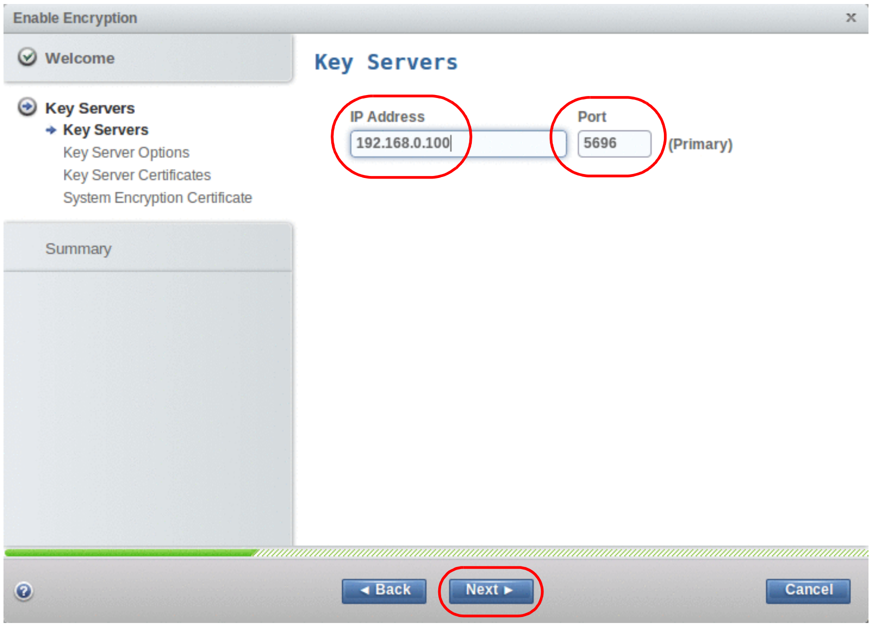
Select the method the system can use to access encryption keys.

● Key servers
○ USB flash drives

Prerequisites
- Key servers are online and accessible.
- For IBM SKLM (KMIP) key servers: The key server's SSL certificate or the Certificate Authority certificate used to sign the key server's certificate must be available.

*Figure 12-27 Selecting Key servers in the Enable Encryption wizard*

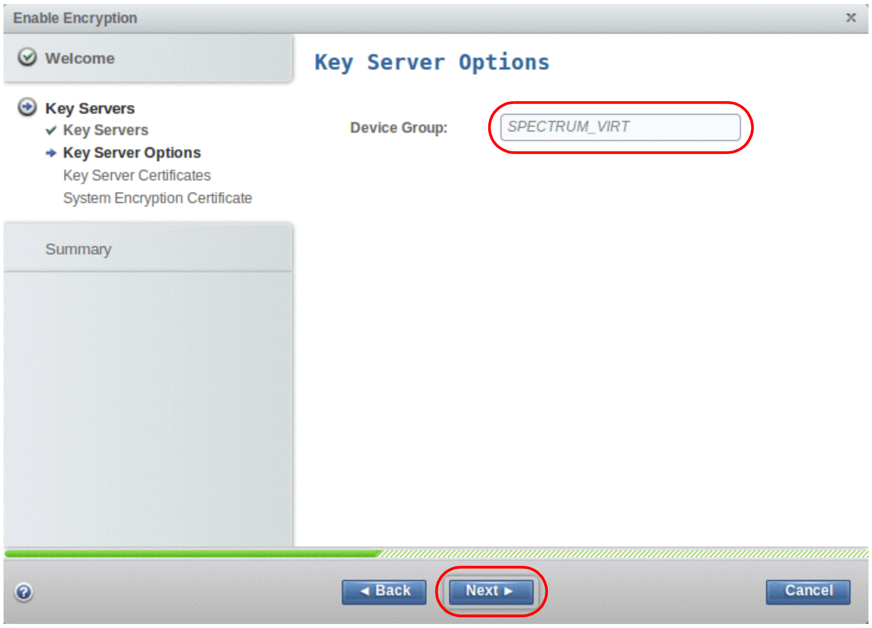
2. The wizard moves to the **Key Servers** tab, as shown in Figure 12-28. Type the **IP address** of the key server. Type the **Port** number for the KMIP protocol, 5696 is the default port. Click **Next** when you are done.



IP Address: 192.168.0.100
Port: 5696 (Primary)

Figure 12-28 Configuring key server IP address and port number

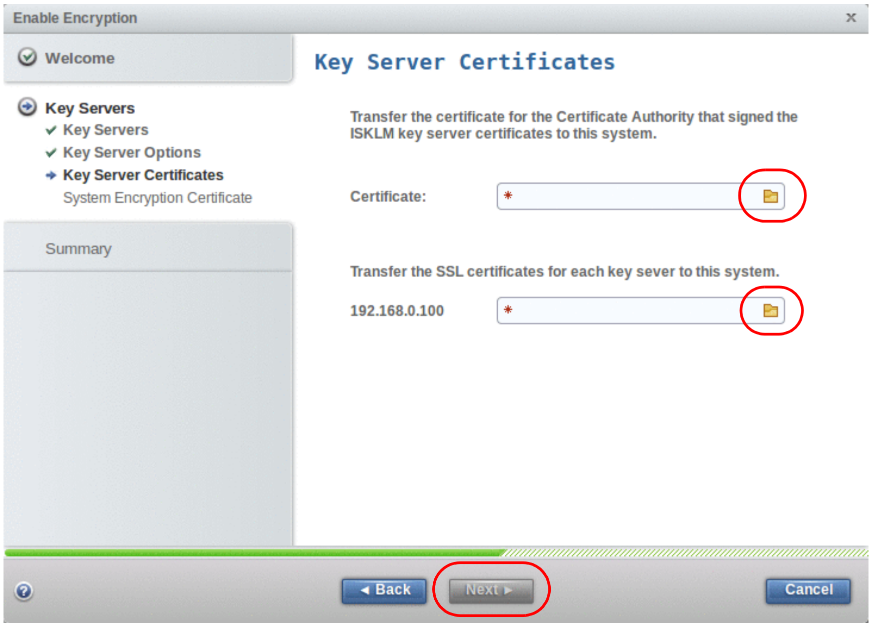
3. The next page in the wizard is a reminder that you should have created a `SPECTRUM_VIRT` device group for Spectrum Virtualize systems on the SKLM key server. Click **Next** to continue, as shown in Figure 12-29.



Device Group: SPECTRUM_VIRT

Figure 12-29 Checking key server device group

4. You should now configure the Spectrum Virtualize system to trust the SKLM key server SSL certificate. This can done by either uploading the Certificate Authority used to signed the SKLM key server certificate or by uploading the SSL certificate of the key server directly. Both options are shown in Figure 12-30. When either file has been selected you will be allowed to click **Next**.



Transfer the certificate for the Certificate Authority that signed the ISKLM key server certificates to this system.

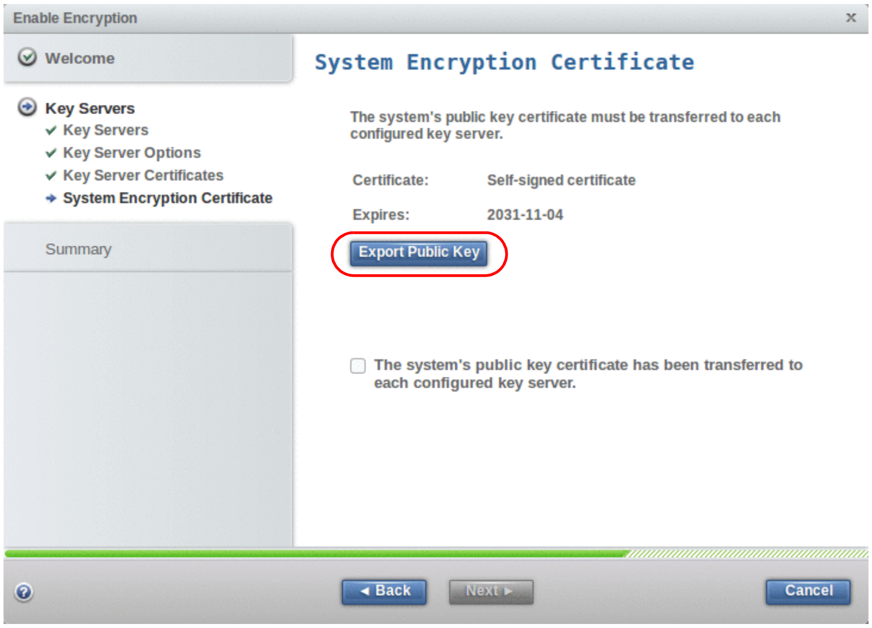
Certificate: *

Transfer the SSL certificates for each key sever to this system.

192.168.0.100 *

*Figure 12-30 Uploading the key server SSL certificate or Certificate Authority*

5. You should now configure the SKLM key server to trust the SSL certificate of the Spectrum Virtualize system. You can Download the Spectrum Virtualize system SSL certificate by clicking **Export Public Key**, as shown in Figure 12-31. You should install this certificate in the SKLM key server under the `SPECTRUM_VIRT` device group.



The system's public key certificate must be transferred to each configured key server.

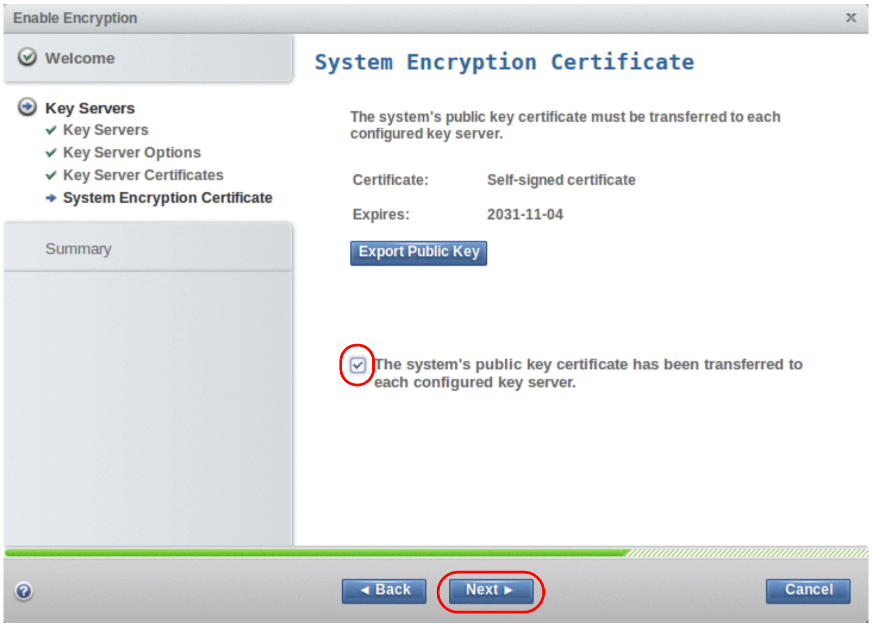
Certificate: Self-signed certificate

Expires: 2031-11-04

□ The system's public key certificate has been transferred to each configured key server.

*Figure 12-31 Downloading the Spectrum Virtualize SSL certificate*

6. When the Spectrum Virtualize system SSL certificate has been installed on the SKLM key server, acknowledge this by ticking the box indicated in Figure 12-32 and click **Next**.



System Encryption Certificate

The system's public key certificate must be transferred to each configured key server.

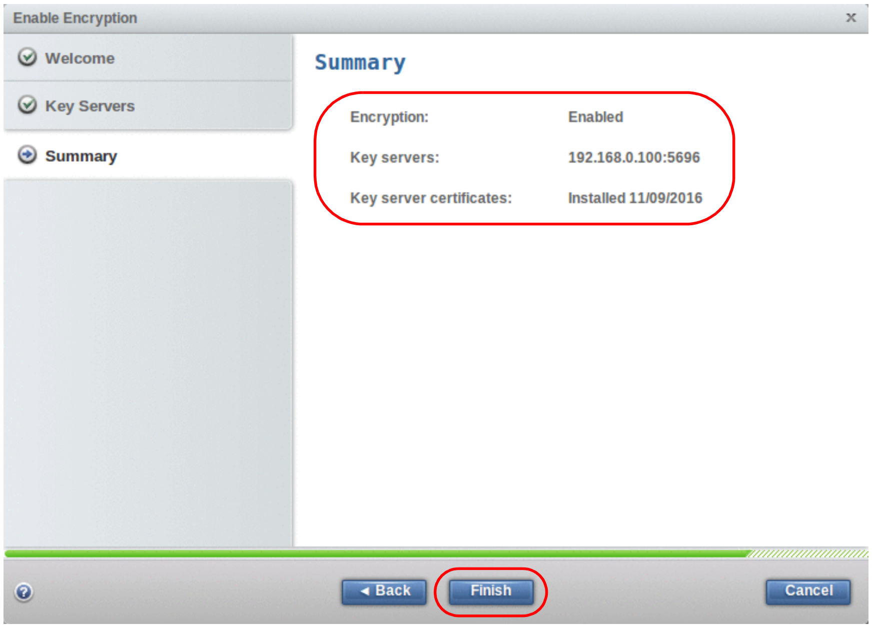
Certificate: Self-signed certificate

Expires: 2031-11-04

☑ The system's public key certificate has been transferred to each configured key server.

*Figure 12-32 Acknowledge Spectrum Virtualize SSL certificate transfer*

7. The key server configuration is shown in the **Summary** tab, as shown in Figure 12-33. Click **Finish** to create the key server object and finalize the encryption enablement.



Summary

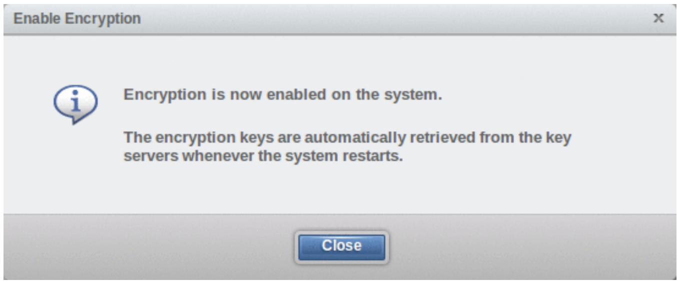
Encryption: Enabled

Key servers: 192.168.0.100:5696

Key server certificates: Installed 11/09/2016

*Figure 12-33 Finish enabling encryption using a key server*

8. If there are no errors while creating the key server object you will receive a message confirming the encryption is now enabled on the system, as shown in Figure 12-34 on page 585.



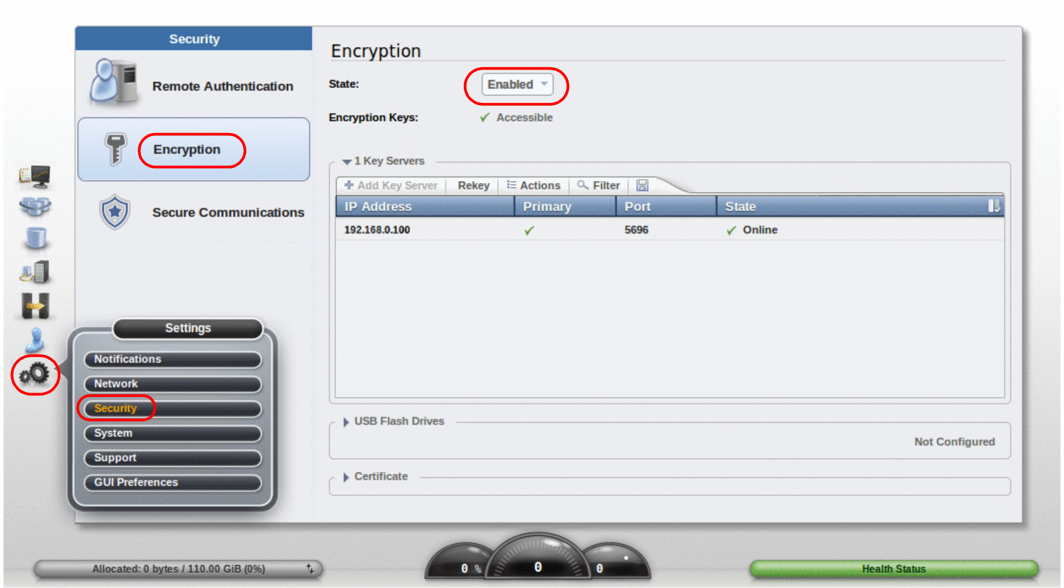
Enable Encryption

Encryption is now enabled on the system.

The encryption keys are automatically retrieved from the key servers whenever the system restarts.

*Figure 12-34 Encryption enabled message using a key server*

9. Confirm encryption is enabled in **Settings** → **Security** → **Encryption**, as shown in Figure 12-35



Encryption

State: Enabled

Encryption Keys: ✓ Accessible

1 Key Servers

| IP Address | Primary | Port | State |
|---|---|---|---|
| 192.168.0.100 | ✓ | 5696 | ✓ Online |

USB Flash Drives: Not Configured

Certificate

*Figure 12-35 Encryption enabled view using a key server*

## 12.4.3 Enabling encryption using USB flash drives

> **Note:** The system needs at least three USB flash drives to be present before you enable encryption using this method. Physical access to system is required. We recommend IBM USB flash drives, although other flash drives may work.
>
> Order IBM USB flash drives in eConfig as Feature Code ACEA: Encryption USB Flash Drives (Four Pack).

The process requires a minimum of three USB flash drives to store the generated encryption keys. This is required to enable encryption and to rekey the system. During these operations, you are responsible for ensuring the security of the system.

If your system is in a secure location, one USB flash drive for each canister can remain inserted in the system. If the location is not secure, all USB flash drives with the master access keys need to be removed from the system and be stored securely.

*A minimum of one USB flash drive with the correct master access key is required to enable encryption for the cluster after a system restart* such as a system wide reboot or power loss. No USB flash drive is required during a warm reboot, such as a node exiting service mode or a single node reboot. The data center power-on procedure needs to include instructions about the USB flash drive locations and how to use them.

During power-on, we recommend that you insert USB flash drives into the USB ports on two supported canisters to guard against any unexpected failure of the node, the node's USB port, or the USB flash drive during the power-on procedure. Use these general guidelines when you enable encryption and manage the USB flash drives that contain encryption keys.

While the system enables encryption, you are prompted to insert the USB flash drives into the system. The system copies the encryption keys to these drives systematically. The system generates and copies the encryption keys to all available USB flash drives.

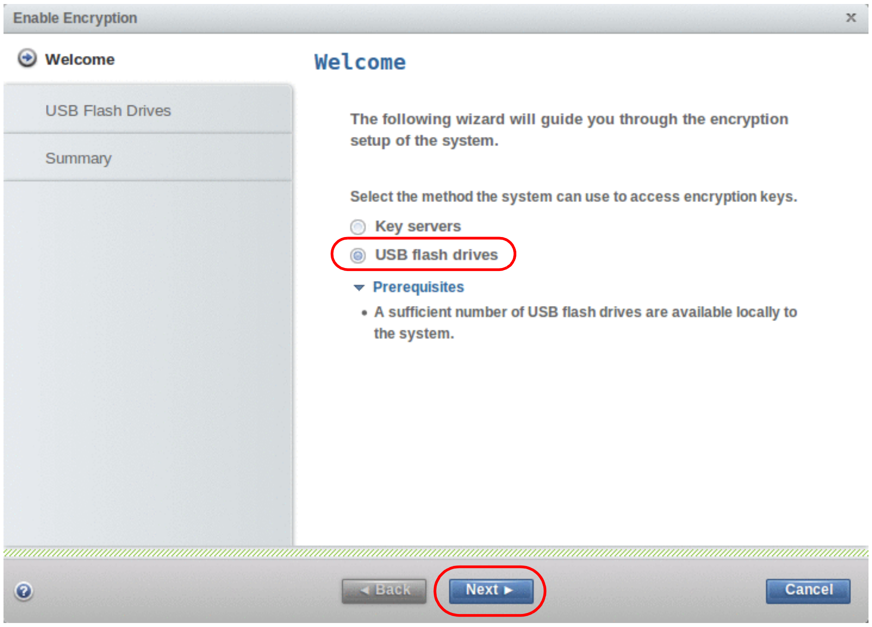
Ensure that each copy of the encryption key is valid before you write any user data to the system. The system validates any key material on a USB flash drive when it is inserted into the canister. If the key material is not valid, the system logs an error. If the USB flash drive is unusable or fails, the system does not display it as output.

Securely store all copies of the encryption key. For example, any USB flash drives that are not left inserted into the system can be locked in a safe. Comparable precautions must be taken to securely protect any other copies of the encryption key that are stored in other forms of storage.

> **Notes:** We recommend creating at least one additional copy on another USB flash drive to store in a secure location. This copy can be in other forms of storage to provide resiliency and to mitigate risk if, for example, the USB flash drives are from a faulty batch of drives.

To enable encryption using USB flash drives follow these steps:

1. In the **Enable Encryption** wizard **Welcome** tab select **USB flash drives** and click **Next**, as shown in Figure 12-36.



The following wizard will guide you through the encryption setup of the system.

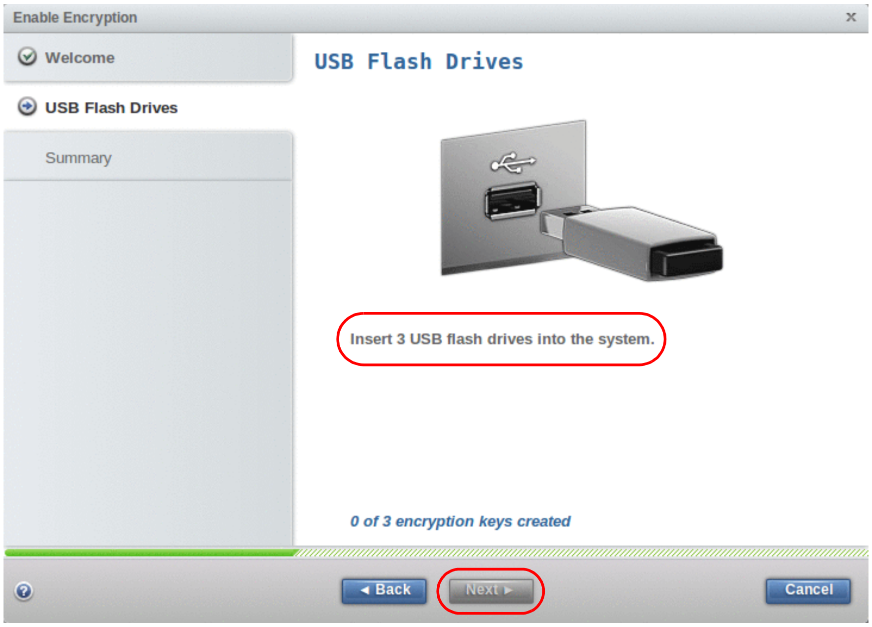
Select the method the system can use to access encryption keys.

○ Key servers
● USB flash drives

- A sufficient number of USB flash drives are available locally to the system.

*Figure 12-36 Selecting USB flash drives in the Enable Encryption wizard*

2. The wizard moves to the **USB Flash Drives** tab, as shown in Figure 12-37. The tab indicates the number of required USB flash drives.



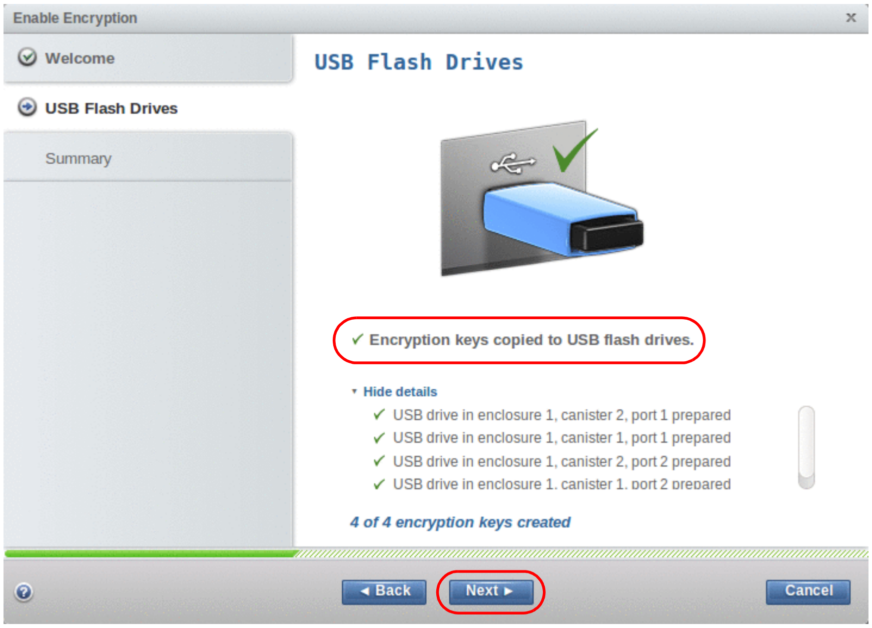
Insert 3 USB flash drives into the system.

*0 of 3 encryption keys created*

*Figure 12-37 Waiting for USB flash drives to be inserted*

> **Note:** The **Next** option remains disabled and the status at the bottom is kept to 0 until three USB flash drives are detected.

3. Insert the USB flash drives into the USB ports as requested.
4. After the minimum requirement is met, the encryption keys are automatically copied on the USB flash drives, as shown in Figure 12-38.



✓ Encryption keys copied to USB flash drives.

▾ Hide details

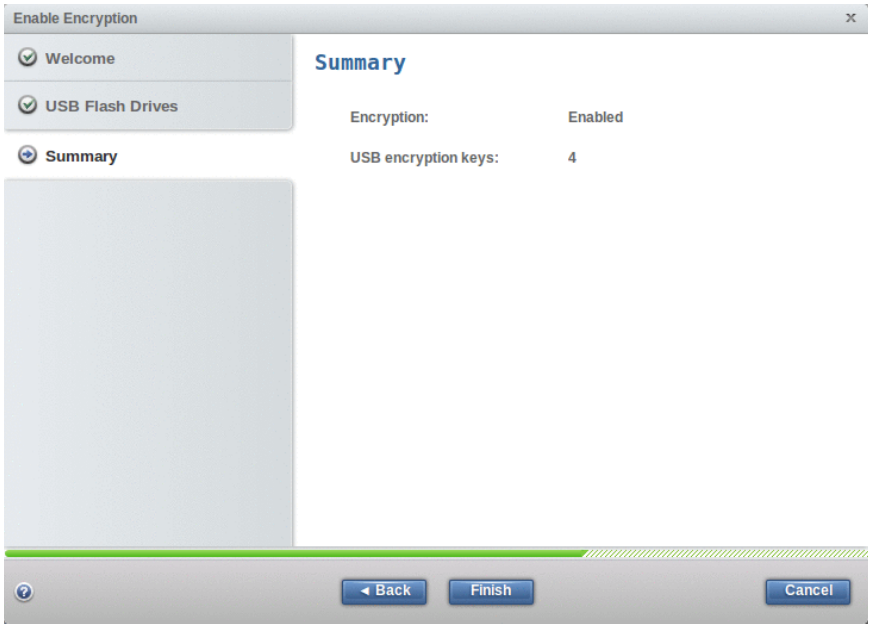
✓ USB drive in enclosure 1, canister 2, port 1 prepared
✓ USB drive in enclosure 1, canister 1, port 1 prepared
✓ USB drive in enclosure 1, canister 2, port 2 prepared
✓ USB drive in enclosure 1, canister 1, port 2 prepared

*4 of 4 encryption keys created*

*Figure 12-38 Writing the master access key to USB flash drives*

You can keep adding USB flash drives or replacing the ones already plugged in to create new copies. When done, click **Next**.

5. The number of keys created are shown in the **Summary** tab, as shown in Figure 12-39. Click **Finish** to finalize the encryption enablement.



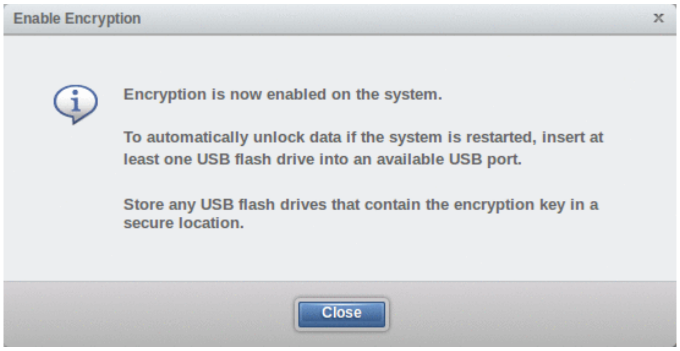
Summary

Encryption: Enabled

USB encryption keys: 4

*Figure 12-39 Commit the encryption enablement*

6. You will receive a message confirming the encryption is now enabled on the system, as shown in Figure 12-40.



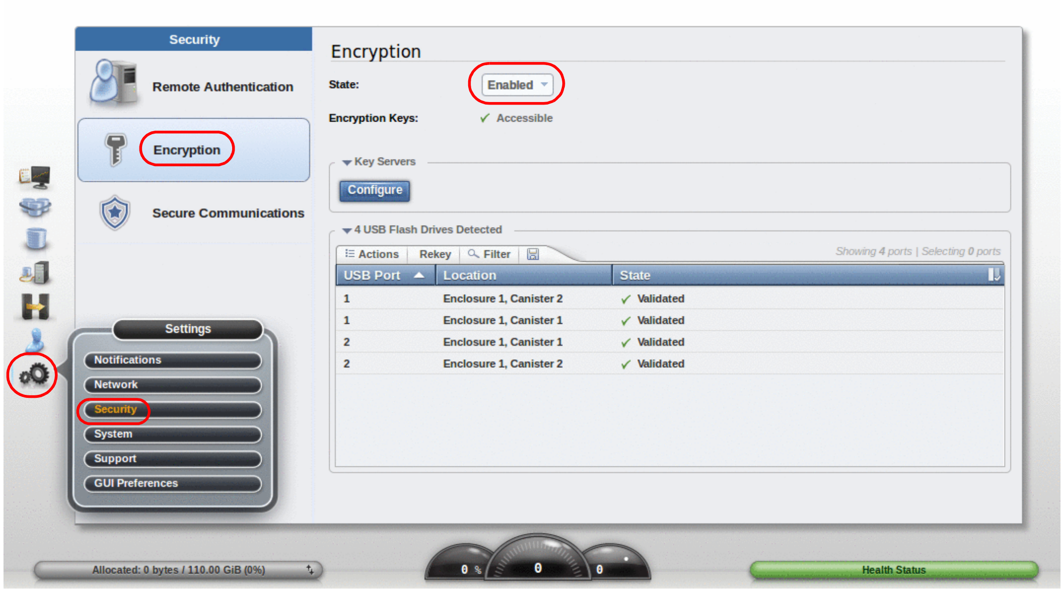
Encryption is now enabled on the system.

To automatically unlock data if the system is restarted, insert at least one USB flash drive into an available USB port.

Store any USB flash drives that contain the encryption key in a secure location.

*Figure 12-40 Encryption enabled message using USB flash drives*

7. Confirm encryption is enabled in **Settings** → **Security** → **Encryption**, as shown in Figure 12-41.



Encryption

State: Enabled

Encryption Keys: ✓ Accessible

Key Servers

Configure

4 USB Flash Drives Detected

| USB Port | Location | State |
|---|---|---|
| 1 | Enclosure 1, Canister 2 | ✓ Validated |
| 1 | Enclosure 1, Canister 1 | ✓ Validated |
| 2 | Enclosure 1, Canister 1 | ✓ Validated |
| 2 | Enclosure 1, Canister 2 | ✓ Validated |

*Figure 12-41 Encryption enabled view using USB flash drives*

# 12.5 Using encryption

The design for encryption is based on the concept that a system must either encrypt or not. The Spectrum Virtualize solution is not intended to encourage both encrypted volumes and unencrypted volumes. Some unsupported configurations are actively policed in code. For example, no support exists for creating unencrypted child pools from encrypted parent pools. However, exceptions exist:

- During the migration (conversion) of volumes from unencrypted to encrypted volumes, a system might report both encrypted and unencrypted volumes.
- Internal storage which is configured by using the CLI, where the client intentionally overrides the encryption default.

**Note:** Encryption support for Distributed RAID is available in IBM Spectrum Virtualize code V7.7 and later.

**Note:** *You must decide whether to encrypt or not encrypt an object when it is created*. You cannot change it after you create the object.

## 12.5.1 Encrypted pools

See Chapter 6, “Storage pools” on page 165 for generic instructions on how to open the **Create Pool** window. Once you opened the **Create Pool** window you can check and disable the encryption status of the new pool. First click on **Advanced**, as shown in Figure 12-42.

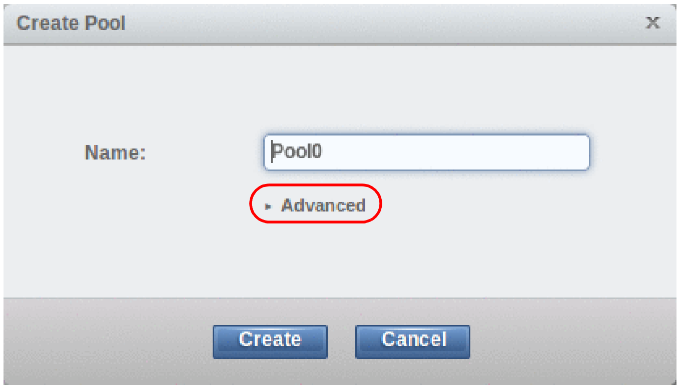
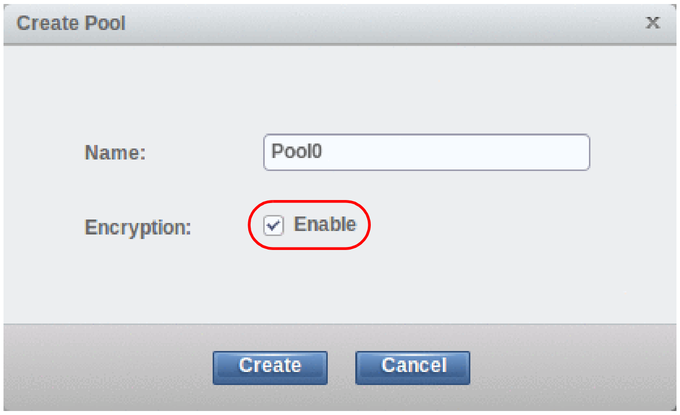


*Figure 12-42 Create Pool window basic*

When encryption is enabled at a system level, the option in the advanced panel defaults to **Encryption Enable**, as shown in Figure 12-43. The exception is if there are nodes in the system that do not support software encryption, in this case it is not possible to create encrypted pools.



Name: Pool0

Encryption: ☑ Enable

*Figure 12-43 Create Pool window advanced*

You can click **Create** to create a new encrypted pool. When doing so, all storage that you plan to add to this encrypted pool must be encrypted.

You can check whether a pool is encrypted in different views. For example, you can navigate to **Pools** → **Pools** and customize the view to show the encryption state by selecting the menu bar, right-clicking, and selecting the encryption key icon, as shown in Figure 12-44.

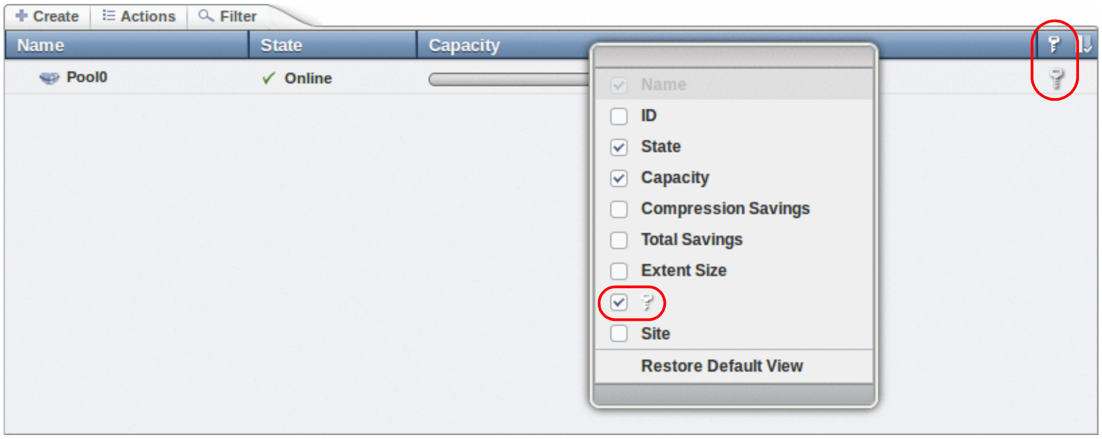
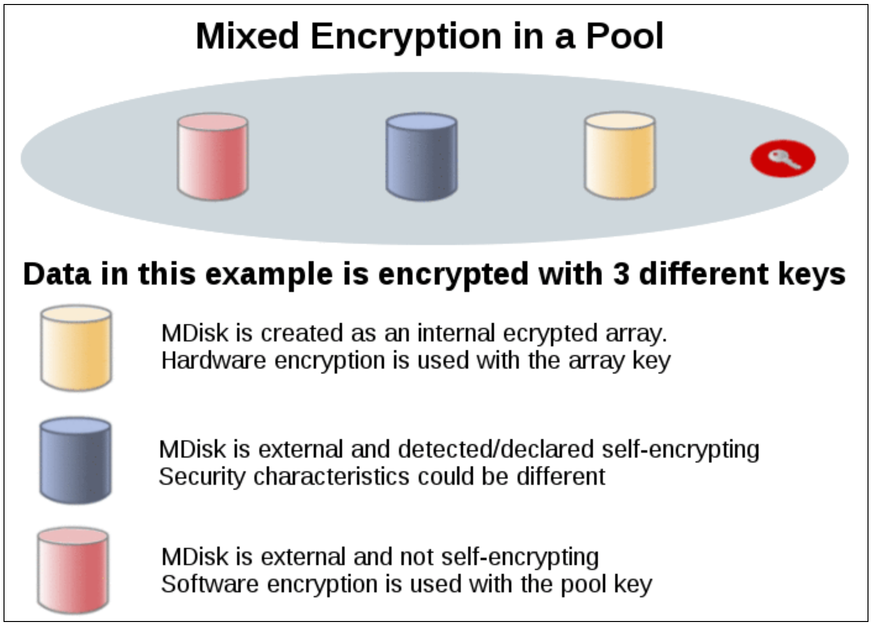


*Figure 12-44 Pool encryption state*

If you create an *unencrypted* pool it will not show the encryption key icon. However, if you only add encrypted arrays or self-encrypting MDisks to the pool it will then show as encrypted based on the fact that all extents are encrypted. It will revert back to the unencrypted state if you then add an unencrypted array or MDisk.

Further information on how to add encrypted storage to encrypted pools is in the following sections. You can mix and match storage encryption types in a pool. The graphic that is shown in Figure 12-45 provides guidance.



**Mixed Encryption in a Pool**

**Data in this example is encrypted with 3 different keys**

MDisk is created as an internal ecrypted array.
Hardware encryption is used with the array key

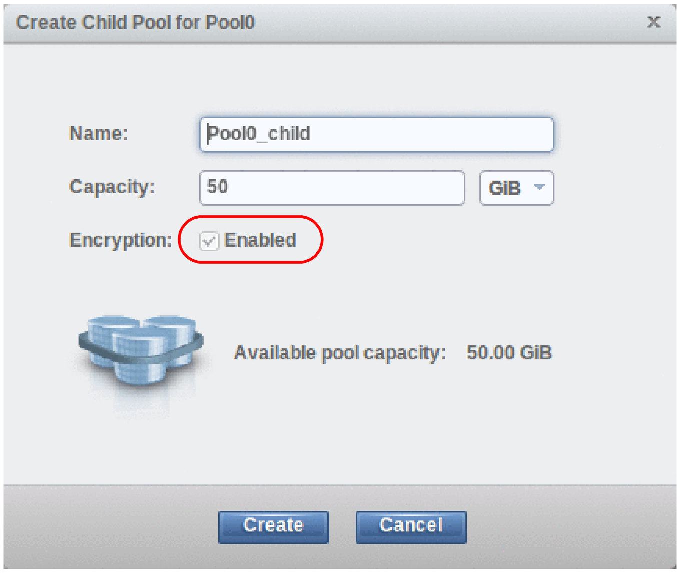
MDisk is external and detected/declared self-encrypting
Security characteristics could be different

MDisk is external and not self-encrypting
Software encryption is used with the pool key

*Figure 12-45 Mix and match encryption in a pool*

## 12.5.2 Encrypted child pools

See Chapter 6, “Storage pools” on page 165 for generic instructions on how to open the **Create Child Pool** window. If the parent pool is encrypted every child pool must be encrypted too, the GUI will enforce this by automatically ticking **Encryption Enabled** in the **Create Child Pool** window, as shown in Figure 12-46.



Create Child Pool for Pool0

Name: Pool0_child

Capacity: 50 GiB

Encryption: ☑ Enabled

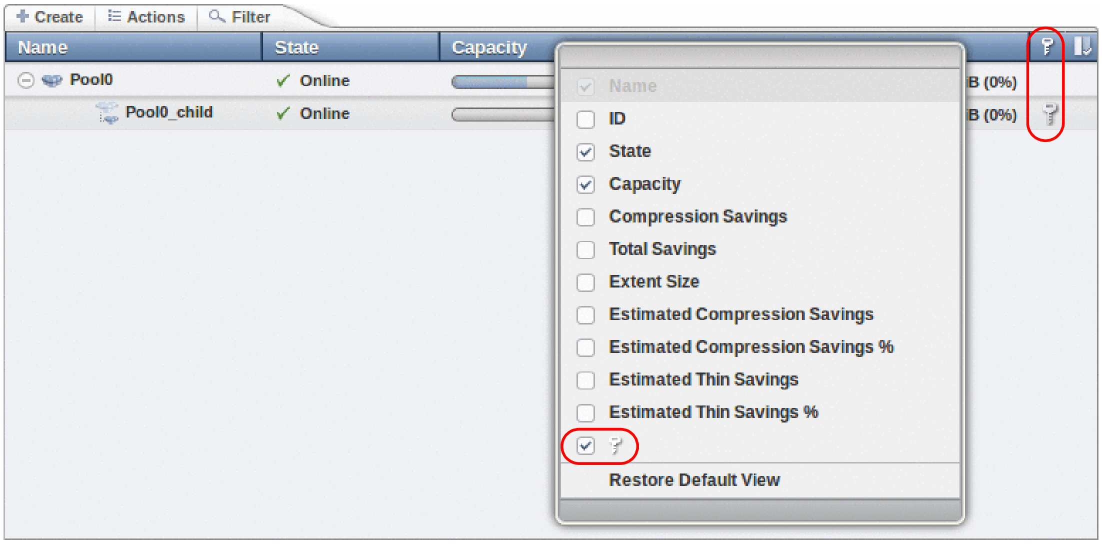
Available pool capacity: 50.00 GiB

*Figure 12-46 Create an encrypted child pool*

If the parent pool is not encrypted, an encrypted child pool can still be created from it by manually ticking **Encryption Enabled**. However, some restrictions apply:

- You cannot create an encrypted child pool from an unencrypted parent pool if the parent pool contains not self-encrypting MDisks and there are nodes in the system that do not support software encryption.
- An encrypted child pool created from an unencrypted parent pool will report as unencrypted if the parent pool contains unencrypted arrays. Remove these arrays to guarantee the child pool is fully encrypted.

To check whether a child pool is encrypted, navigate to **Pools** → **Pools** and customize the view to show the encryption state by selecting the menu bar, right-clicking, and selecting the encryption key icon, as shown in Figure 12-47. The figure also shows the case where an encrypted child pool is created from an unencrypted pool.



| Name | State | Capacity |
|---|---|---|
| Pool0 | ✓ Online | B (0%) |
| Pool0_child | ✓ Online | B (0%) |

☑ Name
☐ ID
☑ State
☑ Capacity
☐ Compression Savings
☐ Total Savings
☐ Extent Size
☐ Estimated Compression Savings
☐ Estimated Compression Savings %
☐ Estimated Thin Savings
☐ Estimated Thin Savings %
☑
Restore Default View

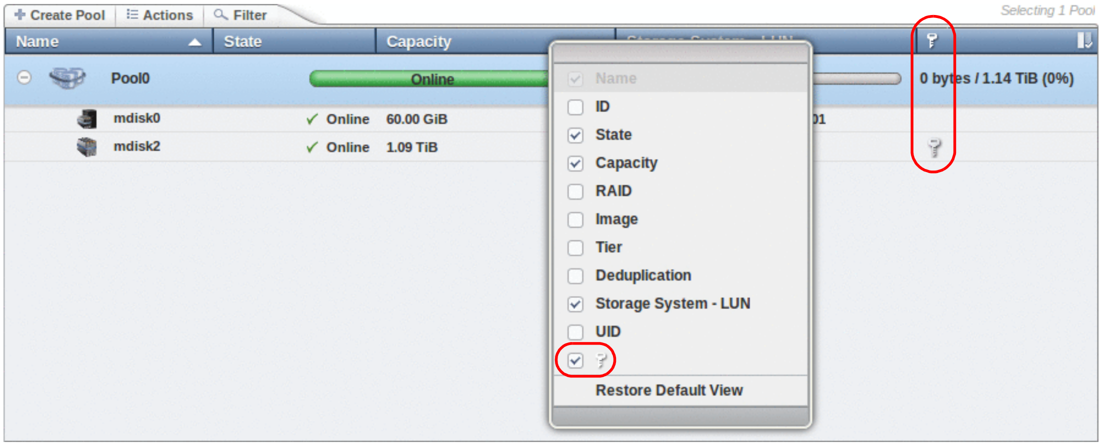
*Figure 12-47 Child pool encryption state*

### 12.5.3 Encrypted arrays

See Chapter 6, “Storage pools” on page 165 for generic instructions on how to add internal storage to a pool. After encryption is enabled, all newly built arrays are hardware encrypted by default. The graphical user interface (GUI) supports only this default option.

> **Note:** You can create an unencrypted array when encryption is enabled by using the command-line interface (CLI) with the `mkarray -encrypt no` command. However, you cannot add unencrypted arrays to an encrypted pool.
>
> To unencrypt an encrypted internal array, you need to migrate its extents to an unencrypted system or to an unencrypted pool that has no encrypted arrays or self-encrypting MDisks. Similarly, extent migration is the only option to encrypt an unencrypted internal array.

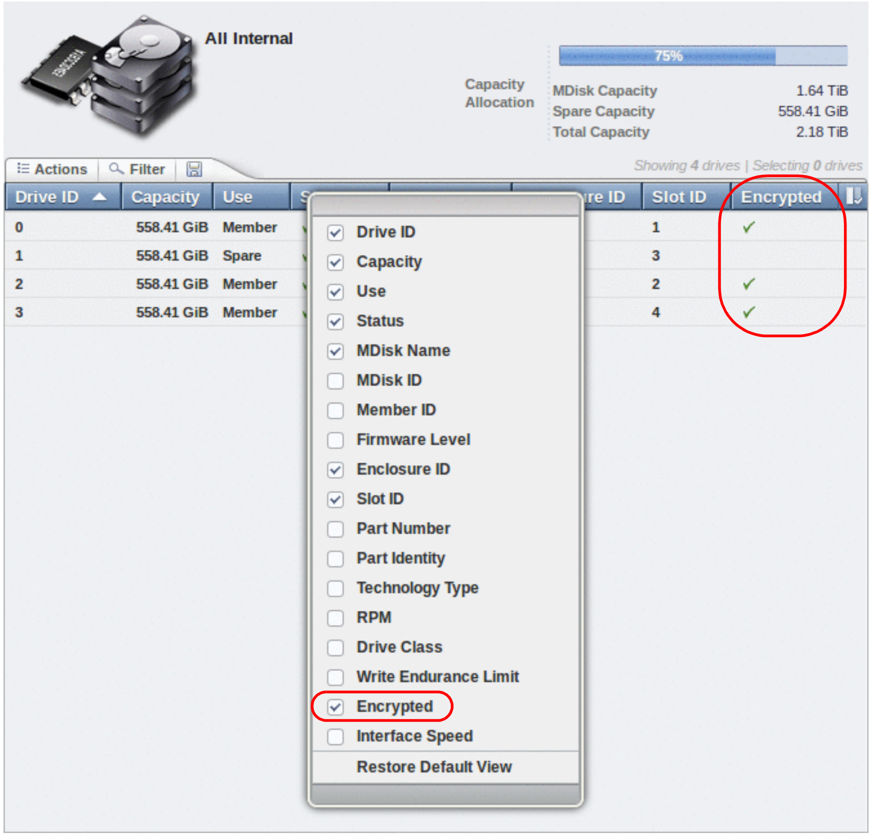
You can check whether an array is encrypted in different views. For example, you can navigate to **Pools** → **MDisk by Pools** and customize the view to show the encryption state by selecting the menu bar, right-clicking, and selecting the encryption key icon, as shown in Figure 12-48. The figure also shows that you can have encrypted arrays in unencrypted pools.



| Name | State | Capacity | |
|---|---|---|---|
| Pool0 | Online | | 0 bytes / 1.14 TiB (0%) |
| mdisk0 | ✓ Online | 60.00 GiB | |
| mdisk2 | ✓ Online | 1.09 TiB | |

Figure 12-48 Array encryption state

You can also check the encryption state of an array looking at its drives by selecting **Pools** → **Internal Storage**. The internal drives associated to an encrypted array are assigned an encrypted property that can be viewed by selecting the menu bar, right-clicking, and selecting the **Encrypted** option, as shown in Figure 12-49.



All Internal

Capacity Allocation 75%

MDisk Capacity 1.64 TiB
Spare Capacity 558.41 GiB
Total Capacity 2.18 TiB

Actions Filter

Showing 4 drives | Selecting 0 drives

| Drive ID | Capacity | Use | S | re ID | Slot ID | Encrypted |
|---|---|---|---|---|---|---|
| 0 | 558.41 GiB | Member | | | 1 | ✓ |
| 1 | 558.41 GiB | Spare | | | 3 | |
| 2 | 558.41 GiB | Member | | | 2 | ✓ |
| 3 | 558.41 GiB | Member | | | 4 | ✓ |

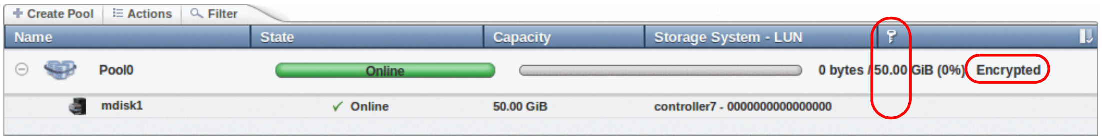
- ☑ Drive ID
- ☑ Capacity
- ☑ Use
- ☑ Status
- ☑ MDisk Name
- □ MDisk ID
- □ Member ID
- □ Firmware Level
- ☑ Enclosure ID
- ☑ Slot ID
- □ Part Number
- □ Part Identity
- □ Technology Type
- □ RPM
- □ Drive Class
- □ Write Endurance Limit
- ☑ Encrypted
- □ Interface Speed
- Restore Default View

*Figure 12-49 Drive encryption state*

## 12.5.4 Encrypted MDisks

See Chapter 6, “Storage pools” on page 165 for generic instructions on how to add internal storage to a pool. Each extent belonging external storage added to an encrypted pool or child pool is automatically encrypted using the pool or child pool key, unless the MDisk is detected or declared as self-encrypting.

There is no way to check which extents are encrypted and which are not. In general, if the parent pool is encrypted all the extents are encrypted. If you navigate to **Pools** → **MDisk by Pools** and customize the view to show the encryption state by selecting the menu bar, right-clicking, and selecting the encryption key icon the mdisk *will not show* the encryption key icon but it is fully encrypted, as shown in Figure 12-50.



Create Pool Actions Filter

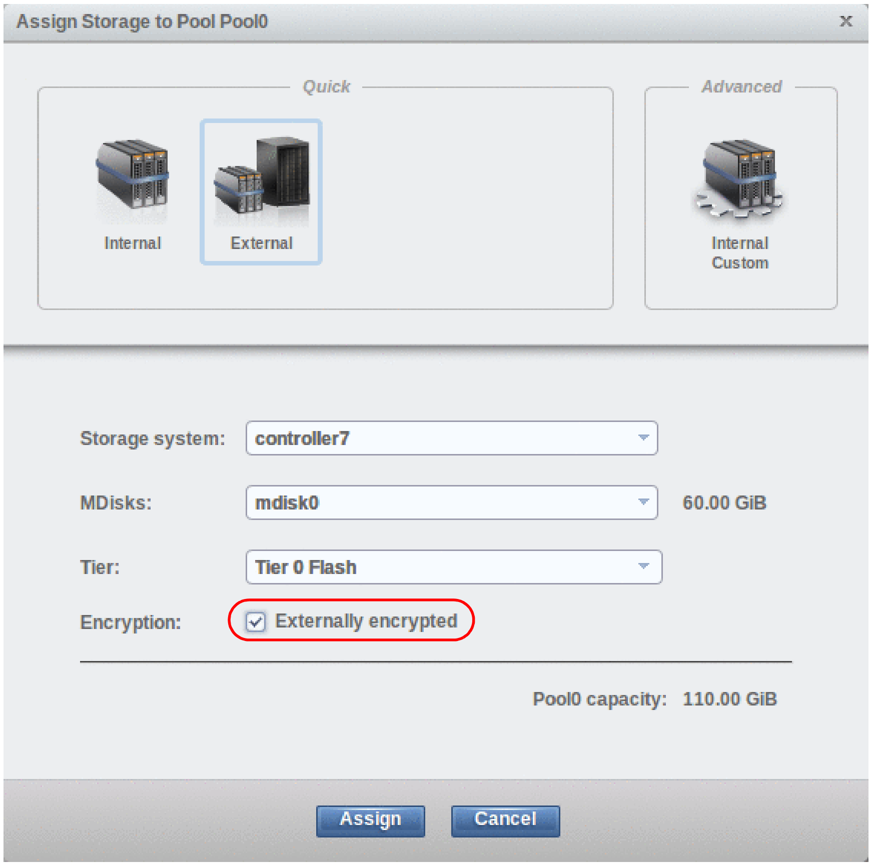
| Name | State | Capacity | Storage System - LUN | |
|---|---|---|---|---|
| Pool0 | Online | | 0 bytes / 50.00 GiB (0%) | Encrypted |
| mdisk1 | ✓ Online | 50.00 GiB | controller7 - 0000000000000000 | |

*Figure 12-50 MDisk encryption state*

### Self-encrypting MDisks

When adding external storage to an encrypted or not encrypted pool you should think carefully if you need to declare the MDisk as self-encrypting. Correctly declaring an MDisk as self-encrypting avoids waste of resources. However, when used improperly it might lead to unencrypted data at rest.

To declare an MDisk as self-encrypting tick **Externally encrypted** when adding external storage in the **Assign Storage** view, as shown in Figure 12-51.


Assign Storage to Pool Pool0

Storage system: controller7

MDisks: mdisk0 60.00 GiB

Tier: Tier 0 Flash

Encryption: ☑ Externally encrypted

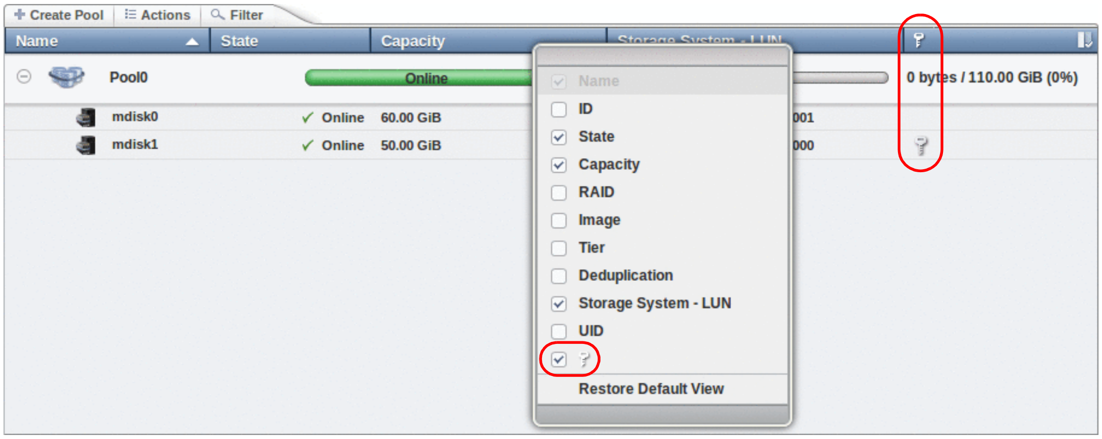
Pool0 capacity: 110.00 GiB

*Figure 12-51 Externally encrypted MDisk*

Spectrum Virtualize products can detect that an MDisk is self-encrypting by using the Inquiry page C2. MDisks provided by other Spectrum Virtualize products will report this page correctly. For these MDisks the **Externally encrypted** box shown in Figure 12-51 *will not* be ticked. However, when added, they will be still considered as self-encrypting.

> **Note:** You can override an MDisk detected as self-encrypting to unencrypted by using the command-line interface (CLI) with the `chmdisk -encrypt no` command. However, you should only do so if you plan to unencrypt the data on the backend or if the backend uses a lower level of security to encrypt data.

To check whether an MDisk has been detected or declared as self-encrypting, navigate to **Pools** → **MDisk by Pools** and customize the view to show the encryption state by selecting the menu bar, right-clicking, and selecting the encryption key icon, as shown in Figure 12-52. The figure also shows that you can have self-encrypting MDisks in unencrypted pools.



| Name | State | Capacity | Storage System - LUN | 🔑 |
|---|---|---|---|---|
| Pool0 | Online | | | 0 bytes / 110.00 GiB (0%) |
| mdisk0 | ✓ Online | 60.00 GiB | 001 | |
| mdisk1 | ✓ Online | 50.00 GiB | 000 | 🔑 |

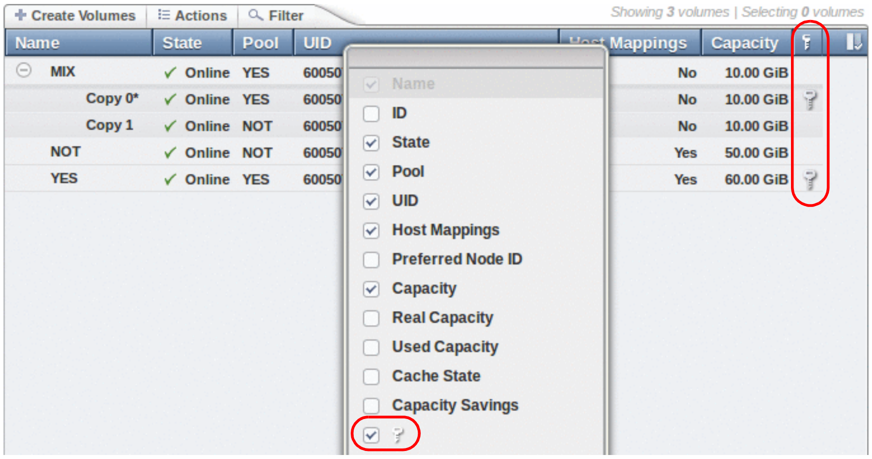
*Figure 12-52 MDisk self-encryption state*

> **Note:** You can change the self-encrypting attribute of an MDisk that is part of an unencrypted pool at any time. However, *you cannot change the self-encrypting attribute of an MDisk after it has been added to an encrypted pool.*

## 12.5.5 Encrypted volumes

See Chapter 7, "Volumes" on page 213 for generic instructions on how to create and manage volumes. The encryption status of a volume depends on the pool encryption status. Volumes created in an encrypted pool are automatically encrypted.

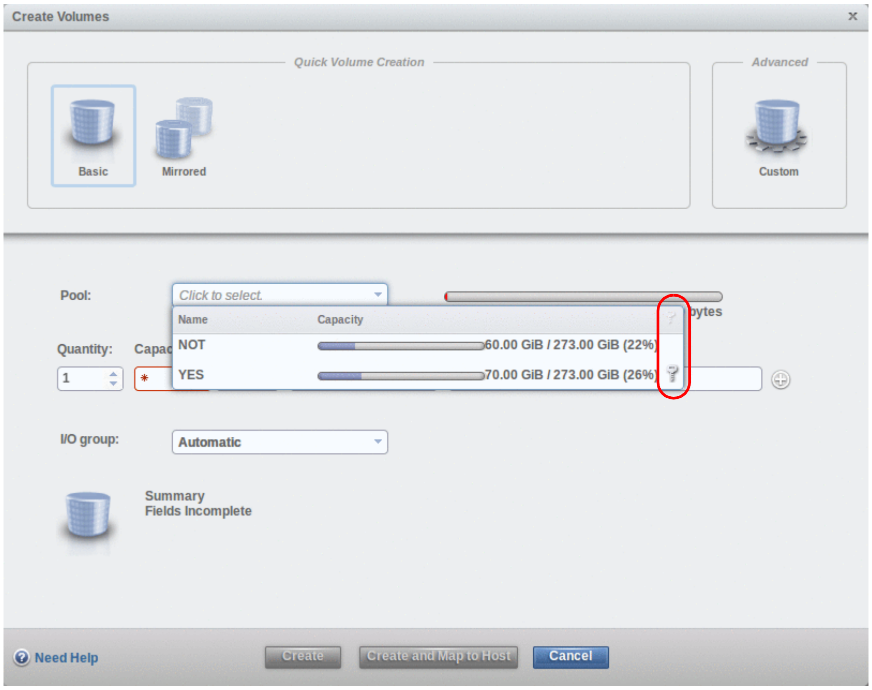
You can check whether a volume is encrypted in different views. For example, you can navigate to **Volumes** → **Volumes** and customize the view to show the encryption state by selecting the menu bar, right-clicking, and selecting the encryption key icon, as shown in Figure 12-53. The figure also shows that a volume is fully encrypted only if all the volume copies are encrypted.



| Name | State | Pool | UID | Host Mappings | Capacity | 🔑 |
|---|---|---|---|---|---|---|
| MIX | ✓ Online | YES | 60050 | No | 10.00 GiB | |
| Copy 0* | ✓ Online | YES | 60050 | No | 10.00 GiB | 🔑 |
| Copy 1 | ✓ Online | NOT | 60050 | No | 10.00 GiB | |
| NOT | ✓ Online | NOT | 60050 | Yes | 50.00 GiB | |
| YES | ✓ Online | YES | 60050 | Yes | 60.00 GiB | 🔑 |

*Figure 12-53 Volume encryption state*

When creating volumes make sure to select encrypted pools to create encrypted volumes, as shown in Figure 12-54.



| Name | Capacity |
|---|---|
| NOT | 60.00 GiB / 273.00 GiB (22%) |
| YES | 70.00 GiB / 273.00 GiB (26%) |

Figure 12-54 Create an encrypted volume by selecting an encrypted pool

You cannot change an existing volume to an encrypted version of itself dynamically. However, this conversion is possible by using two migration options:

- Migrate a volume to an encrypted pool or child pool
- Mirror a volume to an encrypted pool or child pool and delete the unencrypted copy

For more information on either method see Chapter 7, “Volumes” on page 213

## 12.6 Rekeying an encryption-enabled system

Changing the master access key is a security requirement. *Rekeying* is the process of creating a new master access key for the system and store it properly. To rekey, encryption must be enabled on the system. However, the rekey operation works whether or not encrypted objects already exist. The rekey operation uses a very similar process to the one to enable encryption and requires a valid copy of the original master access key on a USB flash drive or a key server to run. Use the rekey operation only if the key is compromised or to implement security policies.

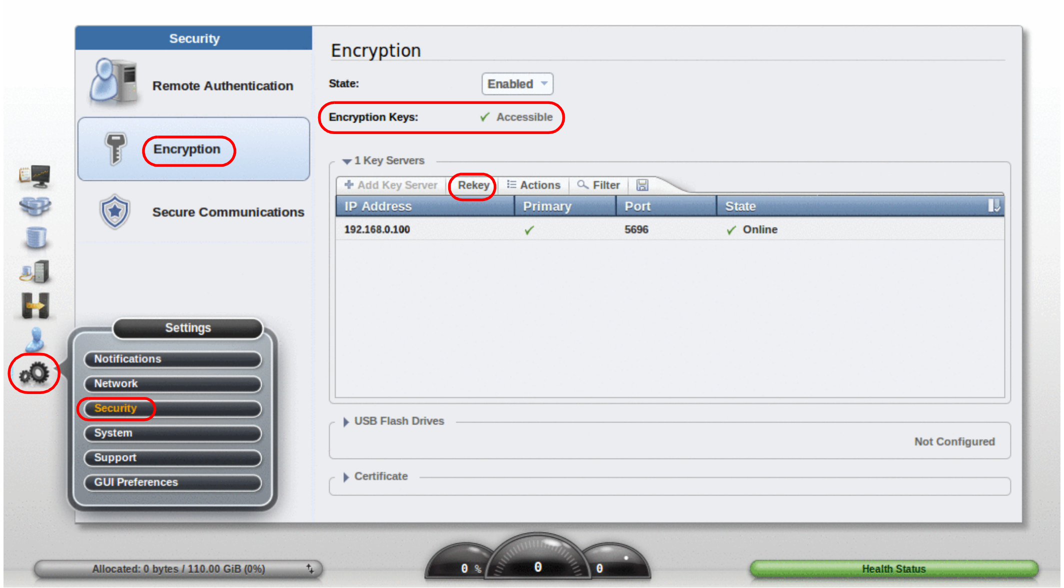
**Important:** Before you create a new master access key, ensure that all nodes are online and that the current master access key is accessible.

### 12.6.1 Rekeying using a key server

Ensure the configured key server is accessible by the system and that *service IPs* are still configured on all your nodes.

To rekey the system using a key server complete these steps:

1. Navigate to **Settings** → **Security** → **Encryption**, ensure that **Encryption Keys** shows that at least one master access key copy is `Accessible`, and click on **Rekey**, as shown in Figure 12-55.



Encryption

State: Enabled

Encryption Keys: ✓ Accessible

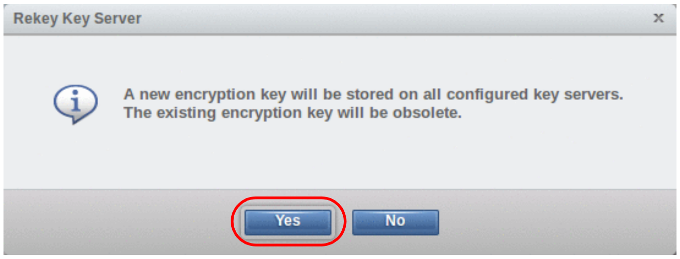
1 Key Servers

| IP Address | Primary | Port | State |
|---|---|---|---|
| 192.168.0.100 | ✓ | 5696 | ✓ Online |

USB Flash Drives: Not Configured

*Figure 12-55 Start key server rekey operation*

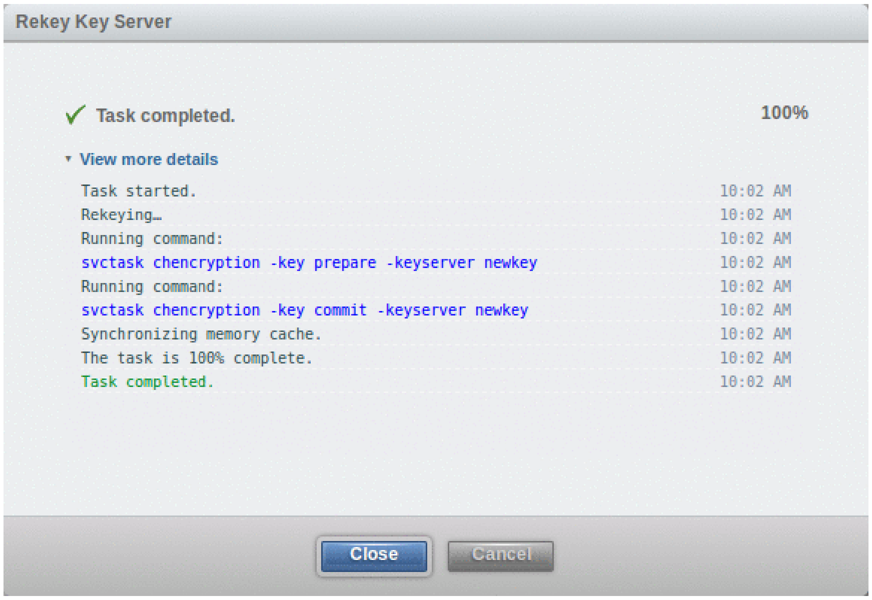
2. Click **Yes** in the next window to confirm the rekey operation, as shown in Figure 12-56.



A new encryption key will be stored on all configured key servers.
The existing encryption key will be obsolete.

*Figure 12-56 Confirm key server rekey operation*

3. You will receive a message confirming the rekey operation was successful, as shown in Figure 12-57.



```
Task started.                                         10:02 AM
Rekeying…                                             10:02 AM
Running command:                                      10:02 AM
svctask chencryption -key prepare -keyserver newkey   10:02 AM
Running command:                                      10:02 AM
svctask chencryption -key commit -keyserver newkey    10:02 AM
Synchronizing memory cache.                           10:02 AM
The task is 100% complete.                            10:02 AM
Task completed.                                       10:02 AM
```

*Figure 12-57 Successful key server rekey operation*

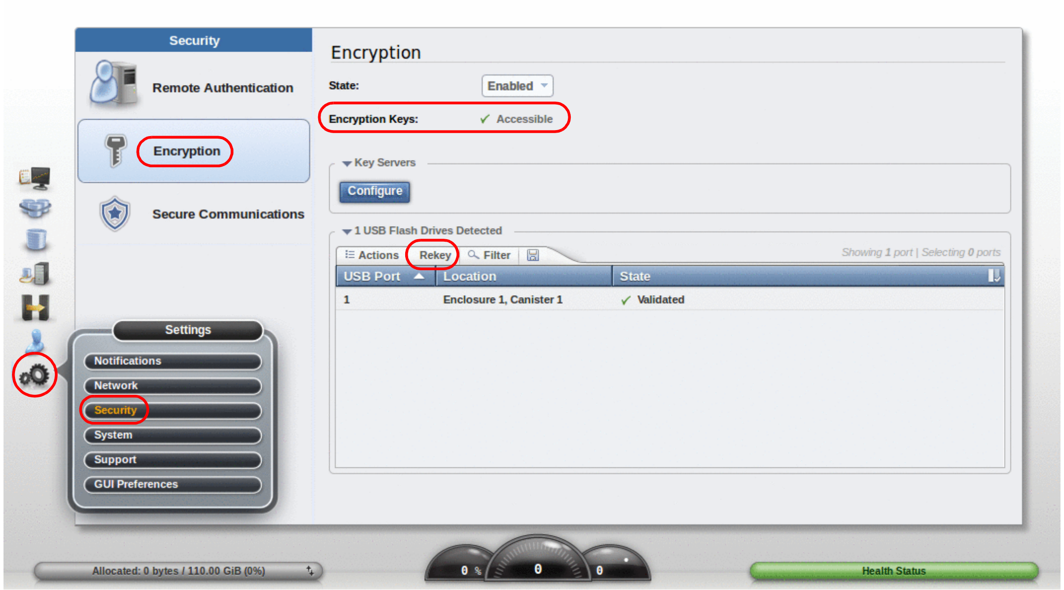
## 12.6.2 Rekeying using USB flash drives

During the rekey process, new keys are generated and copied to the USB flash drives. These keys are then used instead of the current keys. The rekey operation fails if at least one USB flash drive does not contain the current key. To rekey the system, you need at least three USB flash drives to store the master access key copies.

Once the rekey operation is complete, update all other copies of the encryption key that are stored in other forms of storage. Also, take the same precautions to securely store all copies of the encryption key you took when enabling encryption for the first time.

To rekey the system using USB flash drives complete these steps:

1. Navigate to **Settings** → **Security** → **Encryption**, ensure that **Encryption Keys** shows that at least one master access key copy is `Accessible`, and click on **Rekey**, as shown in Figure 12-58.



Encryption

State: Enabled

Encryption Keys: ✓ Accessible

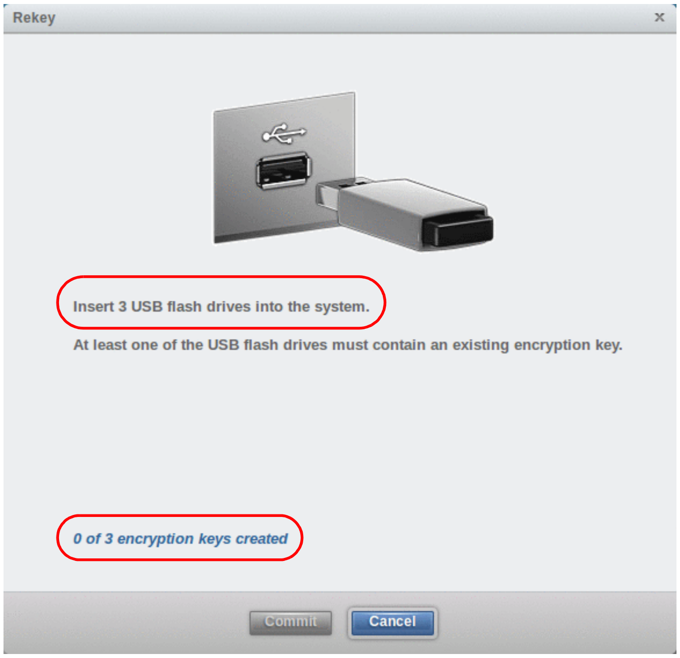
1 USB Flash Drives Detected

Showing 1 port | Selecting 0 ports

| USB Port | Location | State |
|---|---|---|
| 1 | Enclosure 1, Canister 1 | ✓ Validated |

*Figure 12-58 Start USB flash drives rekey operation*

2. The **Rekey** wizard will appear, as shown in Figure 12-59. It indicates the number of required USB flash drives and the number of USB flash drives already inserted.



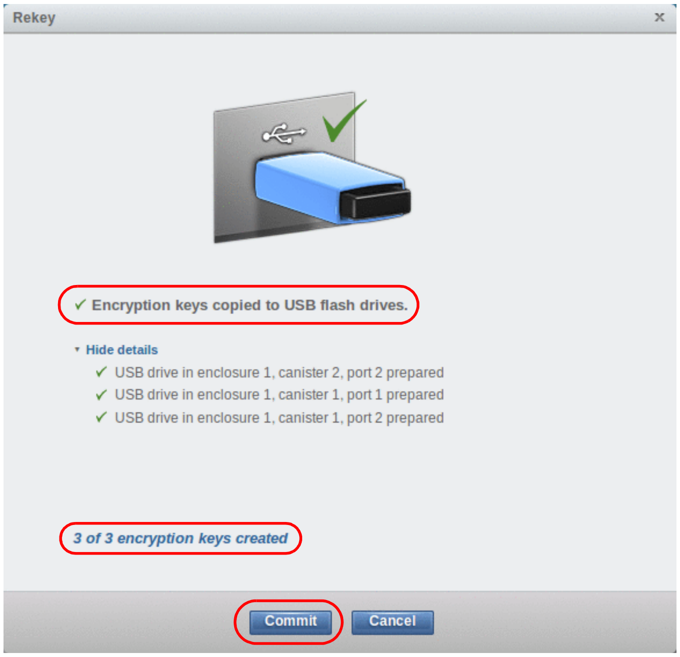
Insert 3 USB flash drives into the system.

At least one of the USB flash drives must contain an existing encryption key.

0 of 3 encryption keys created

*Figure 12-59 Waiting for USB flash drives to be inserted*

3. Insert the USB flash drives into the USB ports as requested.
4. After the minimum requirement is met, new encryption keys are automatically copied on the USB flash drives, as shown in Figure 12-60. Click **Commit** to finalize the rekey operation.



✓ Encryption keys copied to USB flash drives.

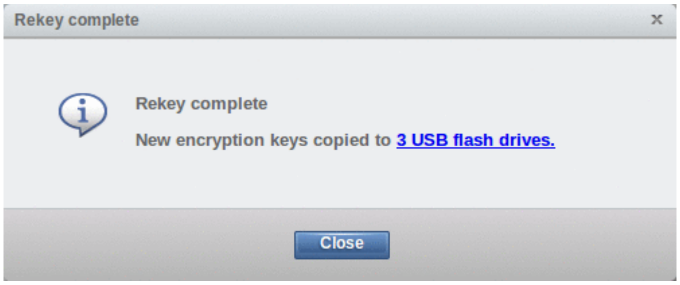
Hide details

✓ USB drive in enclosure 1, canister 2, port 2 prepared
✓ USB drive in enclosure 1, canister 1, port 1 prepared
✓ USB drive in enclosure 1, canister 1, port 2 prepared

3 of 3 encryption keys created

*Figure 12-60 Writing new keys to USB flash drives*

5. You will receive a message confirming the rekey operation was successful, as shown in Figure 12-61.



Rekey complete

New encryption keys copied to 3 USB flash drives.

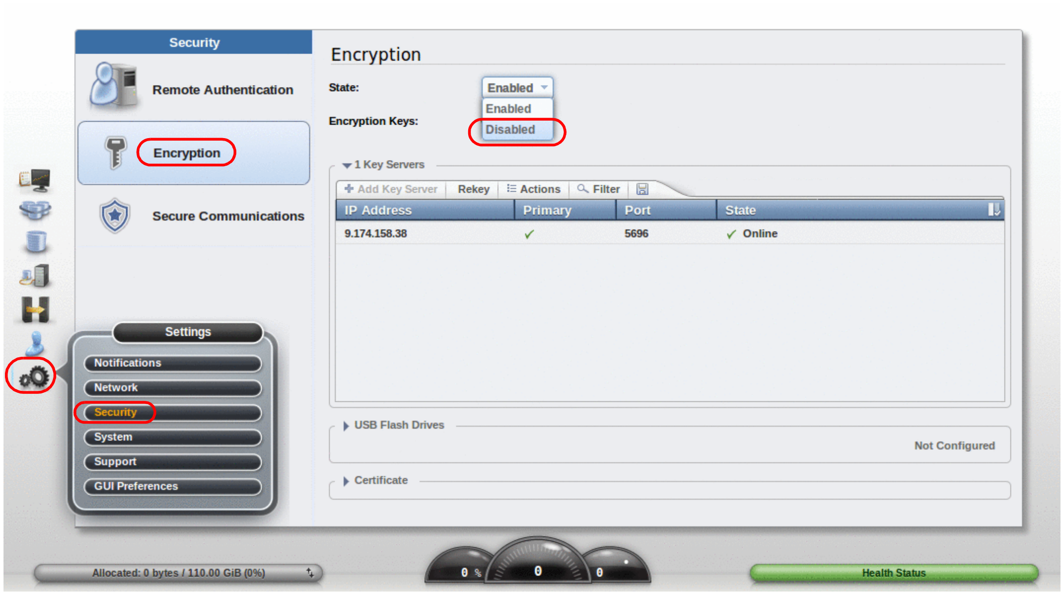
*Figure 12-61 Successful rekey operation using USB flash drives*

## 12.7 Disabling encryption

You are prevented from disabling encryption if encrypted objects exist, with the exception of self-encrypting MDisks. Disabling encryption when using USB flash drives or a key server is very similar.

To disable encryption follow these steps:

1. Navigate to **Settings** → **Security** → **Encryption** and click on **Enabled**. A drop down menu will be displayed unless an encrypted object exists, click on **Disabled**. Figure 12-62 shows an example using a key server.



Encryption

State: Enabled (Enabled / Disabled)

Encryption Keys:

1 Key Servers

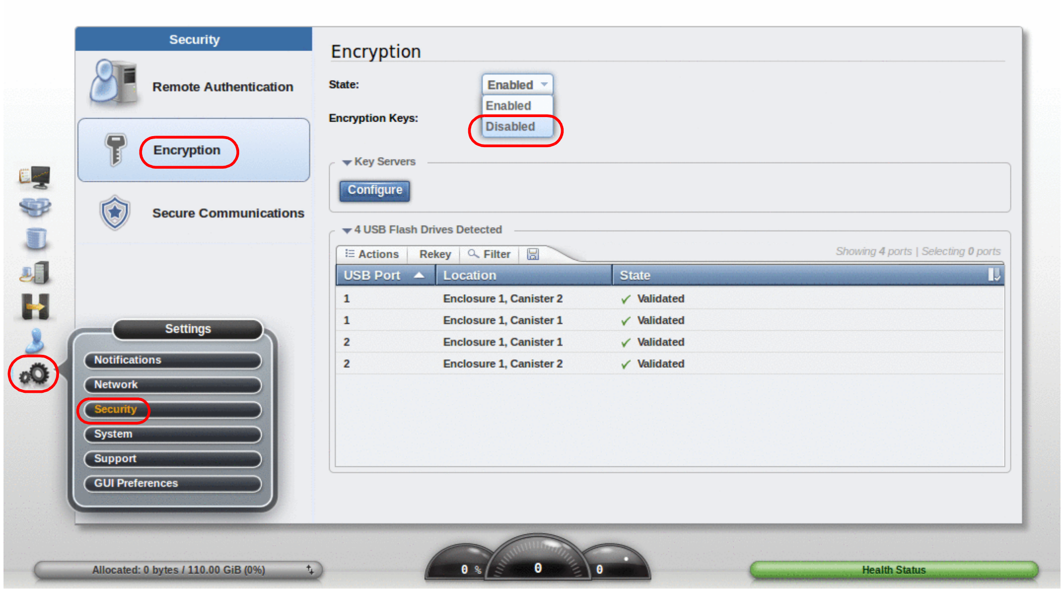
| IP Address | Primary | Port | State |
|---|---|---|---|
| 9.174.158.38 | ✓ | 5696 | ✓ Online |

USB Flash Drives — Not Configured

Certificate

Figure 12-62 Disabling encryption using a key server

Figure 12-63 shows an example using USB flash drives.



Encryption

State: Enabled (Enabled / Disabled)

Encryption Keys:

Key Servers

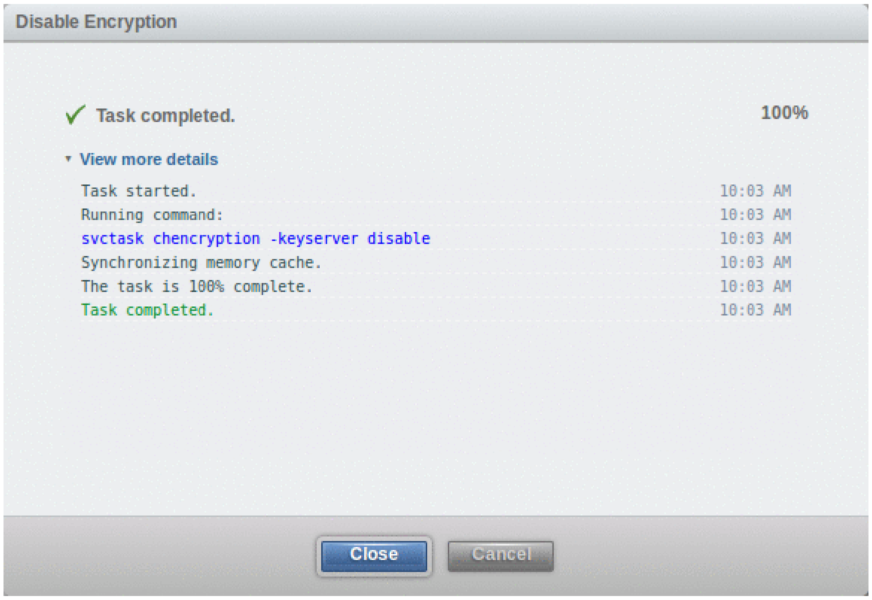
4 USB Flash Drives Detected

Showing 4 ports | Selecting 0 ports

| USB Port | Location | State |
|---|---|---|
| 1 | Enclosure 1, Canister 2 | ✓ Validated |
| 1 | Enclosure 1, Canister 1 | ✓ Validated |
| 2 | Enclosure 1, Canister 1 | ✓ Validated |
| 2 | Enclosure 1, Canister 2 | ✓ Validated |

Figure 12-63 Disabling encryption using USB flash drives

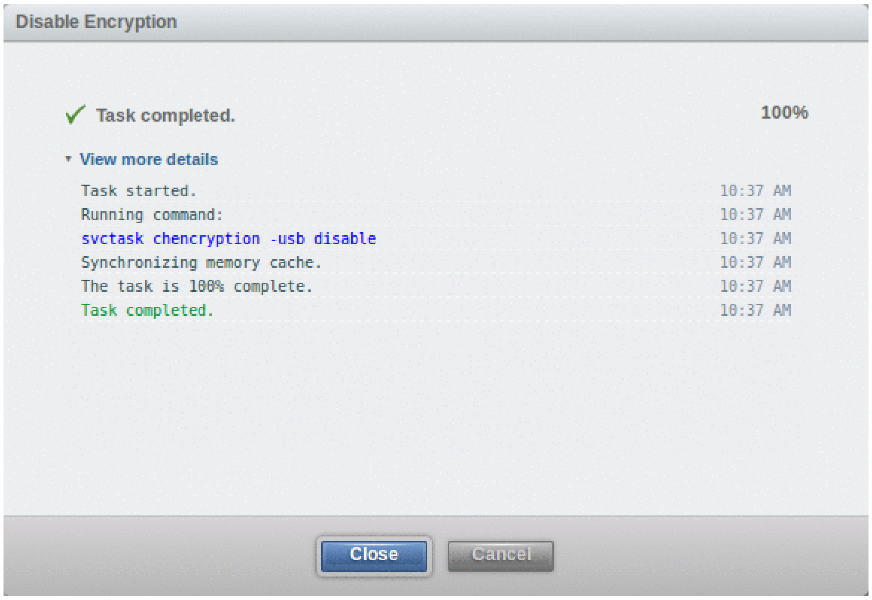
2. You will receive a message confirming encryption has been disabled. Figure 12-64 shows the message when using a key server.



```
Task started.                                  10:03 AM
Running command:                               10:03 AM
svctask chencryption -keyserver disable        10:03 AM
Synchronizing memory cache.                    10:03 AM
The task is 100% complete.                     10:03 AM
Task completed.                                10:03 AM
```

*Figure 12-64 Encryption disabled using a key server*

Figure 12-65 shows the message when using USB flash drives.



```
Task started.                                  10:37 AM
Running command:                               10:37 AM
svctask chencryption -usb disable              10:37 AM
Synchronizing memory cache.                    10:37 AM
The task is 100% complete.                     10:37 AM
Task completed.                                10:37 AM
```

*Figure 12-65 Encryption disabled using USB flash drives*

## 12.8 Restrictions

The following restrictions apply to encryption:

- Image mode volumes cannot be in encrypted pools.
- You cannot add external MDisks not self-encrypting to encrypted pools unless all nodes in the cluster support encryption.
- Nodes that cannot perform software encryption cannot be added to systems with encrypted pools containing external MDisks not self-encrypting.

13

# RAS, monitoring, and troubleshooting

This chapter introduces useful and common procedures to maintain the IBM Spectrum Virtualize.

The following topics are covered within this chapter:

- Hardware descriptions with status indications
- Monitoring from a host
- Monitoring from the IBM Storwize V7000 system
- Event log navigation
- Support options
- Shutting down the infrastructure with the installed IBM Storwize V7000 system
- Service Assistant Tool

## 13.1 Reliability, availability, and serviceability

Reliability, availability, and serviceability (RAS) are important concepts in the design of the IBM Spectrum Virtualize system. Hardware features, software features, design considerations, and operational guidelines all contribute to make the IBM Storwize V7000 system reliable.

Fault tolerance and high levels of availability are achieved by the following methods:

- The Redundant Array of Independent Disks (RAID) capabilities of the underlying disks
- IBM Storwize V7000 nodes clustering using a *Compass* architecture
- Auto-restart of hung nodes
- Integrated uninterruptible power supply (UPS) units to provide memory protection if there is a site power failure
- Host system failover capabilities

High levels of serviceability are available through the following methods:

- Cluster error logging
- Asynchronous error notification
- Dump capabilities to capture software detected failures
- Concurrent diagnostic procedures
- Directed maintenance procedures with simplified drive replacement process
- Concurrent log analysis and memory dump data recovery tools
- Concurrent maintenance of all IBM Storwize V7000 components
- Concurrent upgrade of IBM Storwize V7000 Software and microcode
- Concurrent addition or deletion of node canisters in a clustered system
- Automatic software version correction when replacing a node
- Detailed status and error conditions displayed on the service panel
- Error and event notification through Simple Network Management Protocol (SNMP), syslog, and email

The heart of IBM Storwize V7000 system is a pair of *node canisters*. These two canisters share the data transmitting and receiving load between the attached hosts and the disk arrays. This section examines the RAS features of IBM Storwize V7000 system, monitoring, and troubleshooting.

Throughout this chapter, the term ‘IBM Storwize V7000’ are referred to both models of the product; IBM Storwize V7000 Gen2 and Gen2+.

### 13.1.1 Node canisters

The two node canisters are in the control enclosure, and they work as a clustered system. Figure 13-1 on page 607 shows the ports and indicator lights of a node canister. Rear canister light-emitting diodes (LEDs) are identical to front panel LEDs. The second canister is placed next to the first one in a side-by-side position.

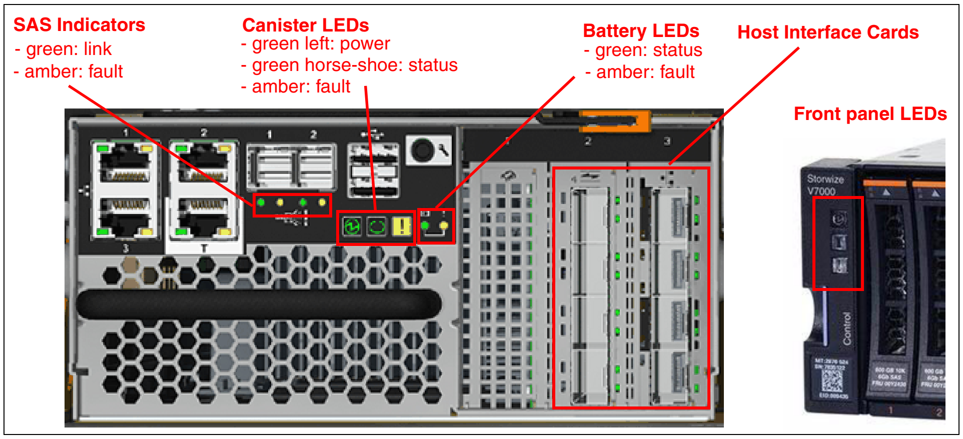


*Figure 13-1 Ports and indicators of node canister (524 controller model)*

## Host interface cards

Two host interface card (HIC) slots are vertically placed on the right side of the canister complemented by an additional slot for a Compression Accelerator card or compression pass-through module. HIC slots can accommodate up to two dual-port 16 gigabit per second (Gbps) Fibre Channel (FC) cards, or one four-port 10 gigabit Ethernet (GbE) adapter stand-alone or with combination with FC card.

The 16 Gbps FC adapter is based on the Emulex Lancer multiprotocol chip, and supports FC or Fibre Channel over Ethernet (FCoE) traffic using a single chip (Converged Network Adapter). It can be configured as dual-port 16 Gbps FC adapter or as a four-port 10 GbE with FCoE support. In case of FC configuration, the meaning of port LEDs is explained in Table 13-1.

*Table 13-1 Fibre Channel link LED statuses*

| Port LED | Color | Meaning |
|---|---|---|
| Link status | Green | Link is up, connection established. |
| Speed | Amber | Link is not up or speed fault. |

## Universal Serial Bus

Two active Universal Serial Bus (USB) connectors are available in the horizontal position. They have no numbers and no indicators are associated with them.

## Ethernet and LED status

Four 10/100/1000 megabits per second (Mbps) Ethernet ports are side by side on the canister. They are marked as 1 and 3 on the left and 2 and T on the right. Each port has two LEDs, and their status values are shown in Table 13-2. Although ports one to three are available for management and Internet Small Computer System Interface (iSCSI) purposes, the T port is strictly dedicated to the technician actions (initial and emergency configuration by local support personnel).

*Table 13-2 Ethernet LED statuses*

| LED | Color | Meaning |
|---|---|---|
| Link state | Green | It is on when there is an Ethernet link. |
| Activity | Amber | It is flashing when there is activity on the link. |

### Serial-attached SCSI ports

Two 12 Gbps serial-attached SCSI (SAS) ports are side by side on the canister with indicator LEDs below them. They are numbered 1 on the left and 2 on the right. Each port is associated with one green and one amber LED indicating its status of the operation, as shown in Table 13-3.

*Table 13-3 SAS LED statuses*

| LED | Meaning |
|---|---|
| Green | Link is connected and up. |
| Orange | Fault on the SAS link (disconnected, wrong speed, errors). |

### Node canister status LEDs

There are three LEDs in a row in the upper middle position of the canister that indicate the status and the functionality of the node (Table 13-4 on page 608).

*Table 13-4 Node canister LEDs*

| Position | Color | Name | State | Meaning |
|---|---|---|---|---|
| Left | Green | Power | On | The node is started and active. It might not be safe to remove the canister. If the fault LED is off, the node is an active member of a cluster or candidate. If the fault LED is also on, node is in service state or in error preventing the software to start. |
| | | | Flashing (2 Hz) | Canister is started and in standby mode. |
| | | | Blinking (4 Hz) | Node is running power-on self-test (POST). |
| | | | Off | No power to the canister or it is running on battery. |
| Middle | Green | Status | On | The node is a member of a cluster. |
| | | | Flashing (2 Hz) | The node is a candidate for or in a service state. |
| | | | Blinking (4 Hz) | The node is performing a fire hose dump. Never unplug the canister at this time. |
| | | | Off | No power to the canister or canister is in standby mode. |
| Right | Amber | Fault | On | The canister is in a service state, or in error, preventing the software from starting. |
| | | | Blinking (2 Hz) | Canister is being identified. |
| | | | Off | Node is either in candidate or active state. |

## 13.1.2 Expansion canisters

As Figure 13-2 shows, two 12 Gbps SAS ports are side by side on the canister of every enclosure. They are numbered 1 on the left and 2 on the right. Similar to the controller canisters, expansion canisters are also installed in the enclosure side-by-side in a vertical position.

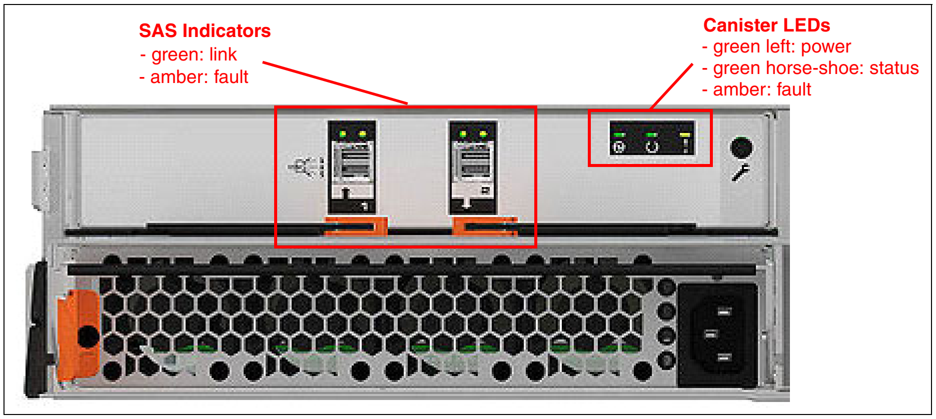


*Figure 13-2 Expansion canister status LEDs*

The interpretation of the SAS status LED indicators has the same meaning as the LED indicators of SAS ports in the control enclosure (Table 13-3 on page 608).

Table 13-5 shows the LED status values of the expansion canister.

*Table 13-5 Expansion canister LEDs statuses*

| Position | Color | Name | State | Meaning |
|---|---|---|---|---|
| Left | Green | Power | On | The canister is powered on. |
| | | | Off | No power available to the canister. |
| Middle | Green | Status | On | The canister is operating normally. |
| | | | Blinking | There is an error with the vital product date (VPD). |
| Right | Amber | Fault | On | There is an error logged against the canister or the system is not running (OSES). |
| | | | Blinking | Canister is being identified. |
| | | | Off | No fault, canister is operating normally. |

## 13.1.3 Dense Drawer Enclosures LED

As Figure 13-3 on page 610 shows, two 12 Gbps SAS ports are side by side on the canister of every enclosure. They are numbered 1 on the right and 2 on the left. Each Dense Drawer has two canisters side by side.

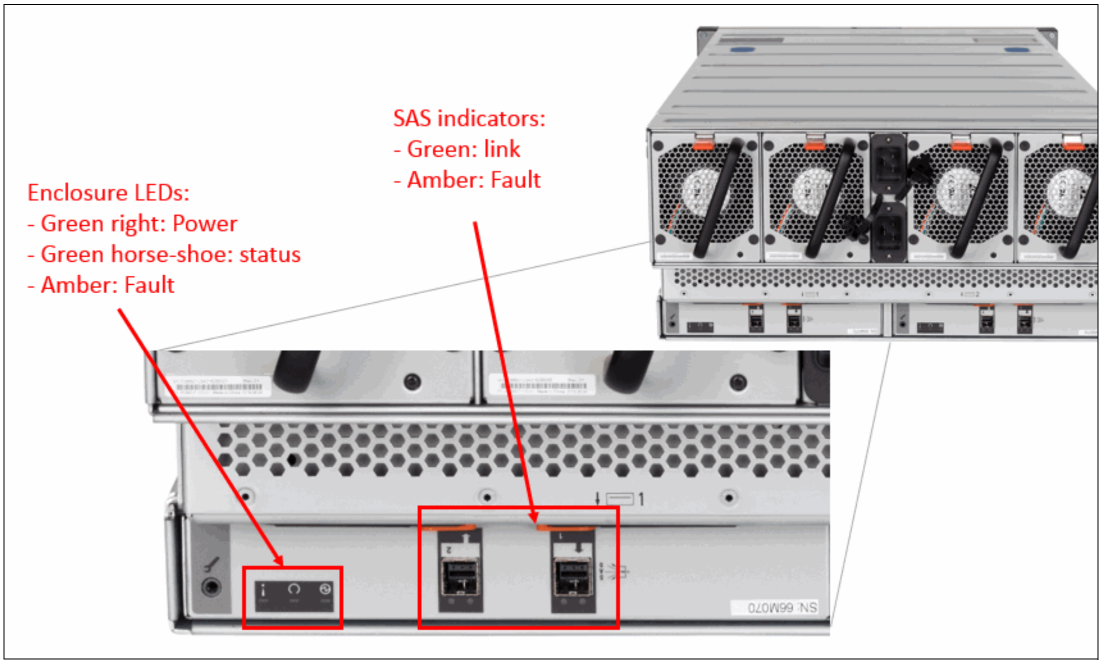


*Figure 13-3 Dense Drawer LEDs*

The interpretation of SAS status LED indicators has the same meaning as the LED indicators of SAS ports mentioned in the previous section (Table 13-5).

Table 13-5 shows the LED status values of the expansion canister.

*Table 13-6 Expansion canister LEDs statuses*

| Position | Color | Name | State | Meaning |
|---|---|---|---|---|
| Right | Green | Power | On | The canister is powered on. |
| | | | Off | No power available to the canister. |
| Middle | Green | Status | On | The canister is operating normally. |
| | | | Blinking | There is an error with the vital product date (VPD). |
| Left | Amber | Fault | On | There is an error logged against the canister or the system is not running (OSES). |
| | | | Blinking | Canister is being identified. |
| | | | Off | No fault, canister is operating normally. |

## 13.1.4 Enclosure SAS cabling

Expansion enclosures are attached to control enclosures using 12 Gbps SAS cables. IBM Storwize V7000 Gen2 control enclosure attaches up to 20 expansion enclosures or up to four Dense Drawer enclosures.

A strand starts with an SAS initiator chip inside an IBM Storwize V7000 node canister and progresses through SAS expanders, which connect disk drives. Each canister contains an expander. Each drive has two ports, each connected to a different expander and strand. This configuration ensures that both nodes in the input/output (I/O) group have direct access to each drive, and that there is no single point of failure.

Figure 13-4 shows how the SAS connectivity works inside the node and expansion canisters.

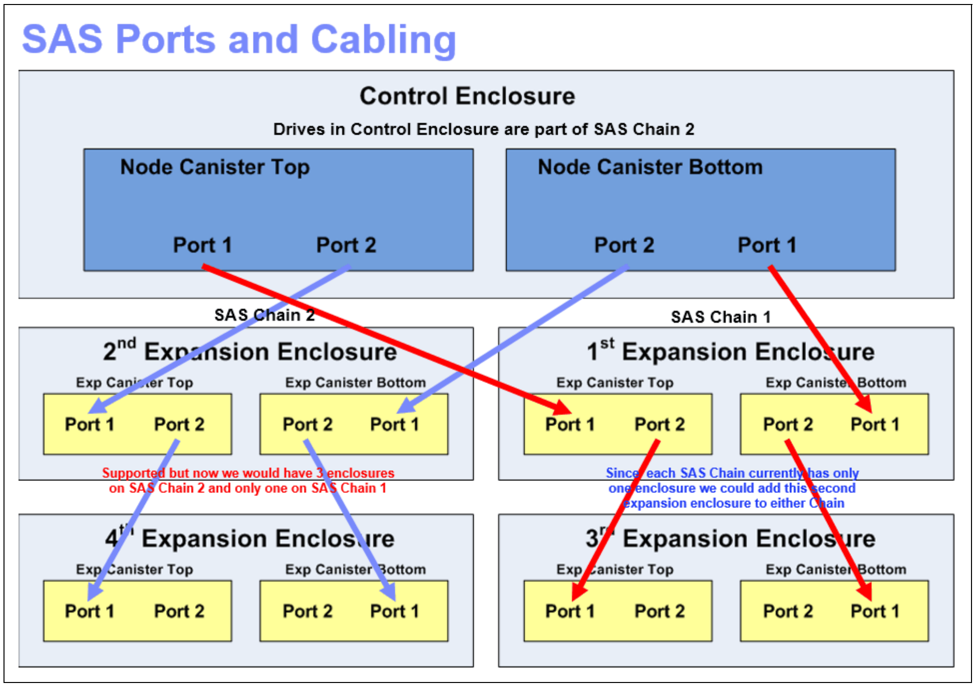


*Figure 13-4 Concept of SAS chaining*

A chain consists of a set of enclosures, correctly interconnected (Figure 13-5 on page 612). Chain 1 of an I/O group is connected to SAS port 1 of both node canisters. Chain 2 is connected to SAS port 2. This configuration means that chain 2 includes the SAS expander and drives of the control enclosure.

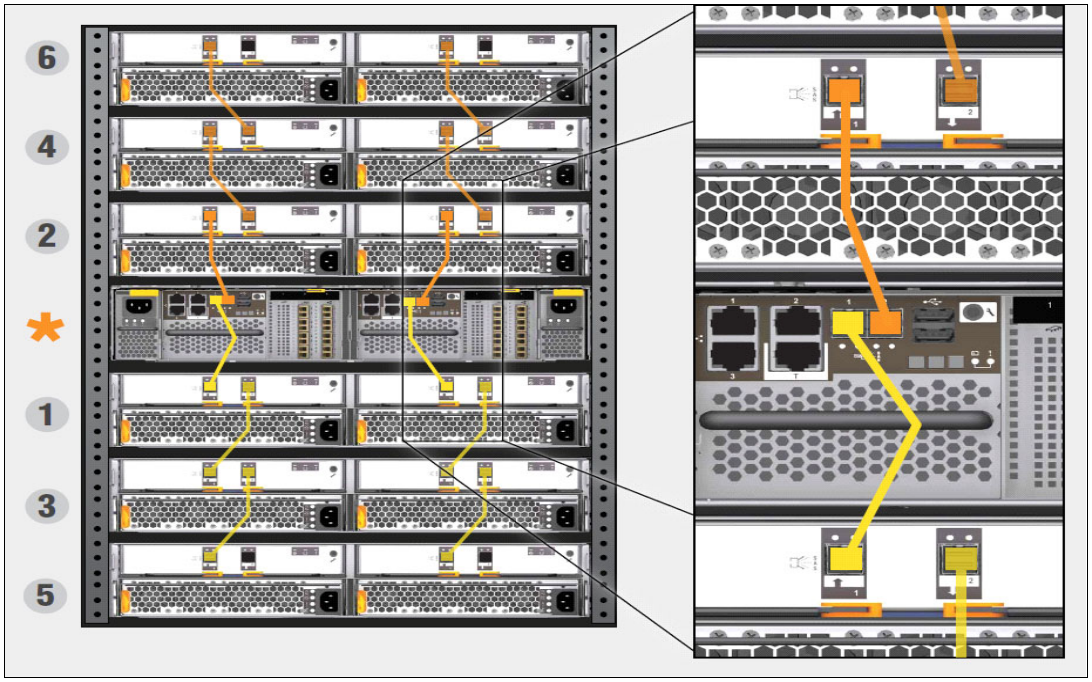


*Figure 13-5 SAS cabling with numbering of enclosures*

At system initialization, when devices are added to or removed from strands, IBM Storwize V7000 performs a discovery process to update the state of the drive and enclosure objects.

## 13.1.5 Power

All enclosures accommodate two power supply units (PSUs) for normal operation. A single PSU can supply the entire enclosure for redundancy. For this reason, it is highly advised to supply AC power to each PSU from different Power Distribution Units (PDUs).

There is a power switch on the power supply and indicator LEDs. The switch must be on for the PSU to be operational. If the power switch is turned off, the PSU stops providing power to the system. For control enclosure PSUs, the battery integrated in the node canister continues to supply power to the node. A fully charged battery is able to perform two fire hose dumps. It supports the power outage for 5 seconds before initiating safety procedures.

Figure 13-6 shows two PSUs present in the control and expansion enclosure. The controller PSU has two green and one amber indication LEDs reporting the status of the PSU.

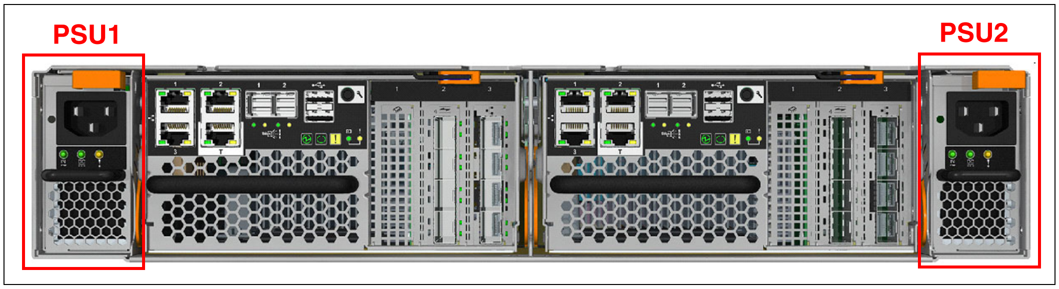


*Figure 13-6 Controller and expansion enclosure LED status indicator*

Figure 13-7 presents the rear overview of the enclosure canister with a PSU. In contrast to the control enclosure, these PSUs do not have a power switch. The enclosure is powered on by the direct attachment of a power cable.

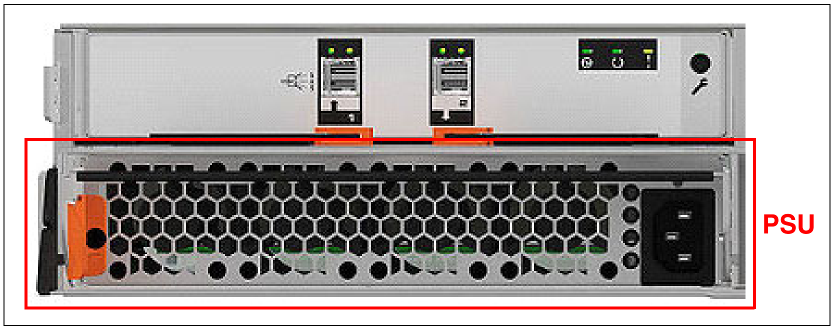


*Figure 13-7 Expansion enclosure power supply unit*

Power supplies in both control and expansion enclosures are hot-swappable and replaceable without a need to shut down a node or cluster. If the power is interrupted in one node for less than 5 seconds, the canister will not perform a fire hose dump, and continues operation from the battery. This is useful for a case of, for example, maintenance of UPS systems in the data center or replugging the power to a different power source or PDU unit. A fully charged battery is able to perform two fire hose dumps.

## 13.2 Shutting down IBM Storwize V7000

You can safely shut down an IBM Storwize V7000 system using both the GUI or the CLI.

> **Important:** Never shut down your IBM Storwize V7000 system by powering off the PSUs, removing both PSUs, or removing both power cables from a running system. It can lead to inconsistency or loss of the data staged in the cache.

Before shutting down IBM Storwize V7000, stop all hosts that have allocated volumes from the device. This step can be skipped for hosts that have volumes that are also provisioned with mirroring (host-based mirror) from different storage devices. However, doing so incurs errors that are related to lost storage paths/disks on the host error log.

You can shut down only one node canister, or you can shut down the entire cluster. When you shut down only one node canister, all activities remain active. When you shut down the entire cluster, you need to power on locally to start the system.

## 13.2.1 Shutting down a node canister

To shut down a single node canister using the GUI, complete the following steps:

1. Hover the cursor over the **Monitoring** function icon and click **System** (Figure 13-8). **Rotate** the system to the rear side by the arrow in the right-bottom corner.

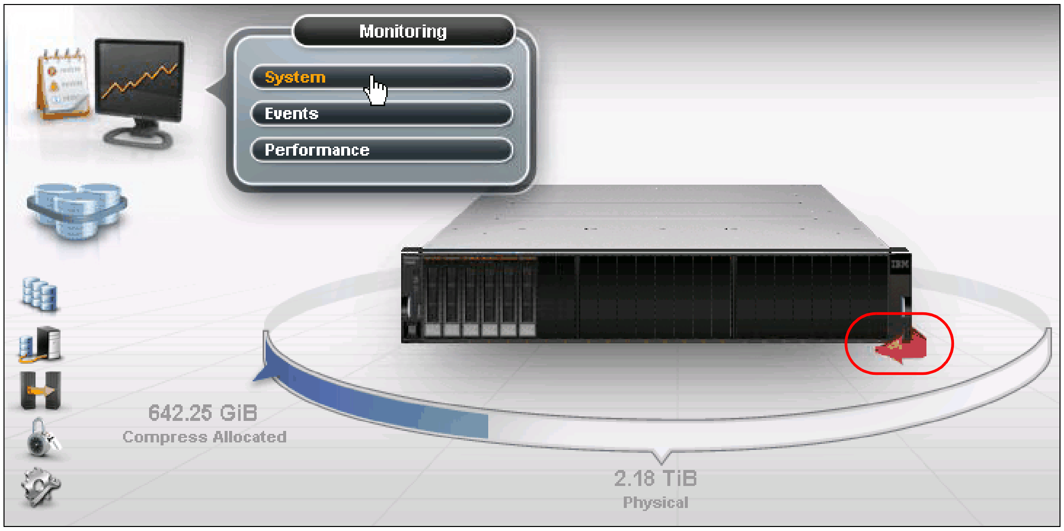
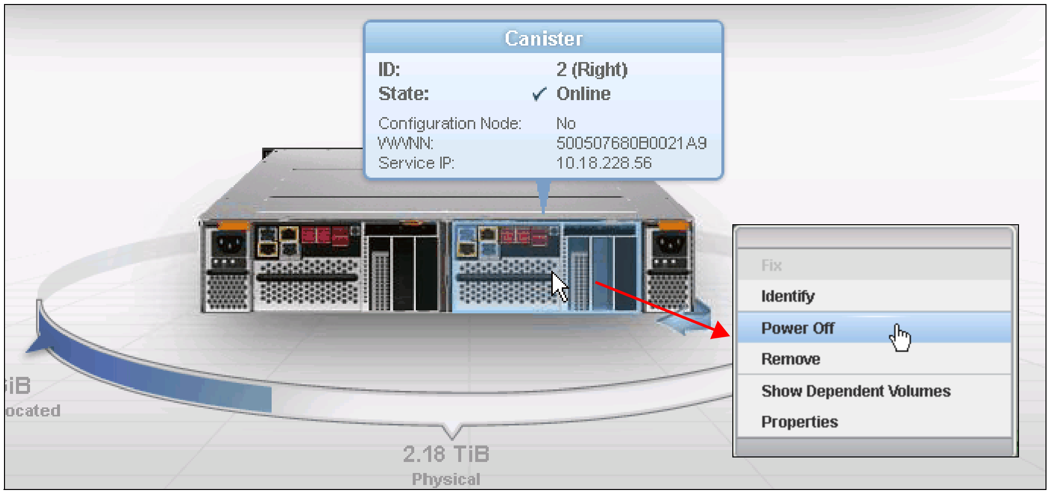


*Figure 13-8 System Device option of the Monitoring function icon*

2. Right-click the **Canisters** that you want to stop. Click **Power Off** in the opened menu (Figure 13-9).

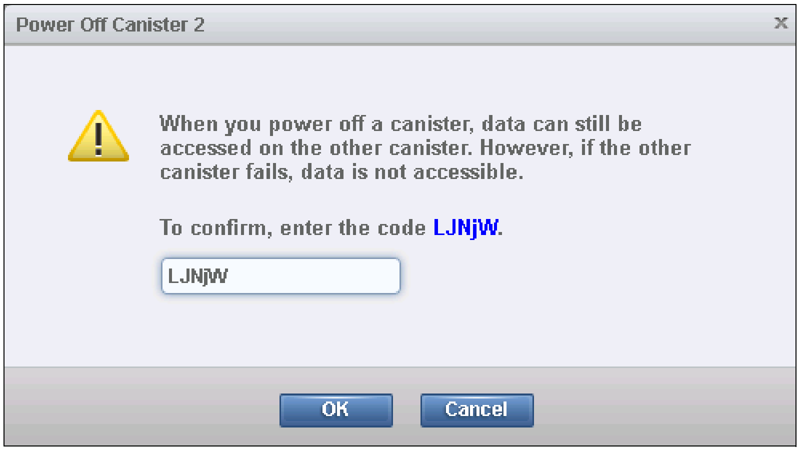


Canister

| | |
|---|---|
| ID: | 2 (Right) |
| State: | ✓ Online |
| Configuration Node: | No |
| WWNN: | 500507680B0021A9 |
| Service IP: | 10.18.228.56 |

*Figure 13-9 Shut down node canister option*

3. The confirmation window opens (Figure 13-10). Confirm whether you want to shut down the node. Type the confirmation code and click **OK**. If the node is active as a Configuration node, the control is moved automatically to the second node canister. Your session to the GUI will probably interrupt. Reestablish it again from the browser after takeover happens.



Power Off Canister 2

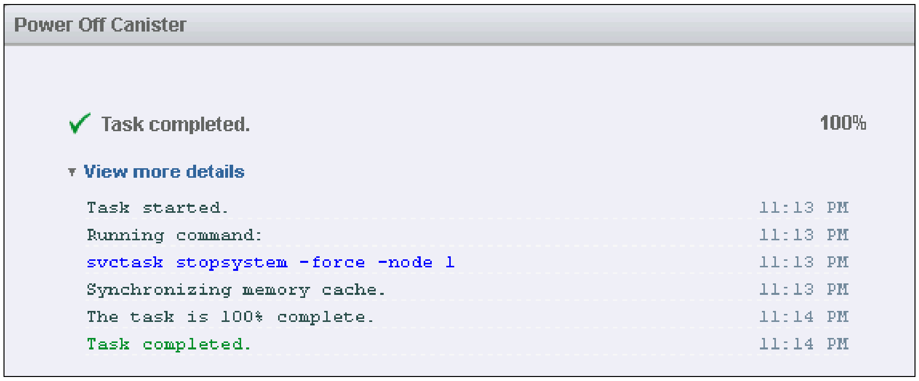
When you power off a canister, data can still be accessed on the other canister. However, if the other canister fails, data is not accessible.

To confirm, enter the code LJNjW.

LJNjW

*Figure 13-10 Confirm Shutdown window*

Shutdown is complete (Figure 13-11).



Power Off Canister

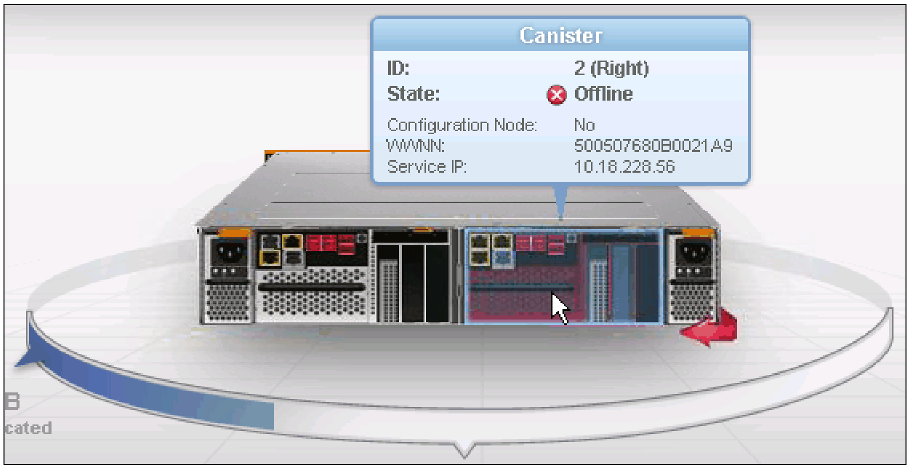
✓ Task completed. 100%

▾ View more details

```
Task started.                                  11:13 PM
Running command:                               11:13 PM
svctask stopsystem -force -node 1              11:13 PM
Synchronizing memory cache.                    11:13 PM
The task is 100% complete.                     11:14 PM
Task completed.                                11:14 PM
```

*Figure 13-11 Shutdown complete*

4. A look at the rear side of the enclosure after canister shutdown is indicated in Figure 13-12.



Canister

| | |
|---|---|
| ID: | 2 (Right) |
| State: | Offline |
| Configuration Node: | No |
| WWNN: | 500507680B0021A9 |
| Service IP: | 10.18.228.56 |

*Figure 13-12 Rear side of enclosure with powered off canister*

To shut down a node canister from the CLI, run the **`svctask stopsystem -node 2`** command.

### 13.2.2 Shutting down a system

The procedure to shut down a system is similar to shutting down a node canister:

1. Rather than rotating the enclosure and selecting a specific canister in the menu, select the whole Storwize V7000 system from the front side. Right-click and select **Power Off** (Figure 13-13).

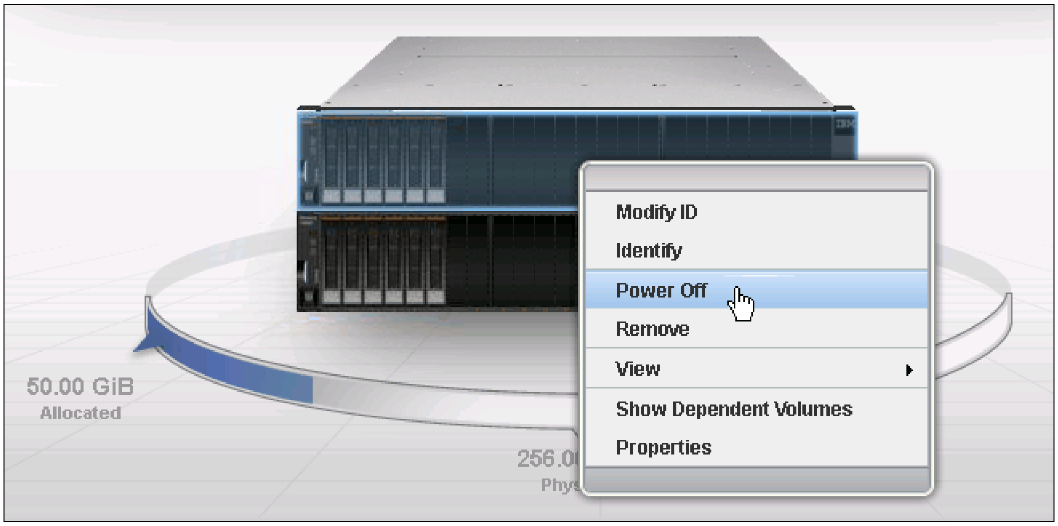
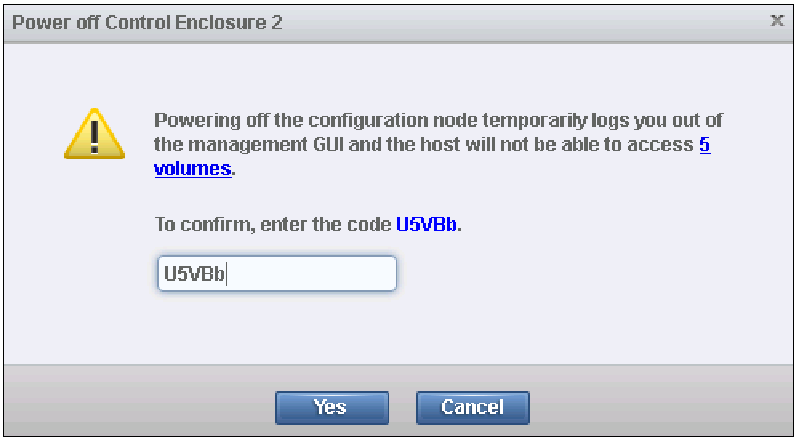


*Figure 13-13 Shutdown of IBM Storwize V7000*

2. Confirm the validity of your decision to shut down IBM Storwize V7000 clustered systems by typing the confirmation code in the pop-up window (Figure 13-14).



Powering off the configuration node temporarily logs you out of the management GUI and the host will not be able to access 5 volumes.

To confirm, enter the code U5VBb.

U5VBb

*Figure 13-14 Shutdown confirmation*

3. The whole system shutdown is typically planned in case of site maintenance (power outage, building construction, and so on), because all components of IBM Storwize V7000 are redundant and replaceable while the system is running. To start the device again after shutdown, you must have physical access to the system and then turn on the switches on the power supplies.

### 13.2.3 Shutting down and powering on an IBM Storwize V7000 infrastructure

When you shut down or power on the entire infrastructure (storage, servers, applications), follow a particular sequence for both the shutdown and the power-on actions. Here is an example sequence of a shutdown, and then a power-on, of an infrastructure that includes IBM Storwize V7000 system.

#### Shutting down

To shut down the infrastructure, complete the following steps:

1. Shut down your servers and all applications.
2. Shut down your IBM Storwize V7000 system:

   a. Shut down the cluster using either the GUI or CLI.
   b. Power off both switches of the controller enclosure.
   c. Power off both switches of all the expansion enclosures.

3. Shut down your SAN switches.

#### Powering on

To power on your infrastructure, complete the following steps:

1. Power on your SAN switches and wait until the boot completes.
2. Power on your storage systems and wait until the systems are up. When the storage systems are up, perform the following substeps in order:

   a. Power on both switches of all the expansion enclosures.
   b. Power on both switches of the controller enclosure.
   a. Power on your servers and start your applications.

## 13.3 Configuration backup

You can download and save the configuration backup file using the IBM Storwize V7000 graphical user interface (GUI) or command-line interface (CLI). On an ad hoc basis, we suggest manually doing this procedure because it is able to save the file directly to your workstation. The command-line option requires login to the system and downloading the dumped file using specific Secure Copy Protocol (SCP). The command-line option is a good practice for an automated backup of the configuration.

> **Important:** Save configuration files of IBM Storwize V7000 regularly. The best approach is to do this daily and automate this task. Always perform the additional backup before any critical maintenance task, such as an update of the Licensed Internal Code (LIC), software version, and so on.

The backup file is updated by the cluster every day. Saving it after any changes to your system configuration is also important. It contains configuration data of arrays, pools, volumes, and so on. The backup never contains any client data.

To successfully perform the configuration backup, follow the prerequisites and requirements:

- All nodes must be online.
- No independent operations that change the configuration can be running in parallel.
- No object name can begin with an underscore.
- All objects should have non-default names.

Although objects should have non-default names at the time that the backup is taken, this prerequisite is not mandatory. The backup command reports an error when the default name is discovered, but the configuration is saved. However, the default object names for controllers, I/O groups, and managed disks (MDisks) do not restore correctly if the ID of the object differs from what is recorded in the current configuration data file. All other objects with default names are renamed during the restore process.

> **Important:** Ad hoc backup of configuration can be done only from the CLI using the **`svcconfig backup`** command. Then, the output of the command can be downloaded from the GUI.

## 13.3.1 Backup using CLI

You can use CLI to trigger configuration backup either manually on an ad hoc basis, or by an automatic process regularly. The **`svcconfig backup`** command generates a new backup file. Triggering a backup using the GUI is not possible, but you can save the output from the GUI.

Example 13-1 shows output of the **`svcconfig backup`** command.

*Example 13-1 Saving configuration using CLI*

```
IBM_Storwize:ITSO_V7000Gen2_2:ITSO_admin>svcconfig backup
..................................................................................
..................................................................................
...........................................................................
CMMVC6155I SVCCONFIG processing completed successfully
IBM_Storwize:ITSO_V7000Gen2_2:ITSO_admin>
```

The **`svcconfig backup`** command generates three files that provide information about the backup process and cluster configuration. These files are dumped into the `/tmp` directory on the configuration node. Use the **`lsdumps`** command to list them. They are typically at the bottom of the list (Example 13-2).

*Example 13-2 Listing backup files in CLI*

```
IBM_Storwize:ITSO_V7000Gen2_2:ITSO_admin>lsdumps
id filename
0  snap.single.78N10WD-1.121221.164843.tgz
1  78N10WD-1.trc.old
2  dump.78N10WD-1.140826.170838
.....
32 svc.config.backup.bak_78N10WD-1
33 svc.config.backup.xml_78N10WD-1
34 svc.config.backup.sh_78N10WD-1
35 svc.config.backup.log_78N10WD-1
36 dpa_heat.78N10WD-1.141107.130950.data
IBM_Storwize:ITSO_V7000Gen2_2:ITSO_admin>
```

Table 13-7 describes the three files that are created by the backup process.

*Table 13-7 File names created by the backup process*

| File name | Description |
|---|---|
| `svc.config.backup.xml` | This file contains your cluster configuration data. |

| File name | Description |
|---|---|
| `svc.config.backup.sh` | This file contains the names of the commands that were issued to create the backup of the cluster. |
| `svc.config.backup.log` | This file contains details about the backup, including any error information that might have been reported. |

## 13.3.2 Backup using the GUI

IBM Storwize V7000 does not offer an option to initiate a backup from the GUI. However, you can download existing daily backups or those manual backups triggered from the CLI.

To download a backup of the configuration using the GUI, complete the following steps:

1. Navigate to the **Settings** icon and click **Support** (Figure 13-15).

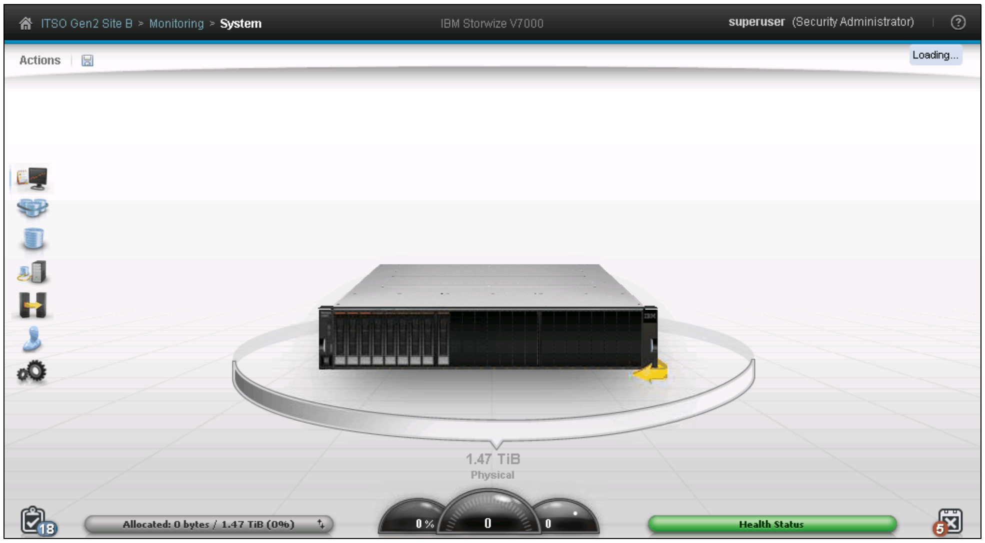
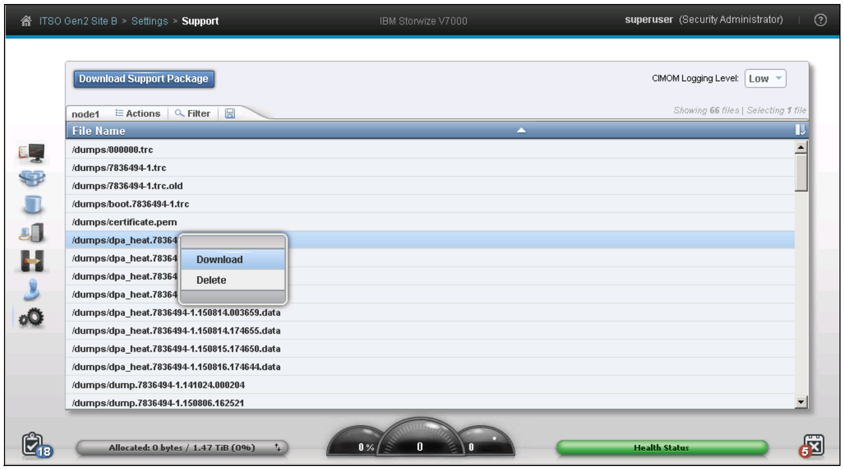

ITSO Gen2 Site B > Monitoring > System
IBM Storwize V7000
superuser (Security Administrator)
Actions
Loading...
1.47 TiB
Physical
18
Allocated: 0 bytes / 1.47 TiB (0%)
0 %
0
0
Health Status
5

*Figure 13-15 Support option*

2. The window shown in Figure 13-16 opens. Click **Show full log listing** to show all log files.
3. Search for and right-click the `/dumps/dpa_heat.<number>.<date and time>` file, and then select **Download** to transfer the file to your workstation.



Showing 66 files | Selecting 1 file

| File Name |
|---|
| /dumps/000000.trc |
| /dumps/7836494-1.trc |
| /dumps/7836494-1.trc.old |
| /dumps/boot.7836494-1.trc |
| /dumps/certificate.pem |
| /dumps/dpa_heat.78364 |
| /dumps/dpa_heat.78364 |
| /dumps/dpa_heat.78364 |
| /dumps/dpa_heat.78364 |
| /dumps/dpa_heat.7836494-1.150814.003659.data |
| /dumps/dpa_heat.7836494-1.150814.174655.data |
| /dumps/dpa_heat.7836494-1.150815.174650.data |
| /dumps/dpa_heat.7836494-1.150816.174644.data |
| /dumps/dump.7836494-1.141024.000204 |
| /dumps/dump.7836494-1.150806.162521 |

*Figure 13-16 Show full log files window*

# 13.4 Software update

In this section, we describe the operations to update your Storwize V7000 software to V7.8.

> **Note:** V7.4 and later no longer use the term *Upgrade*. Rather, they use *Update*.

The format for the software update package name ends in four positive integers that are separated by dots. For example, a software update package might have the name that is shown in the following example:

```
IBM_2145_INSTALL_7.8.0.0
```

## 13.4.1 Precautions before the update

In this section, we describe the precautions you should take before you attempt an update.

> **Important:** Before you attempt any IBM Storwize V7000 code update, read and understand the Storwize V7000 concurrent compatibility and code cross-reference matrix. For more information, see the following website and click **Latest Storwize V7000 code**:
>
> `http://www.ibm.com/support/docview.wss?uid-ssg1S1003705`

During the update, each node in your Storwize V7000 clustered system is automatically shut down and restarted by the update process. Because each node in an I/O Group provides an alternative path to volumes, use the Subsystem Device Driver (SDD) to make sure that all I/O paths between all hosts and storage area networks (SANs) work.

If you do not perform this check, certain hosts might lose connectivity to their volumes and experience I/O errors when the Storwize V7000 node that provides that access is shut down during the update process. You can check the I/O paths by using **`datapath query`** SDD commands.

### 13.4.2 IBM Storwize V7000 update test utility

The software update test utility is a Storwize V7000 software utility that checks for known issues that can cause problems during a Storwize V7000 software update. More information about the utility is available on the following website:

`https://ibm.biz/BdsFtc`

Download the software update utility from this page where you can also download the firmware. This ensures that you get the current version of this utility. You can use the **svcupgradetest** utility to check for known issues that might cause problems during a software update.

The software update test utility can be downloaded in advance of the update process, or it can be downloaded and run directly during the software update, as guided by the update wizard.

You can run the utility multiple times on the same IBM Storwize V7000 system to perform a readiness check in preparation for a software update. We strongly advise running this utility for a final time immediately before you apply the software update to ensure that there were no new releases of the utility since it was originally downloaded.

The installation and use of this utility is nondisruptive, and does not require restart of any IBM Storwize V7000 nodes. Therefore, there is no interruption to host I/O. The utility is only installed on the current configuration node.

System administrators must continue to check whether the version of code that they plan to install is the latest version. You can obtain the current information on the following website:

`https://ibm.biz/BdsEvQ`

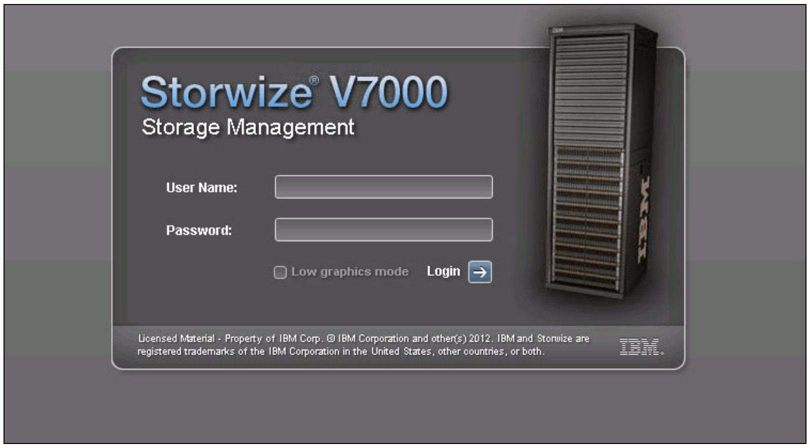
This utility is intended to supplement rather than duplicate the existing tests that are performed by the IBM Spectrum Virtualize update procedure (for example, checking for unfixed errors in the error log).

Concurrent software update of all components is supported through the standard Ethernet management interfaces. However, during the update process, most of the configuration tasks are restricted.

### 13.4.3 Update procedure to V7.8

To update the IBM Spectrum Virtualize software to release V7.8, complete the following steps:

1. Open a supported web browser and navigate to your cluster IP address. A login window opens (Figure 13-17).



Licensed Material - Property of IBM Corp. © IBM Corporation and other(s) 2012. IBM and Storwize are registered trademarks of the IBM Corporation in the United States, other countries, or both.

*Figure 13-17 IBM Storwize V7000 GUI login window*

2. Log in with superuser rights. The IBM Storwize V7000 management home window opens. Move the mouse cursor over **Settings** → **General** (Figure 13-18) and click **General**.

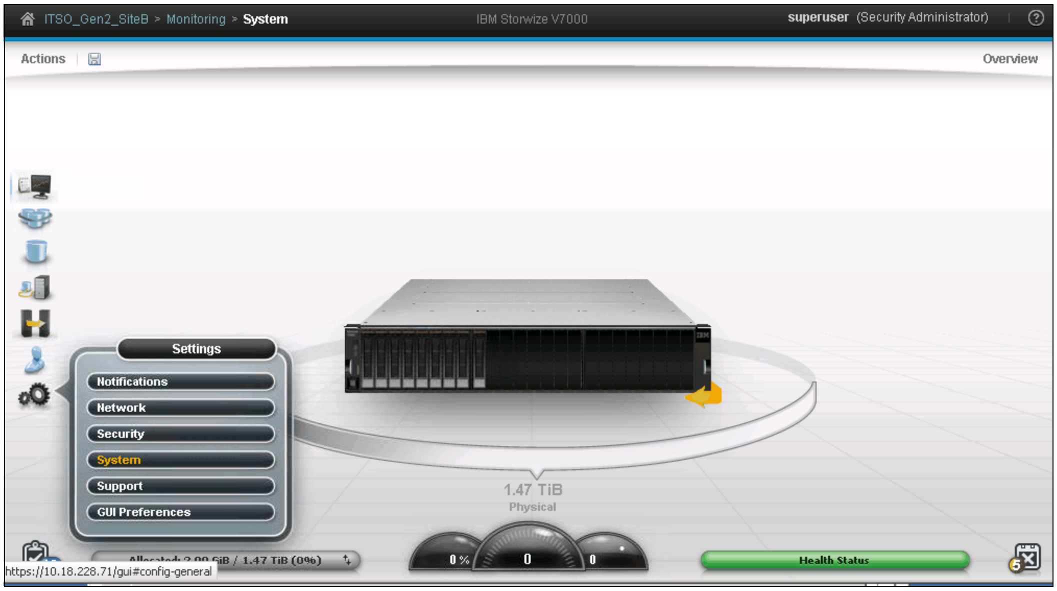
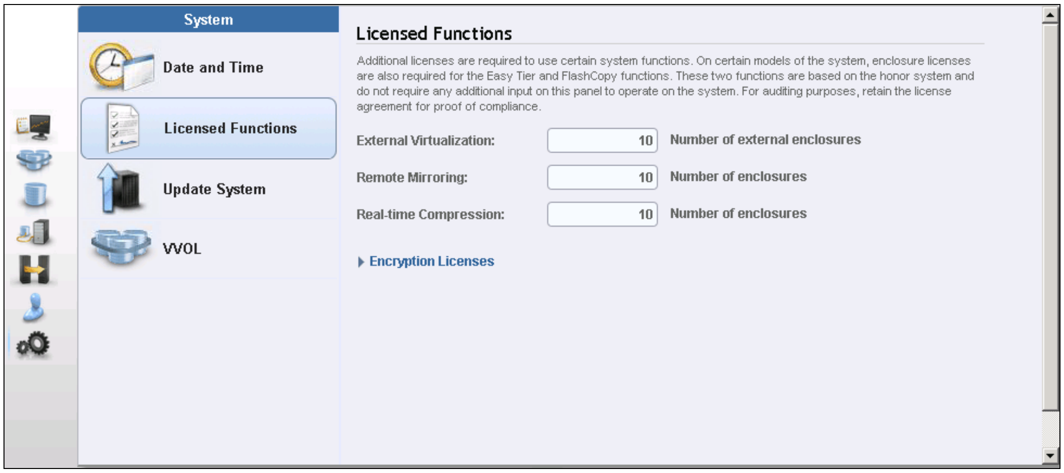

ITSO_Gen2_SiteB > Monitoring > System
IBM Storwize V7000
superuser (Security Administrator)
Actions
Overview
Settings
Notifications
Network
Security
System
Support
GUI Preferences
1.47 TiB
Physical
Allocated: 2.00 GiB / 1.47 TiB (0%)
0 %
0
0
Health Status
https://10.18.228.71/gui#config-general

*Figure 13-18 Settings menu*

3. In the Settings menu, click **System**. The Update Software pane opens (Figure 13-19).



Licensed Functions

Additional licenses are required to use certain system functions. On certain models of the system, enclosure licenses are also required for the Easy Tier and FlashCopy functions. These two functions are based on the honor system and do not require any additional input on this panel to operate on the system. For auditing purposes, retain the license agreement for proof of compliance.

External Virtualization: 10 Number of external enclosures

Remote Mirroring: 10 Number of enclosures

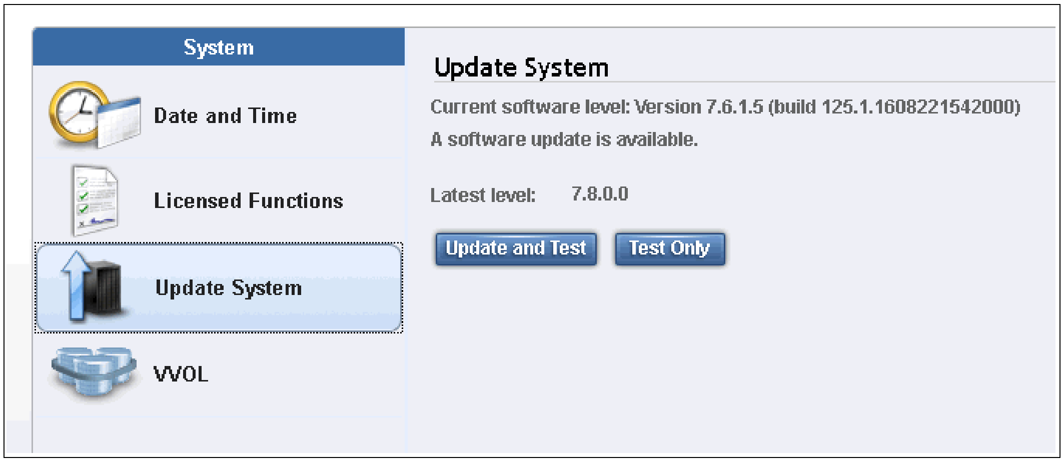
Real-time Compression: 10 Number of enclosures

Encryption Licenses

*Figure 13-19 Update Software menu*

> **My Notifications:** Use the My Notifications tool to receive notifications of new and updated support information to better maintain your system environment, especially in an environment where a direct Internet connection is not possible. Go to the following address (an IBM account is required) and add your IBM Storwize V7000 system to the notifications list to be advised of support information, and to download the current code to your workstation for later upload:
>
> `http://www.ibm.com/software/support/einfo.html`

4. Click **Update System**. You are redirected to the window shown in Figure 13-20. This function starts the software update process.



Update System

Current software level: Version 7.6.1.5 (build 125.1.1608221542000)

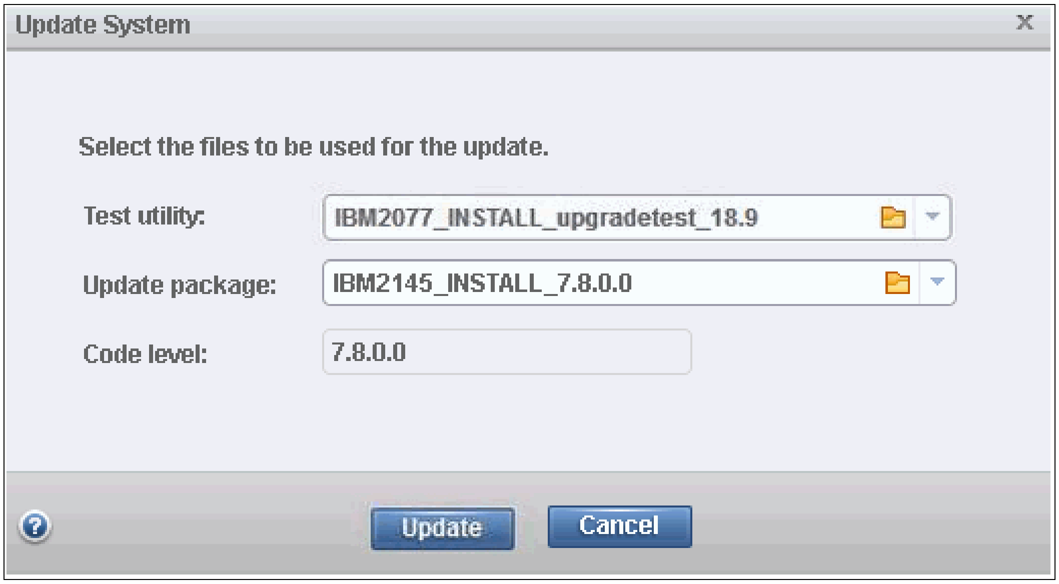
A software update is available.

Latest level: 7.8.0.0

*Figure 13-20 Update System window*

From this window, you have a new feature, which enables you to either select both the update test utility and the code update file, or simply run just the test first.

Providing that you have previously downloaded both files, you can click on the yellow folder and browse to the location where it is saved and upload them to the IBM Storwize V7000. If the files are correct, the GUI detects them as shown in Figure 13-21.



Update System

Select the files to be used for the update.

Test utility: IBM2077_INSTALL_upgradetest_18.9

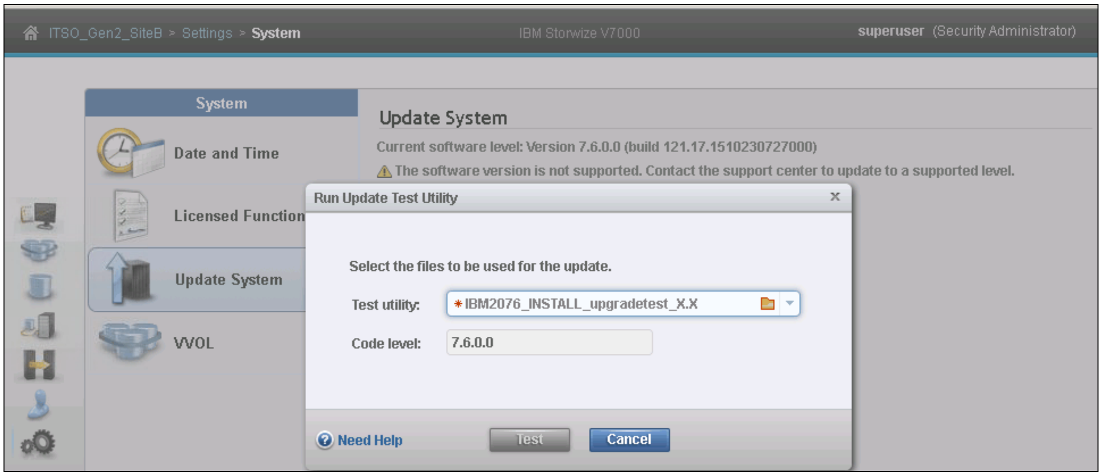
Update package: IBM2145_INSTALL_7.8.0.0

Code level: 7.8.0.0

*Figure 13-21 Upload option for both Test utility and code Update selected*

Choose for **Update** so the system can upload the files to a temporary space in the IBM Spectrum Virtualize and it will automatically starts the update procedure by running the readiness test initially.

If you have previously chosen **Test Only** as shown in Figure 13-20, the GUI prompts you to search for the test file utility as shown in Figure 13-22 and uploads to the cluster.



ITSO_Gen2_SiteB > Settings > System IBM Storwize V7000 superuser (Security Administrator)

Update System

Current software level: Version 7.6.0.0 (build 121.17.1510230727000)

⚠ The software version is not supported. Contact the support center to update to a supported level.

Run Update Test Utility

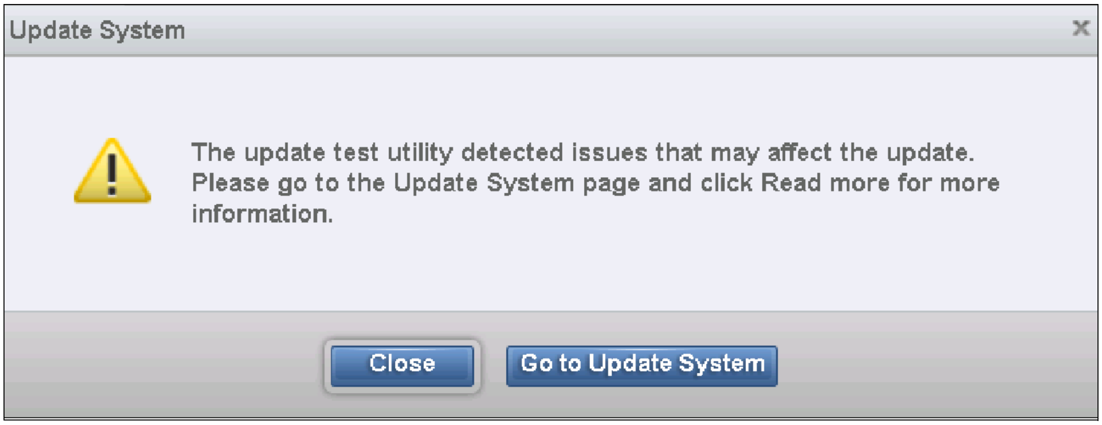
Select the files to be used for the update.

Test utility: *IBM2076_INSTALL_upgradetest_X.X

Code level: 7.6.0.0

*Figure 13-22 selecting and running the update utility test*

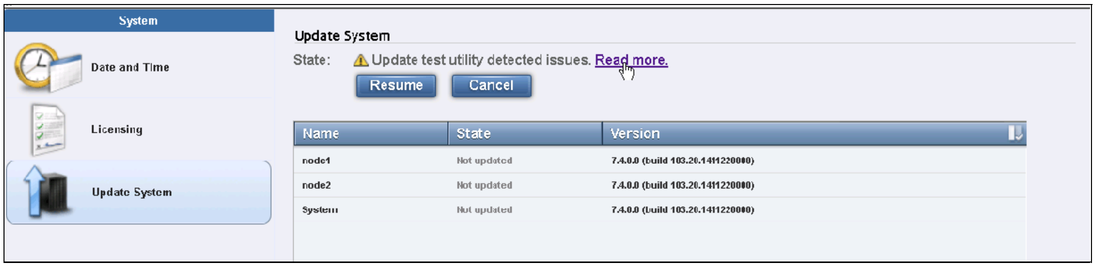
5. The software test utility runs and when the system detects an issue or an error, you are guided by the GUI. See Figure 13-23.



The update test utility detected issues that may affect the update. Please go to the Update System page and click Read more for more information.

*Figure 13-23 Issue detected*

6. Close the window. You come back to the Update System pane. Here, click **Read more** (Figure 13-24).



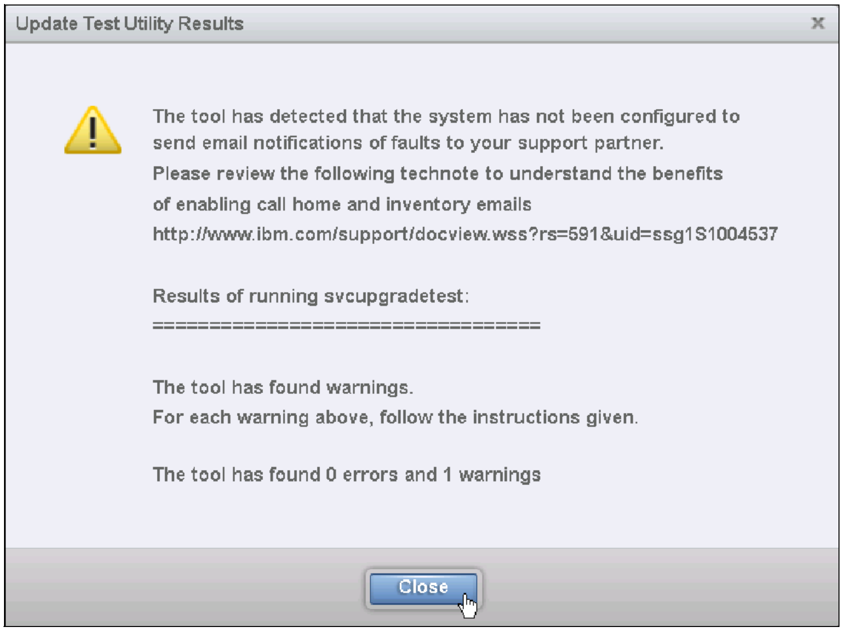
Update System

State: ⚠ Update test utility detected issues. Read more.

| Name | State | Version |
|---|---|---|
| node1 | Not updated | 7.4.0.0 (build 103.20.1411220000) |
| node2 | Not updated | 7.4.0.0 (build 103.20.1411220000) |
| System | Not updated | 7.4.0.0 (build 103.20.1411220000) |

*Figure 13-24 Issues detected by the update test utility*

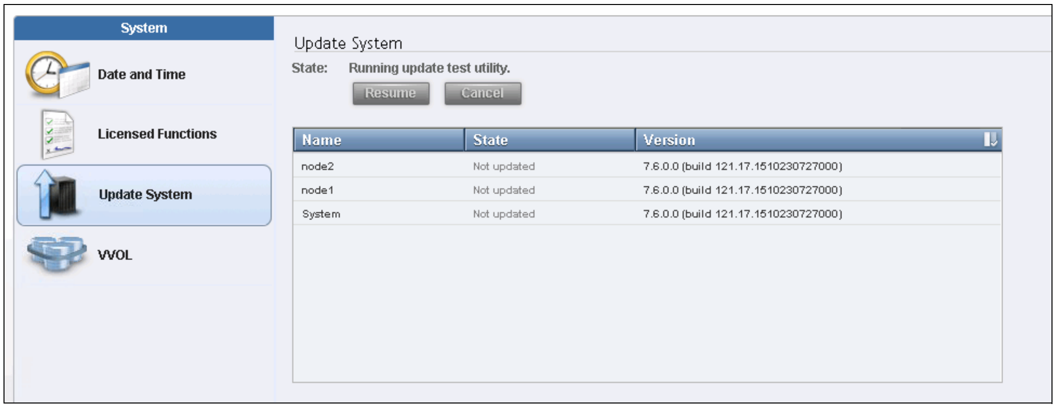
7. The results pane opens and shows you what is wrong (Figure 13-25).



```
The tool has detected that the system has not been configured to
send email notifications of faults to your support partner.
Please review the following technote to understand the benefits
of enabling call home and inventory emails
http://www.ibm.com/support/docview.wss?rs=591&uid=ssg1S1004537

Results of running svcupgradetest:
==================================

The tool has found warnings.
For each warning above, follow the instructions given.

The tool has found 0 errors and 1 warnings
```

*Figure 13-25 Description of the warning from the test utility*

8. In our case, the warning is only a warning that we did not enable email notification. Therefore, we can click **Close** and proceed with the update. For this, click **Resume** and the update proceeds, as shown in Figure 13-26.



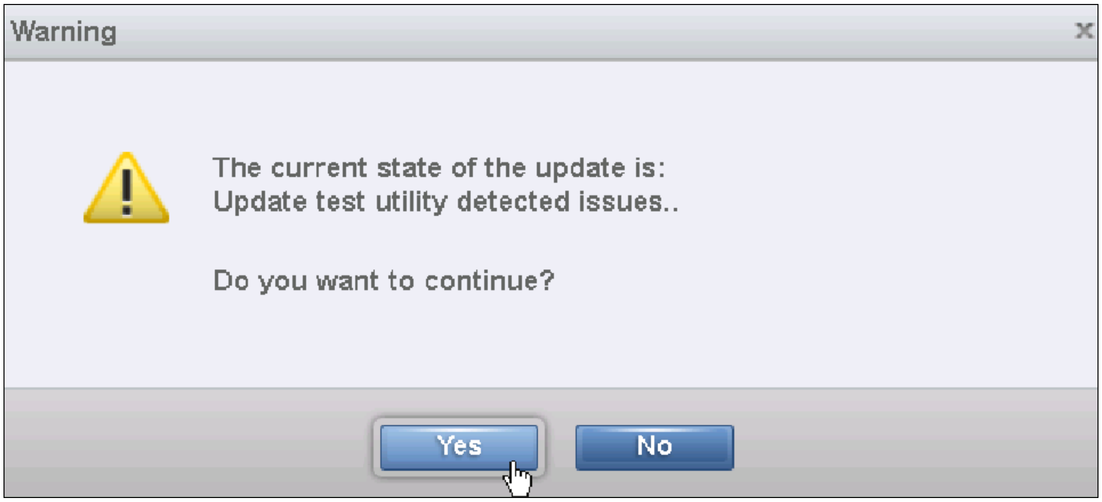
Update System

State: Running update test utility.

| Name | State | Version |
|---|---|---|
| node2 | Not updated | 7.6.0.0 (build 121.17.1510230727000) |
| node1 | Not updated | 7.6.0.0 (build 121.17.1510230727000) |
| System | Not updated | 7.6.0.0 (build 121.17.1510230727000) |

*Figure 13-26 Resuming the update utility test*

9. Due to the earlier error, another warning comes up, as shown in Figure 13-27. We proceed and click **Yes**.



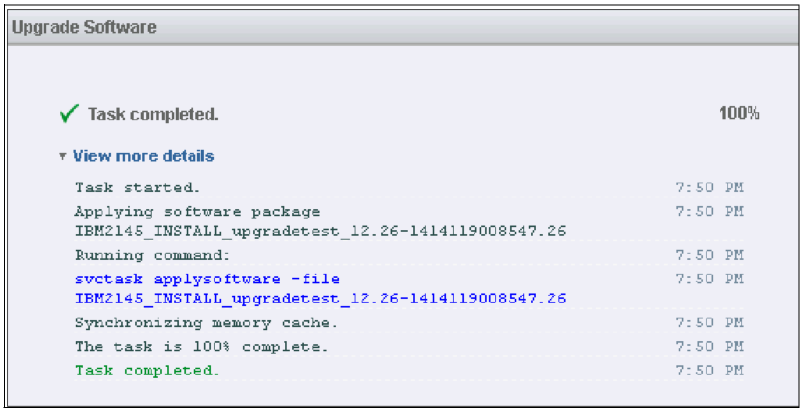
Warning

The current state of the update is:
Update test utility detected issues..

Do you want to continue?

*Figure 13-27 Warning before you can continue*

Figure 13-28 shows Update Test utility progress.



Upgrade Software

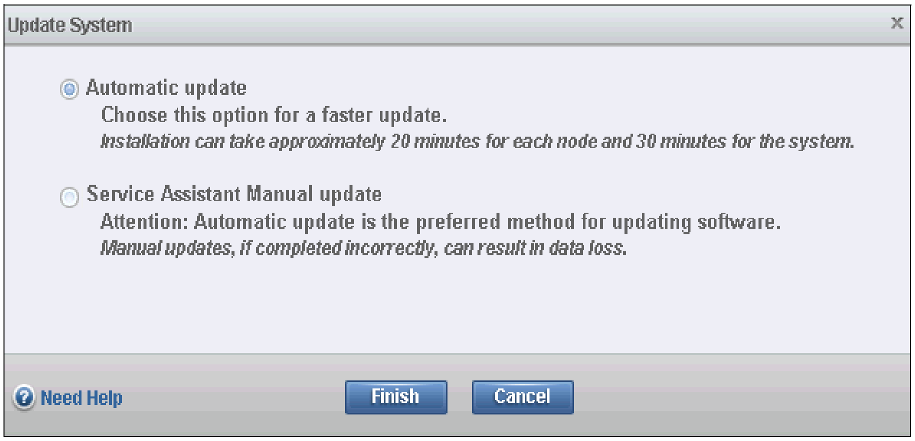
✓ Task completed. 100%

▾ View more details

```
Task started.                                                  7:50 PM
Applying software package                                      7:50 PM
IBM2145_INSTALL_upgradetest_12.26-1414119008547.26
Running command:                                               7:50 PM
svctask applysoftware -file                                    7:50 PM
IBM2145_INSTALL_upgradetest_12.26-1414119008547.26
Synchronizing memory cache.                                    7:50 PM
The task is 100% complete.                                     7:50 PM
Task completed.                                                7:50 PM
```

Figure 13-28 Update test utility completion

10.When prompted, choose **Automatic update** rather than **Service Assistant Manual update**. Manual update is eligible for the cases when the action is suggested and monitored by IBM support personnel (Figure 13-29).



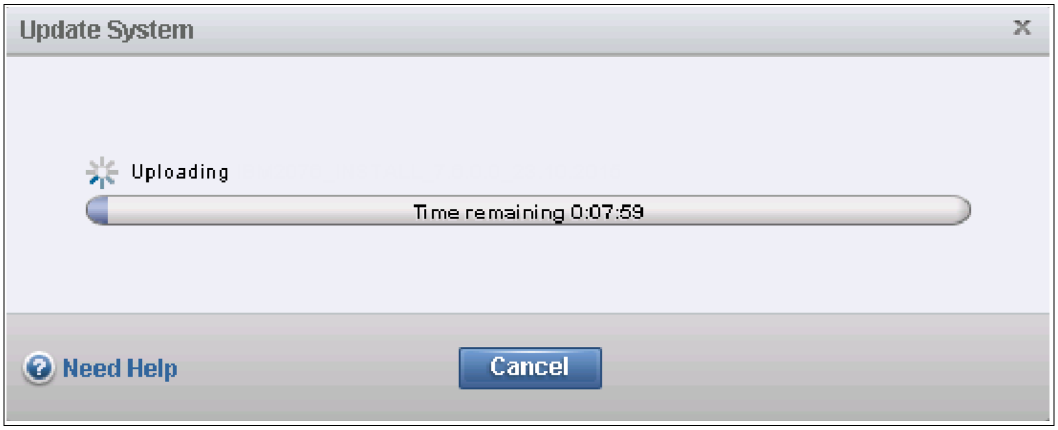
Update System

● Automatic update
Choose this option for a faster update.
*Installation can take approximately 20 minutes for each node and 30 minutes for the system.*

○ Service Assistant Manual update
Attention: Automatic update is the preferred method for updating software.
*Manual updates, if completed incorrectly, can result in data loss.*

Figure 13-29 The automatic update selection

11.Click **Finish**. The software upgrade starts. You are redirected to the window illustrated in Figure 13-30.



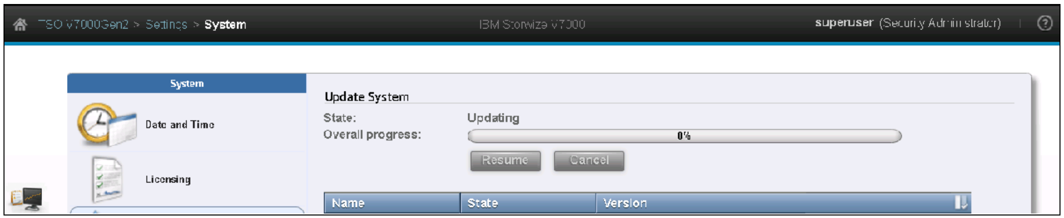
Update System

Uploading

Time remaining 0:07:59

Figure 13-30 Update software code in progress

12. Finally, now the Update process starts. Figure 13-31.



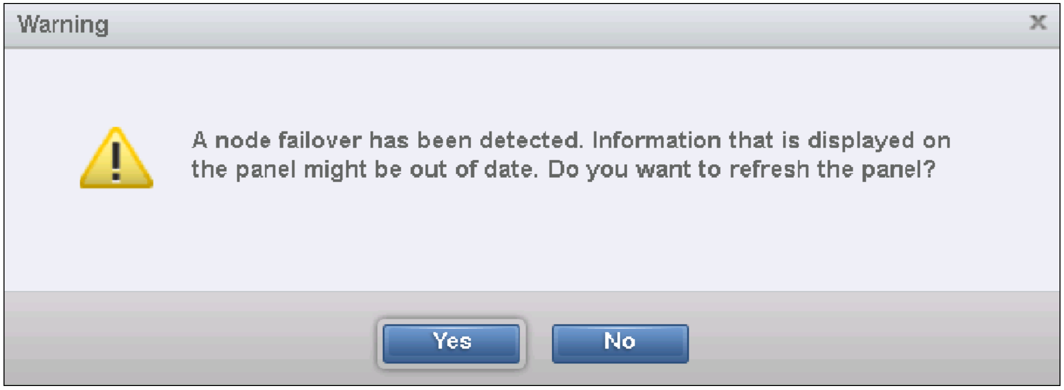
Update System

State: Updating

Overall progress: 0%

*Figure 13-31 Update process starts*

13. During the update process, a node failover occurs and you temporarily lose connection to the GUI. A window displays, prompting you to confirm that you want to refresh the current session using a connection to the cluster node that failed over, as shown in Figure 13-32.

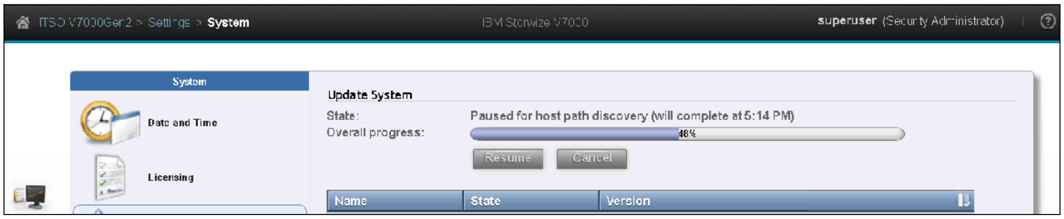


Warning

A node failover has been detected. Information that is displayed on the panel might be out of date. Do you want to refresh the panel?

*Figure 13-32 Node failover*

When the update for the first node is complete, the system pauses for awhile (approx. 30 minutes) to ensure that all paths are reestablished to the now updated node. See Figure 13-33.



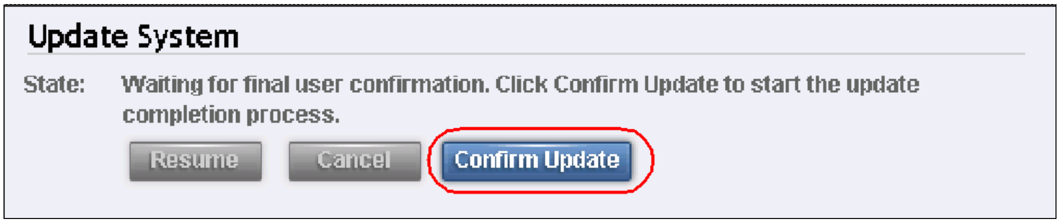
Update System

State: Paused for host path discovery (will complete at 5:14 PM)

Overall progress: 48%

*Figure 13-33 System paused to reestablish the paths*

14. When both nodes have been rebooted, the system waits for your confirmation of the update steps and commits the update. After the code load and initial update of both nodes and the system completes, the confirmation is required. Click **Confirm Update** as shown in Figure 13-34.



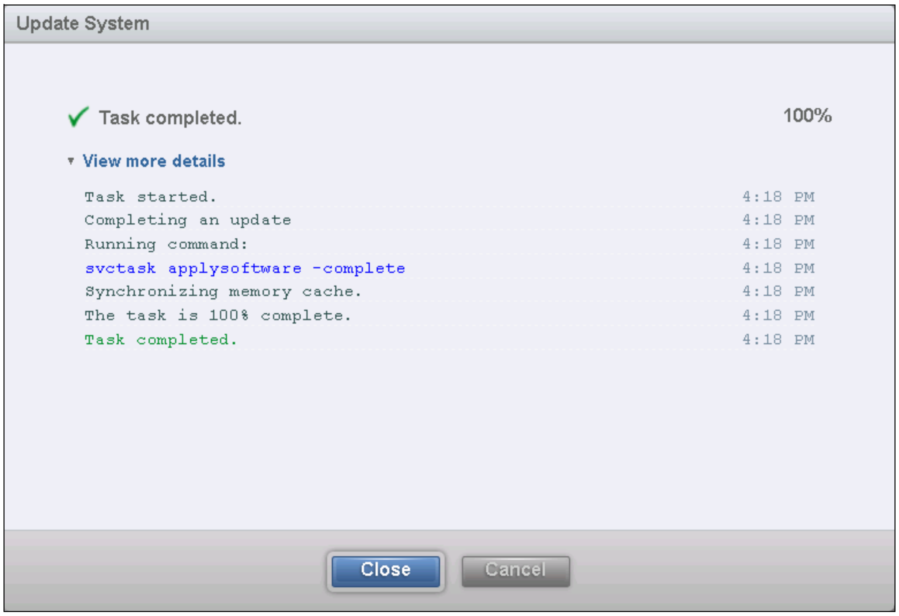
Update System

State: Waiting for final user confirmation. Click Confirm Update to start the update completion process.

*Figure 13-34 Update confirmation*

No other configuration activity is available until the update process is confirmed and committed to the system. Before confirming the update, make sure that there is no unresolved critical HW error logged in the system.

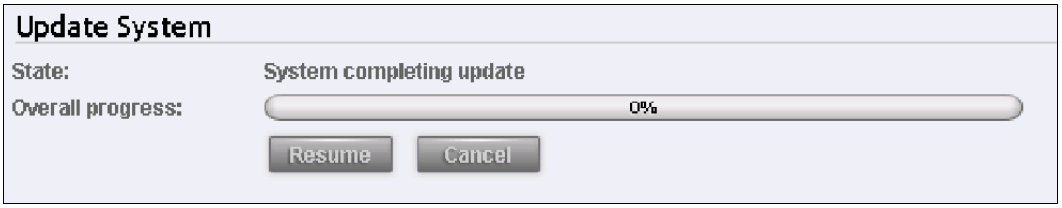
However, with IBM Spectrum Virtualize V7.6, the system allows the update even if it has recognized a failed disk drive in any expansion unit. Figure 13-35 shows update completion.



```
Task started.                               4:18 PM
Completing an update                        4:18 PM
Running command:                            4:18 PM
svctask applysoftware -complete             4:18 PM
Synchronizing memory cache.                 4:18 PM
The task is 100% complete.                  4:18 PM
Task completed.                             4:18 PM
```

*Figure 13-35 Confirmation of software update*

After successful confirmation, IBM Storwize V7000 starts committing upgrades to each node and system separately in a sequence, as shown in Figure 13-36. You might temporarily lose connection to the GUI again due to the failover.



State: System completing update
Overall progress: 0%

*Figure 13-36 Update commitment completing*

15. The update process completes when both nodes and the system unit are confirmed. The final status indicates the level of the code installed in the system.

## 13.4.4 Updating the IBM Storwize V7000 cluster manually

To update IBM Storwize V7000 software manually, complete the following steps:

1. Download the current Licensed Internal Code package and Test Utility from Fix Central and save them to your PC:

   `http://www.ibm.com/support/fixcentral`

2. Make sure the you have the PuTTY Secure Copy (PSCP) client or WinSCP File Transfer Protocol (FTP) tool. If not, download PSCP from the following link:

http://www.chiark.greenend.org.uk/~sgtatham/putty/download.html

3. Using PSCP, copy the update utility test and Licensed Internal Code files from the PC to the Storwize V7000 by issuing the following command:

   ```
   pscp -load <saved_putty_configuration>
   <directory_software_upgrade_files>\<software_upgrade_file_name>
   <username>@<cluster_ip_address>:/home/admin/upgrade
   ```

   In this example command, <*saved_putty_configuration*> is the name of the PuTTY configuration session, <*directory_software_upgrade_files*> is the location of the software upgrade files, <*software_upgrade_file_name*> is the name of the software upgrade file, <*username*> is the name that you want to use on the IBM SAN Volume Controller, and <*cluster_ip_address*> is an Internet Protocol (IP) address of your clustered system (Example 13-3).

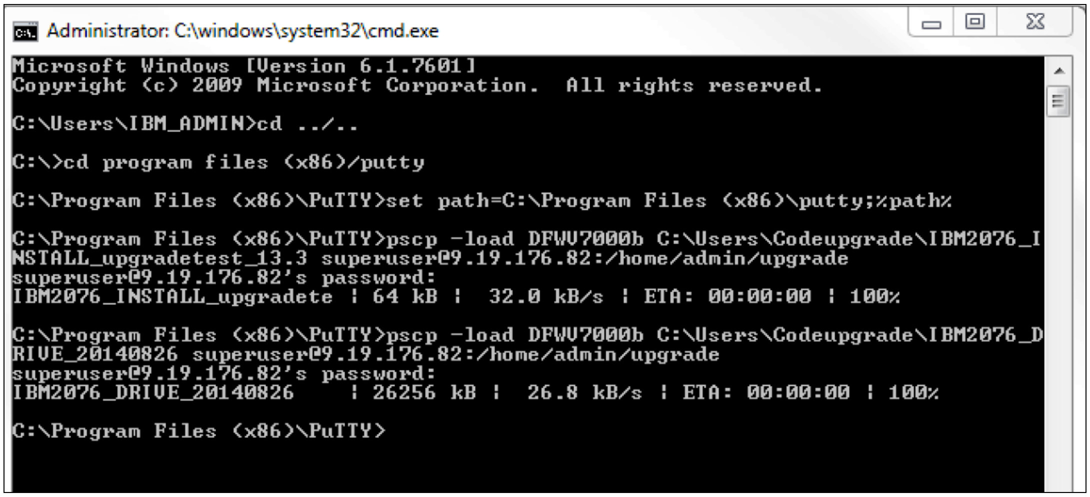
   *Example 13-3 Example PSCP command*

   ```
   pscp -load DFWV7000b C:\Users\Codeupgrade\IBM2076_DRIVE_20140826
   superuser@9.19.176.82:/home/admin/update
   ```

4. After issuing the command, you are prompted to enter the superuser's password.

   After you have successfully entered it, the program starts copying the file to the IBM Storwize V7000.

   Figure 13-37 shows an example of PSCP panel.



Administrator: C:\windows\system32\cmd.exe

```
Microsoft Windows [Version 6.1.7601]
Copyright (c) 2009 Microsoft Corporation.  All rights reserved.

C:\Users\IBM_ADMIN>cd ../..

C:\>cd program files (x86)/putty

C:\Program Files (x86)\PuTTY>set path=C:\Program Files (x86)\putty;%path%

C:\Program Files (x86)\PuTTY>pscp -load DFWV7000b C:\Users\Codeupgrade\IBM2076_I
NSTALL_upgradetest_13.3 superuser@9.19.176.82:/home/admin/upgrade
superuser@9.19.176.82's password:
IBM2076_INSTALL_upgradete | 64 kB |  32.0 kB/s | ETA: 00:00:00 | 100%

C:\Program Files (x86)\PuTTY>pscp -load DFWV7000b C:\Users\Codeupgrade\IBM2076_D
RIVE_20140826 superuser@9.19.176.82:/home/admin/upgrade
superuser@9.19.176.82's password:
IBM2076_DRIVE_20140826    | 26256 kB |  26.8 kB/s | ETA: 00:00:00 | 100%

C:\Program Files (x86)\PuTTY>
```

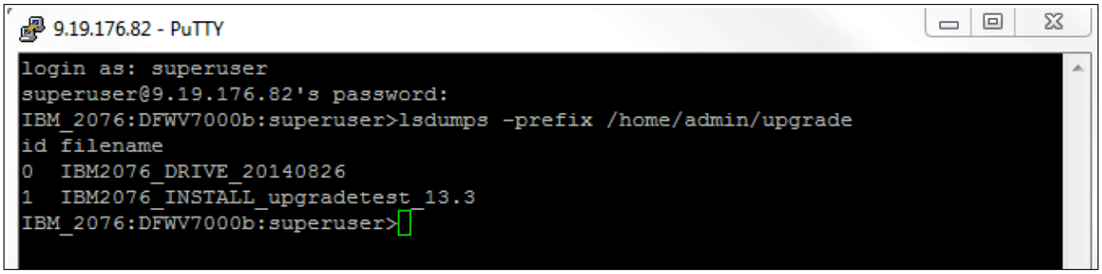
*Figure 13-37 PSCP file copy in progress @ 100% completion*

> **Note:** The Test Utility can be transferred through the GUI or CLI, but the Licensed Internal Code must be transferred from the CLI. In Figure 13-37 it shows both the Test Utility and the Licensed Internal Code.

5. After the progress bar reaches 100%, you will find the file in the target directory. To check that the file is transferred, run the following command:

   ```
   lsdumps -prefix /home/admin/update
   ```

   Figure 13-38 shows the install test.



```
login as: superuser
superuser@9.19.176.82's password:
IBM_2076:DFWV7000b:superuser>lsdumps -prefix /home/admin/upgrade
id filename
0  IBM2076_DRIVE_20140826
1  IBM2076_INSTALL_upgradetest_13.3
IBM_2076:DFWV7000b:superuser>
```

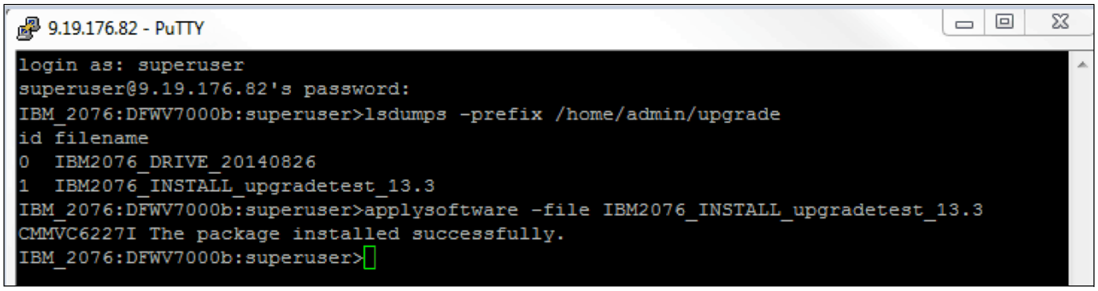
*Figure 13-38 Installing the update utility test*

6. Run the Test Utility from the CLI by issuing the command shown in Example 13-4.

   *Example 13-4 Install test*

   ```
   applysoftware -file IBM2076_INSTALL_upgradetest_13.3
   ```

   Figure 13-39 shows applying the update utility test.



```
login as: superuser
superuser@9.19.176.82's password:
IBM_2076:DFWV7000b:superuser>lsdumps -prefix /home/admin/upgrade
id filename
0  IBM2076_DRIVE_20140826
1  IBM2076_INSTALL_upgradetest_13.3
IBM_2076:DFWV7000b:superuser>applysoftware -file IBM2076_INSTALL_upgradetest_13.3
CMMVC6227I The package installed successfully.
IBM_2076:DFWV7000b:superuser>
```

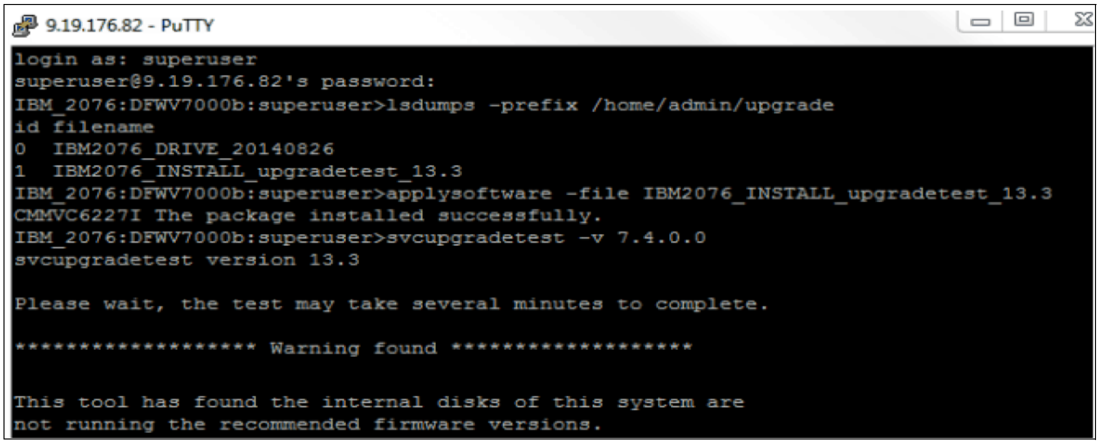
*Figure 13-39 running the update utility test*

7. Run the upgrade test utility by logging onto the Storwize V7000 CLI and running the **svcupgradetest -v <*V.R.M.F*>** command, where <*V.R.M.F*> is the version number of the new software being installed, as shown in Example 13-5.

   *Example 13-5 Run the utility*

   ```
   svcupgradetest -v7.8.0.0
   ```

   Figure 13-40 shows running update utility test.



```
login as: superuser
superuser@9.19.176.82's password:
IBM_2076:DFWV7000b:superuser>lsdumps -prefix /home/admin/upgrade
id filename
0  IBM2076_DRIVE_20140826
1  IBM2076_INSTALL_upgradetest_13.3
IBM_2076:DFWV7000b:superuser>applysoftware -file IBM2076_INSTALL_upgradetest_13.3
CMMVC6227I The package installed successfully.
IBM_2076:DFWV7000b:superuser>svcupgradetest -v 7.4.0.0
svcupgradetest version 13.3

Please wait, the test may take several minutes to complete.

******************* Warning found *******************

This tool has found the internal disks of this system are
not running the recommended firmware versions.
```

*Figure 13-40 Update utility test completion*

After transferring the update utility test and Licensed Internal Code to the cluster using PSCP, and running the update utility test, complete the rest of the manual procedure by performing the following steps:

1. Open a supported web browser and navigate to your cluster IP address. A login window opens (Figure 13-41).

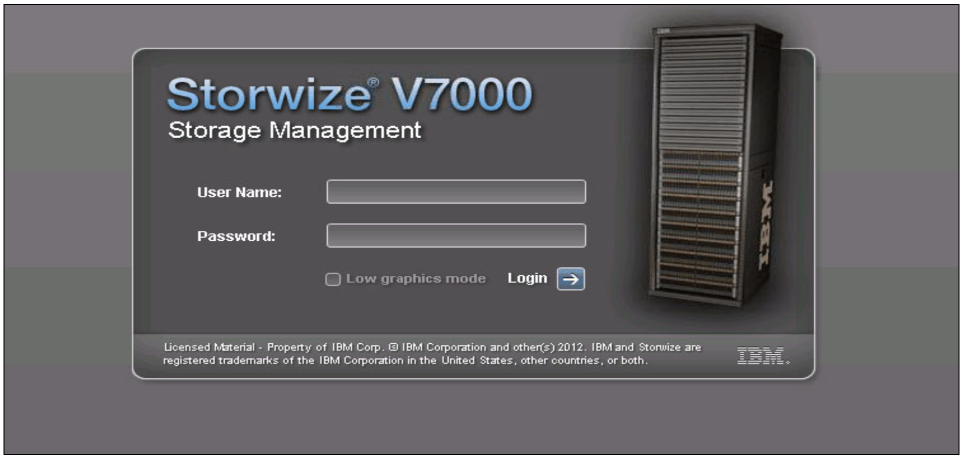

Storwize® V7000
Storage Management
User Name:
Password:
Low graphics mode
Login
Licensed Material - Property of IBM Corp. © IBM Corporation and other(s) 2012. IBM and Storwize are registered trademarks of the IBM Corporation in the United States, other countries, or both.
IBM.

*Figure 13-41 IBM Storwize V7000 GUI login window*

2. Log in with superuser rights. IBM Storwize V7000 management home window opens. Move the mouse cursor over **System** (Figure 13-42).

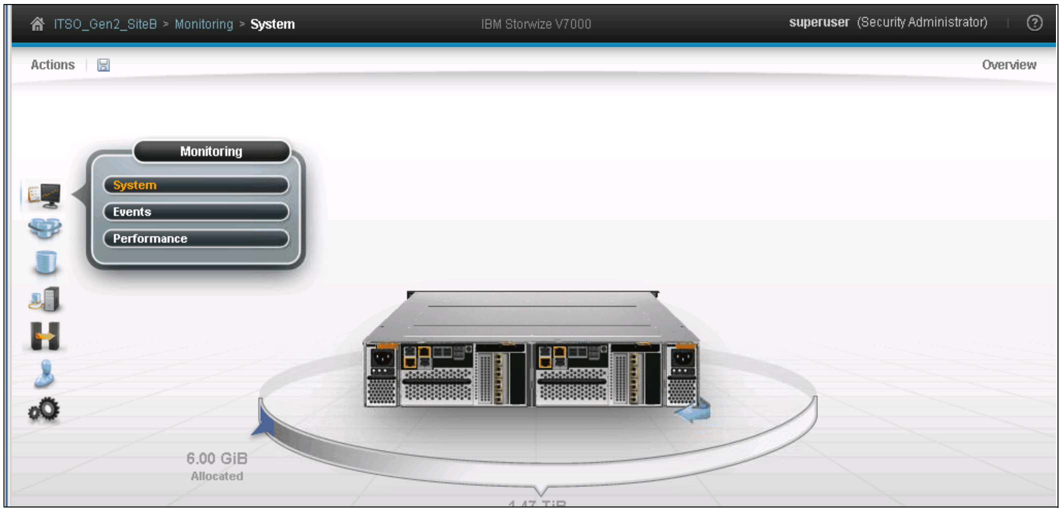


*Figure 13-42 Systems menu*

3. From the window shown in Figure 13-43, right-click the non-config node and click **Remove**.

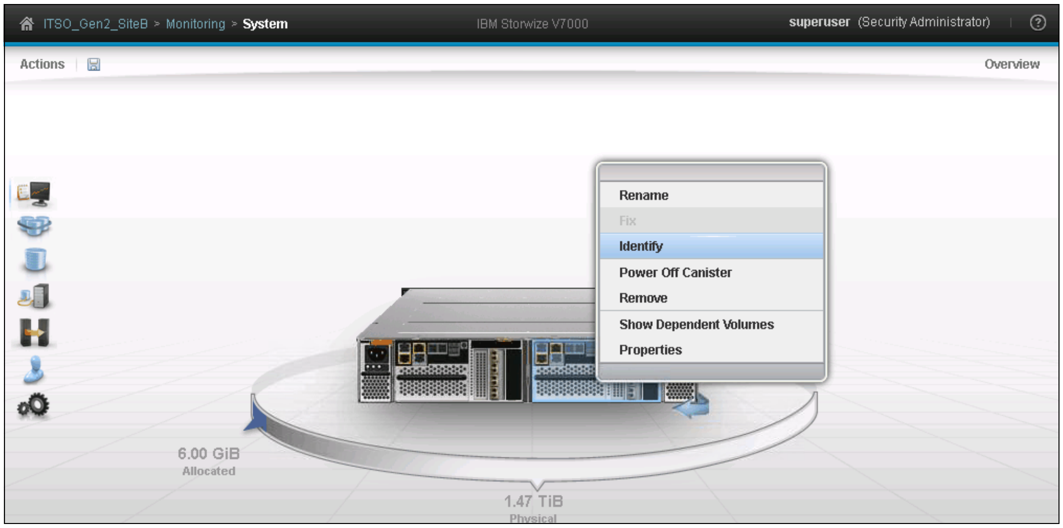


*Figure 13-43 Removing the non-configuration node*

> **Note:** The non-configuration node can also be removed using the CLI.

4. Run the **svcinfo lsnode** command from the CLI to verify the ID of the non-configuration node, as shown in Example 13-6.

*Example 13-6 Verify non-configuration node*

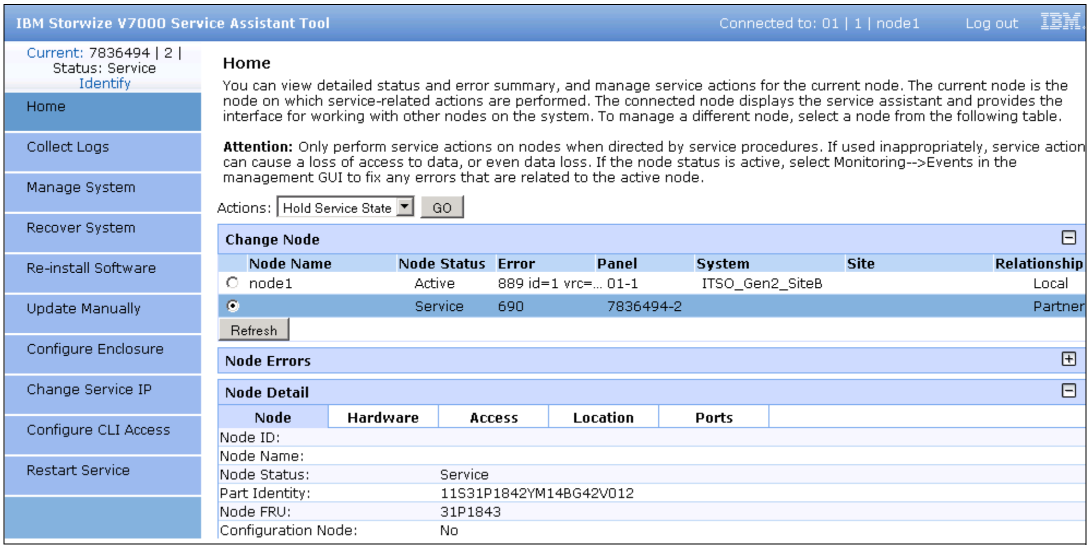
```
IBM_Storwize:ITSO_Gen2_SiteB:superuser>svcinfo lsnode
id name  UPS_serial_number WWNN             status IO_group_id IO_group_name
con         fig_node UPS_unique_id hardware iscsi_name
isc             si_alias panel_name enclosure_id canister_id
enclosure_serial_number site_id sit            e_name
1  node1                   500507680B0021A8 online 0           io_grp0
yes                                    400
iqn.1986-03.com.ibm:2145.itsogen2siteb.node1                       01-1
1            1           7836494
2 node2           500507680B0021A9 online 0      io_grp0    no
400      iqn.1986-03.com.ibm:2145.itsogen2siteb.node2
01-2       1             2           7836494
```

The output results of the **svcinfo lsnode** command shows that the non-configuration node ID is `2 node2`.

5. Remove the node by running the **rmnodecanister 2** command.

   This removes node 2 the non-configuration node, leaving it in service or candidate mode.

6. Proceed by logging in to the service GUI to view the status of the removed node, as shown in Figure 13-44.



Home

You can view detailed status and error summary, and manage service actions for the current node. The current node is the node on which service-related actions are performed. The connected node displays the service assistant and provides the interface for working with other nodes on the system. To manage a different node, select a node from the following table.

**Attention:** Only perform service actions on nodes when directed by service procedures. If used inappropriately, service action can cause a loss of access to data, or even data loss. If the node status is active, select Monitoring-->Events in the management GUI to fix any errors that are related to the active node.

Actions: Hold Service State GO

**Change Node**

| Node Name | Node Status | Error | Panel | System | Site | Relationship |
|---|---|---|---|---|---|---|
| ○ node1 | Active | 889 id=1 vrc=... | 01-1 | ITSO_Gen2_SiteB | | Local |
| ● | Service | 690 | 7836494-2 | | | Partner |

Refresh

**Node Errors**

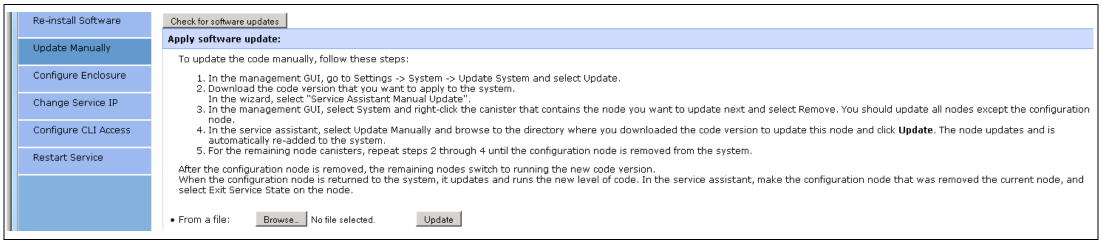
**Node Detail**

Node | Hardware | Access | Location | Ports

| | |
|---|---|
| Node ID: | |
| Node Name: | |
| Node Status: | Service |
| Part Identity: | 11S31P1842YM14BG42V012 |
| Node FRU: | 31P1843 |
| Configuration Node: | No |

*Figure 13-44 Removed node in service status*

7. With the removed node (node ID 2) selected, click **Update Manually** from the left pane, as shown in Figure 13-45.
8. Click **Browse** to locate the code that you have saved to your PC and uploaded already to the cluster via PSCP.
9. Select and upload the code, then click **Update**, as shown in Figure 13-45.



**Apply software update:**

To update the code manually, follow these steps:

1. In the management GUI, go to Settings -> System -> Update System and select Update.
2. Download the code version that you want to apply to the system.
   In the wizard, select "Service Assistant Manual Update".
3. In the management GUI, select System and right-click the canister that contains the node you want to update next and select Remove. You should update all nodes except the configuration node.
4. In the service assistant, select Update Manually and browse to the directory where you downloaded the code version to update this node and click **Update**. The node updates and is automatically re-added to the system.
5. For the remaining node canisters, repeat steps 2 through 4 until the configuration node is removed from the system.

After the configuration node is removed, the remaining nodes switch to running the new code version.
When the configuration node is returned to the system, it updates and runs the new level of code. In the service assistant, make the configuration node that was removed the current node, and select Exit Service State on the node.

- From a file: Browse... No file selected. Update

*Figure 13-45 Update*

When the code has completed updating, a message displays indicating that the software update completed, directly above the **Check for software updates** button. The node will reboot and return to candidate status.

10. Add the node back to the cluster using the cluster GUI or CLI.

    If adding using the CLI, use the **svctask addnode** command, as shown in Example 13-7.

*Example 13-7 Add node*

```
svctask addnode -panelname <panel_name> -iogrp <io_group_name>
svctask addnode -wwnodename 5005076801e08b -iogrp io_grp0
```

11. Upon completing the addition of the candidate node, verify that it has been updated to the new code.
12. When confirmed, proceed by removing the configuration node, and it becomes candidate or service.

    The newly updated node takes over as the configuration node. This keeps the cluster running, and its status changes to `candidate` because this was the original configuration node, and you do not have to apply the update manually.

13. Add the node back, and it automatically updates itself with the new code and joins the cluster.

### 13.4.5 Manually updating the Storwize V7000

This example assumes that we have an 8-node cluster of the IBM Storwize V7000 cluster, as illustrated in Table 13-8.

*Table 13-8 The iogrp*

| iogrp (O) | iogrp (1) | iogrp (2) | iogrp (3) |
|---|---|---|---|
| node 1 (config node) | node 3 | node 5 | node 7 |
| node 2 (non config) | node 4 | node 6 | node 8 |

After uploading the update utility test and Licensed Internal Code to the cluster using PSCP, and running the utility test, proceed with the following steps:

1. Start by removing node 2, which is the partner node of the configuration node in iogrp 0, using either the cluster GUI or CLI.
2. Log in to the service GUI to verify that the removed node is in `candidate` status.
3. Select the candidate node and click **Update Manually** from the left pane.
4. Browse and locate the code already downloaded and saved to your PC.
5. Upload the code and click **Update**.

   When the update is completed, a message caption indicating software update completion displays. The node will then reboot, and appear again in the service GUI after approximately 20-25 minutes in candidate status.
6. Select and verify that the node has updated to the new code.
7. Then add the node back using either the cluster GUI or the CLI.
8. After adding the node, select node 3 from iogrp1.
9. Repeat steps 1 - 7 by removing node 3, updating it manually, verifying the code, and adding it back to the cluster.
10. Proceed to node 5 in iogrp 2.
11. Repeat steps 1 - 7 by removing node 5, updating it manually, verifying the code, and adding it back to the cluster.
12. Move on to node 7 in iogrp 3.
13. Repeat steps 1 - 7 by removing node 5, updating it manually, verifying the code, and adding it back to the cluster.

> **Note:** At this point, the update is 50% completed, as we now have one node from each iogrp updated with the new code manually. Always leave the configuration node for last during a manual Licensed Internal Code update.

14. Next, select node 4 from iogrp 1.
15. Repeat steps 1 - 7 by removing node 4, updating it manually, verifying the code, and adding it back to the cluster.
16. Again, select node 6 from iogrp 2.

17. Repeat steps 1 - 7 by removing node 6, updating it manually, verifying the code, and adding it back to the cluster.
18. Next, select node 8 in iogrp 3.
19. Repeat steps 1 - 7 by removing node 8, updating it manually, verifying the code, and adding it back to the cluster.
20. Lastly, select and remove node 1, which is the configuration node in iogrp 0.

    > **Note:** The partner node, ID 2, becomes the configuration node, because the config node is removed from the cluster, keeping the cluster intact.

    The removed configuration node becomes candidate, and you do not have to apply the code update manually. Simply add the node back to the cluster. It automatically updates itself and then adds itself back to the cluster with the new code.
21. After all the nodes are updated, you must confirm the update to complete the process. The confirmation restarts each node in order and takes about 30 minutes to complete.

The update is complete.

## 13.5 Critical Fix Notification feature

The *Critical Fix Notification* function enables IBM to warn IBM Storwize V7000 and IBM SAN Volume Controller users when a critical issue exists in the level of code that they are using. The system notifies users when they log on to the GUI using a web browser connected to the Internet.

Consider the following information about this function:

- It warns users only about critical fixes, and does not warn them that they are running a previous version of the software.
- It works only if the browser also has access to the Internet (IBM Storwize V7000 and IBM SAN Volume Controller systems themselves do not need to be connected to the Internet).
- The function cannot be disabled, and each time it displays a warning, it must be acknowledged (with the option to not warn the user again for that issue).

The decision about what is a *critical* fix is subjective and requires judgment, which is exercised by the development team. As a result, clients might still encounter bugs in code that were not deemed critical. They should continue to review information about new code levels to determine if they are supposed to upgrade even without a critical fix notification.

## 13.6 Troubleshooting and fix procedures

The management GUI of IBM Storwize V7000 is a browser-based GUI for configuring and managing all aspects of your system. It provides extensive facilities to help troubleshoot and correct problems. We explain how to effectively use its features to avoid service disruption of your IBM Storwize V7000.

Figure 13-46 shows the menu to start the Monitoring menu for **System** information, viewing **Events**, or seeing real-time **Performance** statistics.

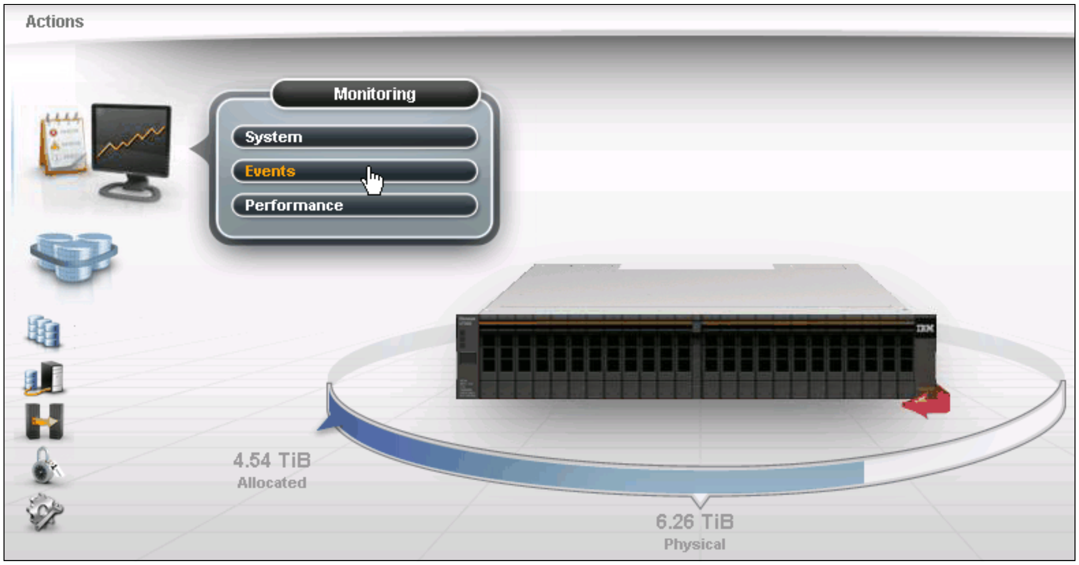


*Figure 13-46 Monitoring options*

Use the management GUI to manage and service your system. Select **Monitoring** → **Events** to access problems that must be fixed and maintenance procedures that walk you through the process of correcting problems. Information in the Events panel can be filtered in three ways:

- Recommended Actions

  Shows only the alerts that require attention. Alerts are listed in priority order and should be fixed sequentially by using the available fix procedures. For each problem that is selected, you can do these tasks:

  – Run a fix procedure
  – View the properties

- Unfixed Messages and Alerts

  Displays only the alerts and messages that are not fixed. For each entry that is selected, you can perform these tasks:

  – Run a fix procedure
  – Mark an event as fixed
  – Filter the entries to show them by specific minutes, hours, or dates
  – Reset the date filter
  – View the properties

- Show All

  Displays all event types whether they are fixed or unfixed. For each entry that is selected, you can perform these tasks:

  – Run a fix procedure
  – Mark an event as fixed
  – Filter the entries to show them by specific minutes, hours, or dates
  – Reset the date filter
  – View the properties

Some events require a certain number of occurrences in 25 hours before they are displayed as unfixed. If they do not reach this threshold in 25 hours, they are flagged as *expired*. Monitoring events are below the coalesce threshold, and are usually transient.

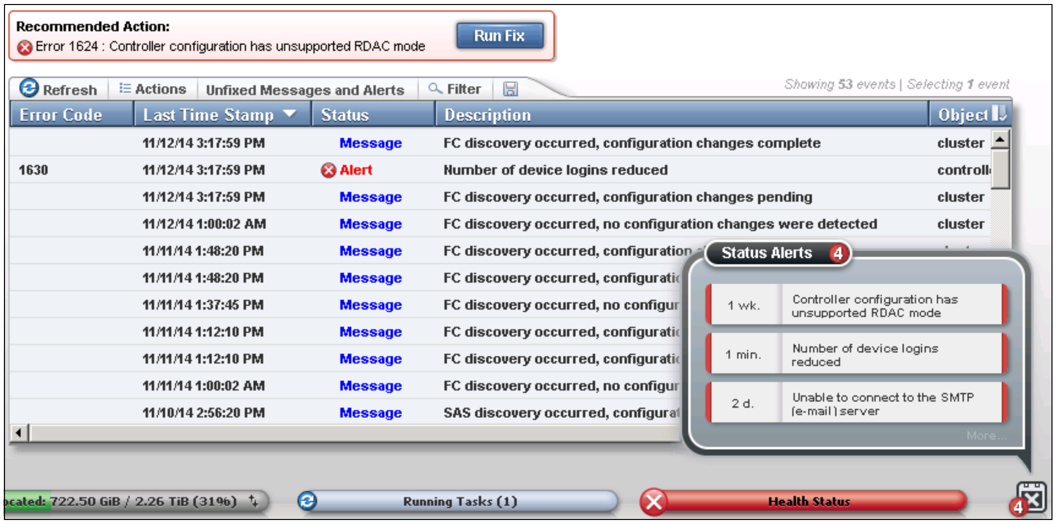
> **Important:** The management GUI is the primary tool that is used to *operate* and *service* your system. The real-time *monitoring* should be established via SNMP traps, email notifications, or syslog messaging on an automatic manner.

## 13.6.1 Managing event log

Regularly check the status of the system using the management GUI. If you suspect a problem, first use the management GUI to diagnose and resolve the problem.

Use the views that are available in the management GUI to verify the status of the system, the hardware devices, the physical storage, and the available volumes:

1. Select **Monitoring** → **Events** to see all problems that exist on the system (Figure 13-47).



Recommended Action:
Error 1624 : Controller configuration has unsupported RDAC mode

Showing 53 events | Selecting 1 event

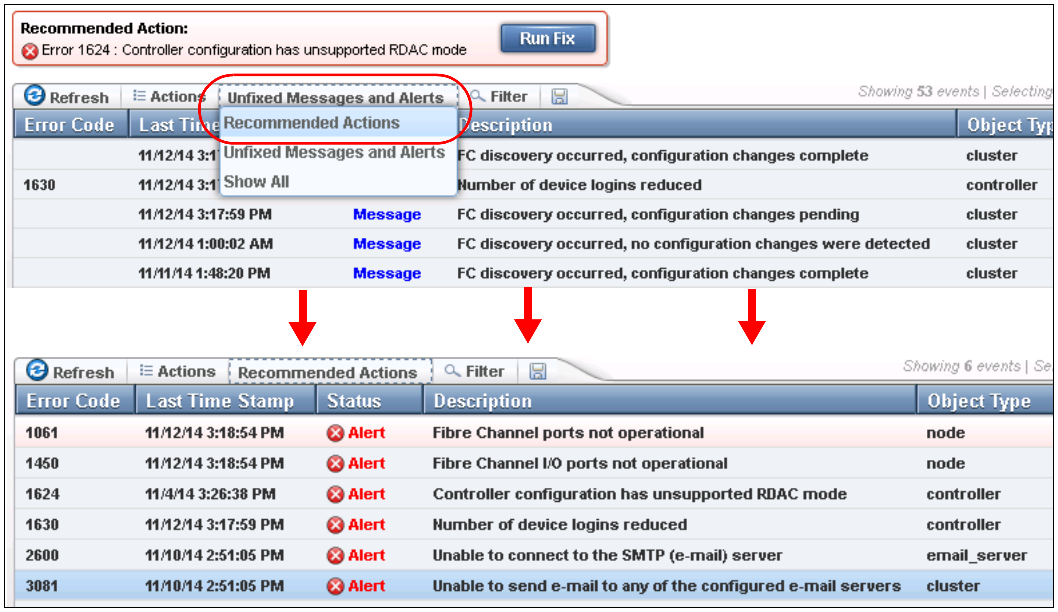
| Error Code | Last Time Stamp | Status | Description | Object |
|---|---|---|---|---|
| | 11/12/14 3:17:59 PM | Message | FC discovery occurred, configuration changes complete | cluster |
| 1630 | 11/12/14 3:17:59 PM | Alert | Number of device logins reduced | controll |
| | 11/12/14 3:17:59 PM | Message | FC discovery occurred, configuration changes pending | cluster |
| | 11/12/14 1:00:02 AM | Message | FC discovery occurred, no configuration changes were detected | cluster |
| | 11/11/14 1:48:20 PM | Message | FC discovery occurred, configuration | |
| | 11/11/14 1:48:20 PM | Message | FC discovery occurred, configuratio | |
| | 11/11/14 1:37:45 PM | Message | FC discovery occurred, no configur | |
| | 11/11/14 1:12:10 PM | Message | FC discovery occurred, configuratio | |
| | 11/11/14 1:12:10 PM | Message | FC discovery occurred, configuratio | |
| | 11/11/14 1:00:02 AM | Message | FC discovery occurred, no configur | |
| | 11/10/14 2:56:20 PM | Message | SAS discovery occurred, configurat | |

Status Alerts 4

| | |
|---|---|
| 1 wk. | Controller configuration has unsupported RDAC mode |
| 1 min. | Number of device logins reduced |
| 2 d. | Unable to connect to the SMTP (e-mail) server |

*Figure 13-47 Messages in the event log*

2. Select **Show All** → **Recommended Actions** to display the most important events to be resolved (Figure 13-48). The Recommended Actions tab shows the highest priority maintenance procedure that must be run. Use the troubleshooting wizard so that IBM Storwize V7000 system can determine the proper order of maintenance procedures.

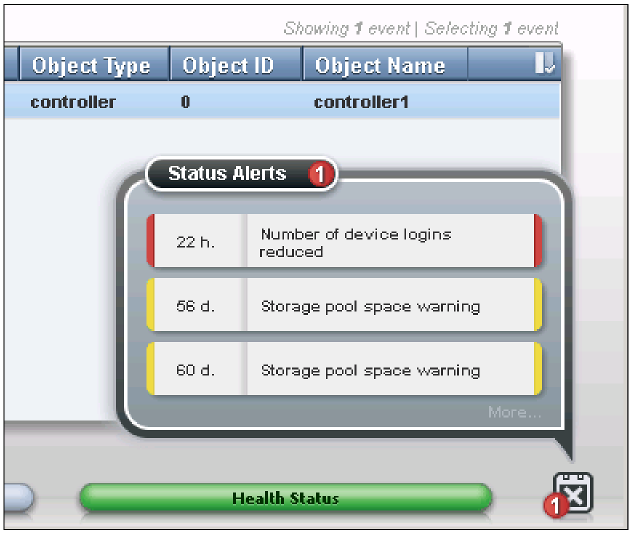

Recommended Action:
Error 1624 : Controller configuration has unsupported RDAC mode
Run Fix
Refresh
Actions
Unfixed Messages and Alerts
Filter
Showing 53 events | Selecting
Recommended Actions
Unfixed Messages and Alerts
Show All
Refresh
Actions
Recommended Actions
Filter
Showing 6 events | Se

| Error Code | Last Time | | Description | Object Typ |
|---|---|---|---|---|
| | 11/12/14 3:1 | | FC discovery occurred, configuration changes complete | cluster |
| 1630 | 11/12/14 3:1 | | Number of device logins reduced | controller |
| | 11/12/14 3:17:59 PM | Message | FC discovery occurred, configuration changes pending | cluster |
| | 11/12/14 1:00:02 AM | Message | FC discovery occurred, no configuration changes were detected | cluster |
| | 11/11/14 1:48:20 PM | Message | FC discovery occurred, configuration changes complete | cluster |

| Error Code | Last Time Stamp | Status | Description | Object Type |
|---|---|---|---|---|
| 1061 | 11/12/14 3:18:54 PM | Alert | Fibre Channel ports not operational | node |
| 1450 | 11/12/14 3:18:54 PM | Alert | Fibre Channel I/O ports not operational | node |
| 1624 | 11/4/14 3:26:38 PM | Alert | Controller configuration has unsupported RDAC mode | controller |
| 1630 | 11/12/14 3:17:59 PM | Alert | Number of device logins reduced | controller |
| 2600 | 11/10/14 2:51:05 PM | Alert | Unable to connect to the SMTP (e-mail) server | email_server |
| 3081 | 11/10/14 2:51:05 PM | Alert | Unable to send e-mail to any of the configured e-mail servers | cluster |

*Figure 13-48 Recommended Actions*

In this example, the *number of device logins reduced* is listed (service error code 1630). Review the physical FC cabling to determine the issue and then click **Run Fix**. At any time and from any GUI panel, you can directly navigate to this menu by using the Status Alerts icon at the lower right corner of the GUI (Figure 13-49).



| Object Type | Object ID | Object Name |
|---|---|---|
| controller | 0 | controller1 |

| | |
|---|---|
| 22 h. | Number of device logins reduced |
| 56 d. | Storage pool space warning |
| 60 d. | Storage pool space warning |

*Figure 13-49 Status alerts*

## 13.6.2 Running a fix procedure

If there is a service error code for the alert, you can run a fix procedure that assists you in resolving the problem. These fix procedures analyze the system and provide more information about the problem. They suggest actions to take and walk you through the actions that automatically manage the system where necessary. Finally, they check that the problem is resolved.

If an error is reported, always use the fix procedures from the management GUI to resolve the problem. Always use the fix procedures for both software configuration problems and hardware failures. The fix procedures analyze the system to ensure that the required changes do not cause volumes to be inaccessible to the hosts. The fix procedures automatically perform configuration changes that are required to return the system to its optimum state.

> **Good practice:** Before running a fix procedure for the most critical errors, take a backup of the system configuration, as suggested in 13.3, “Configuration backup” on page 617.

The fix procedure displays information that is relevant to the problem, and provides various options to correct the problem. Where possible, the fix procedure runs the commands that are required to reconfigure the system.

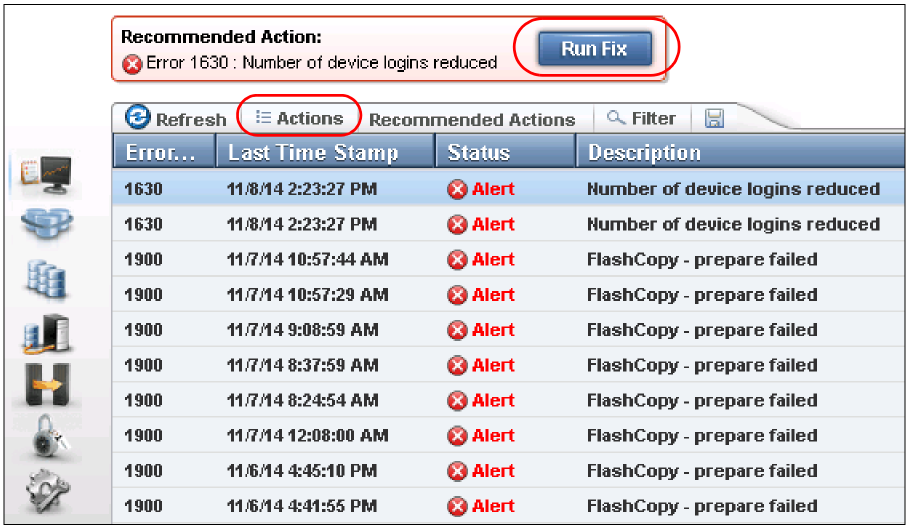
Always use the recommended action for an alert, because these actions ensure that all required steps are performed. Use the recommended actions even in cases where the service action seems obvious, such as a drive showing a fault. In this case, the drive must be replaced and reconfiguration must be performed. The fix procedure performs the reconfiguration for you.

The fix procedure also checks that another existing problem does not result in a fix procedure that causes volume data to be lost. For example, if a power supply unit in a node enclosure must be replaced, the fix procedure checks and warns you if the integrated battery in the other power supply unit is not sufficiently charged to protect the system.

> **Hint:** If possible, fix the alerts in the order shown to resolve the most serious issues first. Often, other alerts are corrected automatically, because they were the result of a more serious issue.

The following example demonstrates how to clear the error that is related to the malfunctioning FC connectivity between control canisters of IBM Storwize V7000:

1. From the dynamic menu (the icons on the left), select **Monitoring** → **Events**, and then focus on the errors with the highest priority first. List only the recommended actions by selecting the filters in the **Actions** menu (Figure 13-50). Click **Run Fix**.

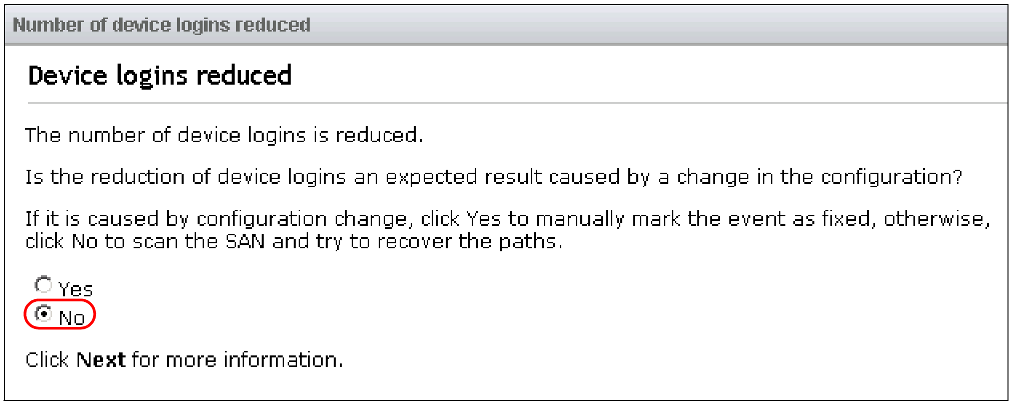


**Recommended Action:**
Error 1630 : Number of device logins reduced

| Error... | Last Time Stamp | Status | Description |
|---|---|---|---|
| 1630 | 11/8/14 2:23:27 PM | Alert | Number of device logins reduced |
| 1630 | 11/8/14 2:23:27 PM | Alert | Number of device logins reduced |
| 1900 | 11/7/14 10:57:44 AM | Alert | FlashCopy - prepare failed |
| 1900 | 11/7/14 10:57:29 AM | Alert | FlashCopy - prepare failed |
| 1900 | 11/7/14 9:08:59 AM | Alert | FlashCopy - prepare failed |
| 1900 | 11/7/14 8:37:59 AM | Alert | FlashCopy - prepare failed |
| 1900 | 11/7/14 8:24:54 AM | Alert | FlashCopy - prepare failed |
| 1900 | 11/7/14 12:08:00 AM | Alert | FlashCopy - prepare failed |
| 1900 | 11/6/14 4:45:10 PM | Alert | FlashCopy - prepare failed |
| 1900 | 11/6/14 4:41:55 PM | Alert | FlashCopy - prepare failed |

*Figure 13-50 Initiate Run Fix procedure from the management GUI*

2. The pop-up window prompts you to indicate whether the issue was caused by a planned change or maintenance task, or whether it appeared in an uncontrolled manner (Figure 13-51).



Number of device logins reduced

Device logins reduced

The number of device logins is reduced.

Is the reduction of device logins an expected result caused by a change in the configuration?

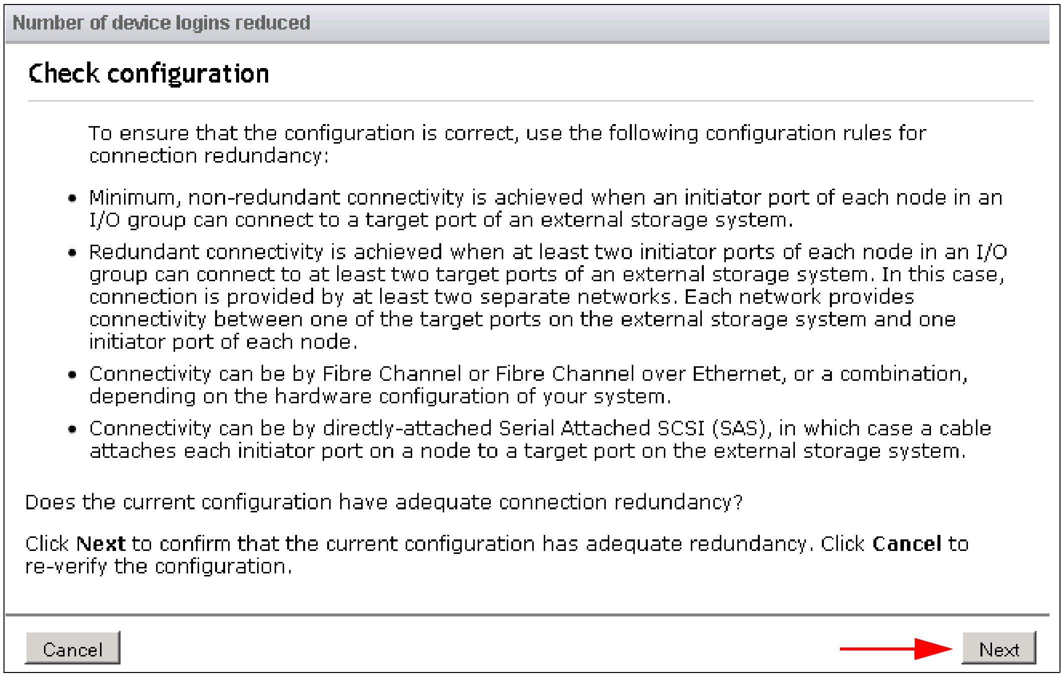
If it is caused by configuration change, click Yes to manually mark the event as fixed, otherwise, click No to scan the SAN and try to recover the paths.

○ Yes
● No

Click **Next** for more information.

*Figure 13-51 Determination of planned action*

3. If you answer **Yes**, the fix procedure finishes, assuming that all changes in the system are done on purpose and no other action is necessary. However, our example simulates a broken FC cable and we follow the complete fix procedure. Select **No** and click **Next**.
4. In the next window (Figure 13-52), IBM Storwize V7000 lists suggested actions and which components must be checked to fix and close the error. When you are sure that all possible technical requirements are met (in our case we replaced a broken FC cable), click **Next**.



Number of device logins reduced

Check configuration

To ensure that the configuration is correct, use the following configuration rules for connection redundancy:

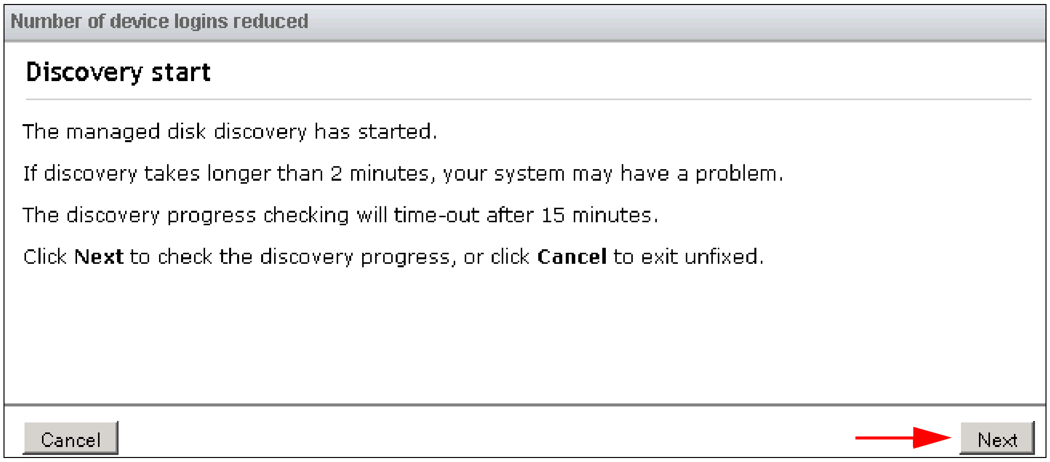
- Minimum, non-redundant connectivity is achieved when an initiator port of each node in an I/O group can connect to a target port of an external storage system.
- Redundant connectivity is achieved when at least two initiator ports of each node in an I/O group can connect to at least two target ports of an external storage system. In this case, connection is provided by at least two separate networks. Each network provides connectivity between one of the target ports on the external storage system and one initiator port of each node.
- Connectivity can be by Fibre Channel or Fibre Channel over Ethernet, or a combination, depending on the hardware configuration of your system.
- Connectivity can be by directly-attached Serial Attached SCSI (SAS), in which case a cable attaches each initiator port on a node to a target port on the external storage system.

Does the current configuration have adequate connection redundancy?

Click **Next** to confirm that the current configuration has adequate redundancy. Click **Cancel** to re-verify the configuration.

*Figure 13-52 Verification steps to eliminate single point of failure*

The discovery of managed disks starts (Figure 13-53).


Number of device logins reduced

Discovery start

The managed disk discovery has started.

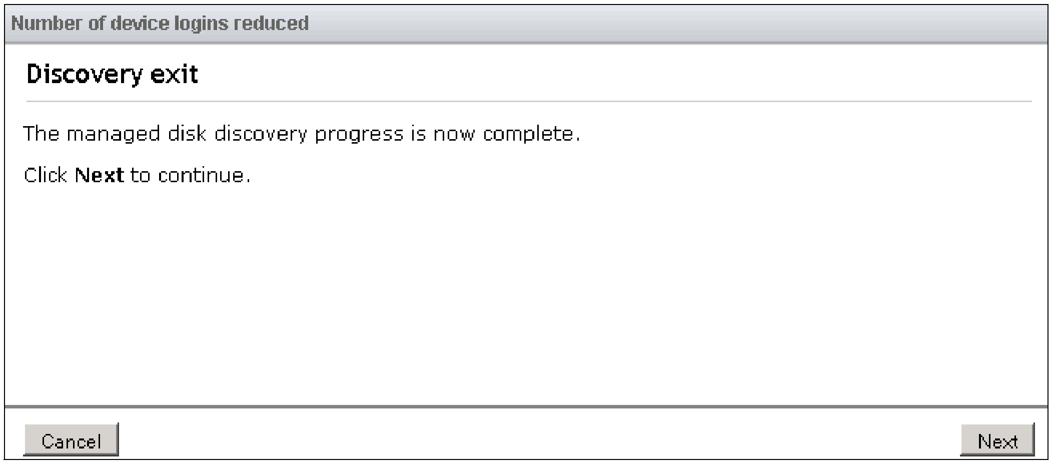
If discovery takes longer than 2 minutes, your system may have a problem.

The discovery progress checking will time-out after 15 minutes.

Click **Next** to check the discovery progress, or click **Cancel** to exit unfixed.

*Figure 13-53 Starting the discovery of managed disks*

If no other important issue exists, it finishes maximally within 2 minutes, depending on the number of enclosures and installed disk drives (Figure 13-54).


Number of device logins reduced

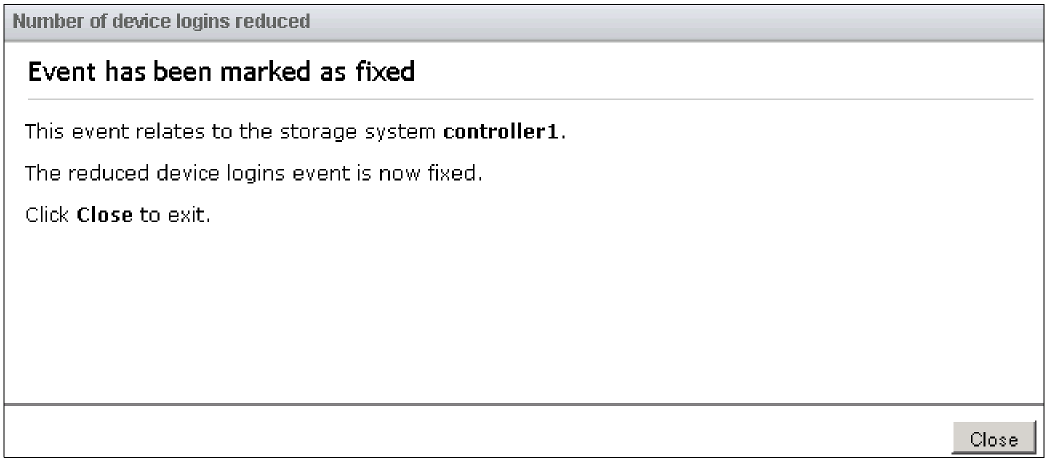
Discovery exit

The managed disk discovery progress is now complete.

Click **Next** to continue.

*Figure 13-54 Discovery complete*

5. An event has been marked as fixed, and you can safely finish the fix procedure. Click **Close** and the event is removed from the list of events (Figure 13-55).



Number of device logins reduced

Event has been marked as fixed

This event relates to the storage system **controller1**.

The reduced device logins event is now fixed.

Click **Close** to exit.

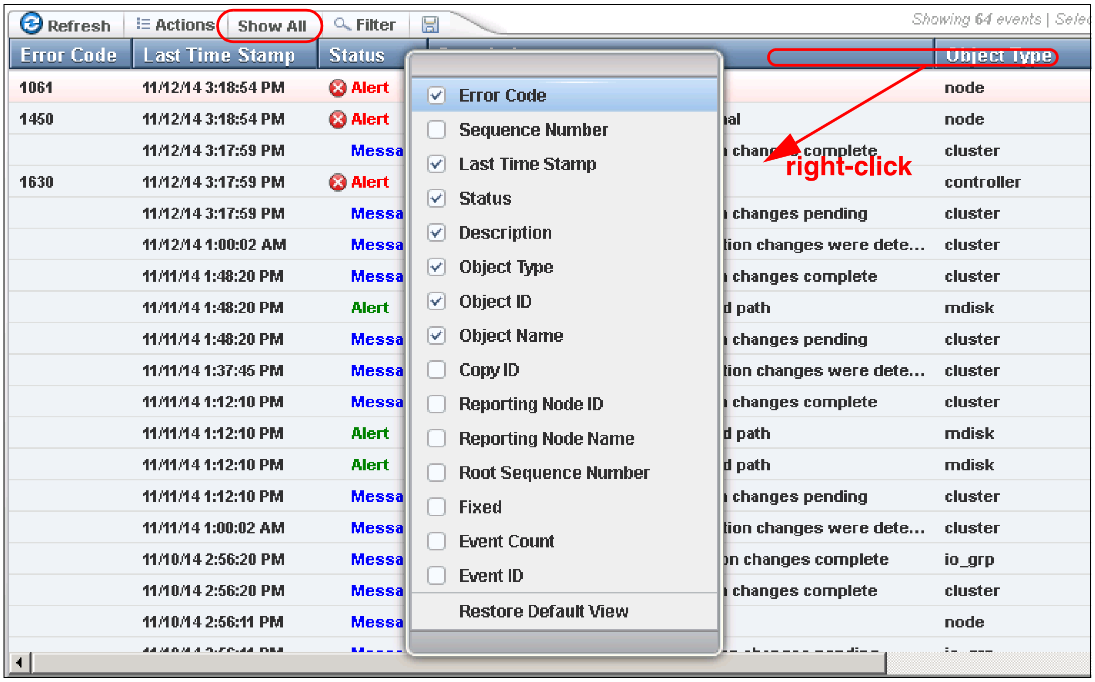
*Figure 13-55 Correctly finished fix procedure*

### Resolve alerts in a timely manner

Perform the recommended actions as quickly as possible after the problem is reported. Your system is designed to be resilient to most single hardware failures. However, if it operates for any period of time with a hardware failure, the possibility increases that a second hardware failure can result in some volume data that is unavailable. If several unfixed alerts exist, fixing any one alert might become more difficult because of the effects of the others.

## 13.6.3 Event log details

Multiple views of the events and recommended actions are available. The GUI works like a typical Windows pop-up menu, so the event log grid is manipulated through the row that contains the column headings (Figure 13-56). When you right-click a table heading, a menu for the column choices opens.



| Error Code | Last Time Stamp | Status | | Object Type |
|---|---|---|---|---|
| 1061 | 11/12/14 3:18:54 PM | Alert | | node |
| 1450 | 11/12/14 3:18:54 PM | Alert | al | node |
| | 11/12/14 3:17:59 PM | Messa | chang complete | cluster |
| 1630 | 11/12/14 3:17:59 PM | Alert | | controller |
| | 11/12/14 3:17:59 PM | Messa | changes pending | cluster |
| | 11/12/14 1:00:02 AM | Messa | tion changes were dete... | cluster |
| | 11/11/14 1:48:20 PM | Messa | changes complete | cluster |
| | 11/11/14 1:48:20 PM | Alert | d path | mdisk |
| | 11/11/14 1:48:20 PM | Messa | changes pending | cluster |
| | 11/11/14 1:37:45 PM | Messa | tion changes were dete... | cluster |
| | 11/11/14 1:12:10 PM | Messa | changes complete | cluster |
| | 11/11/14 1:12:10 PM | Alert | d path | mdisk |
| | 11/11/14 1:12:10 PM | Alert | d path | mdisk |
| | 11/11/14 1:12:10 PM | Messa | changes pending | cluster |
| | 11/11/14 1:00:02 AM | Messa | tion changes were dete... | cluster |
| | 11/10/14 2:56:20 PM | Messa | on changes complete | io_grp |
| | 11/10/14 2:56:20 PM | Messa | changes complete | cluster |
| | 11/10/14 2:56:11 PM | Messa | | node |

☑ Error Code
□ Sequence Number
☑ Last Time Stamp
☑ Status
☑ Description
☑ Object Type
☑ Object ID
☑ Object Name
□ Copy ID
□ Reporting Node ID
□ Reporting Node Name
□ Root Sequence Number
□ Fixed
□ Event Count
□ Event ID
Restore Default View

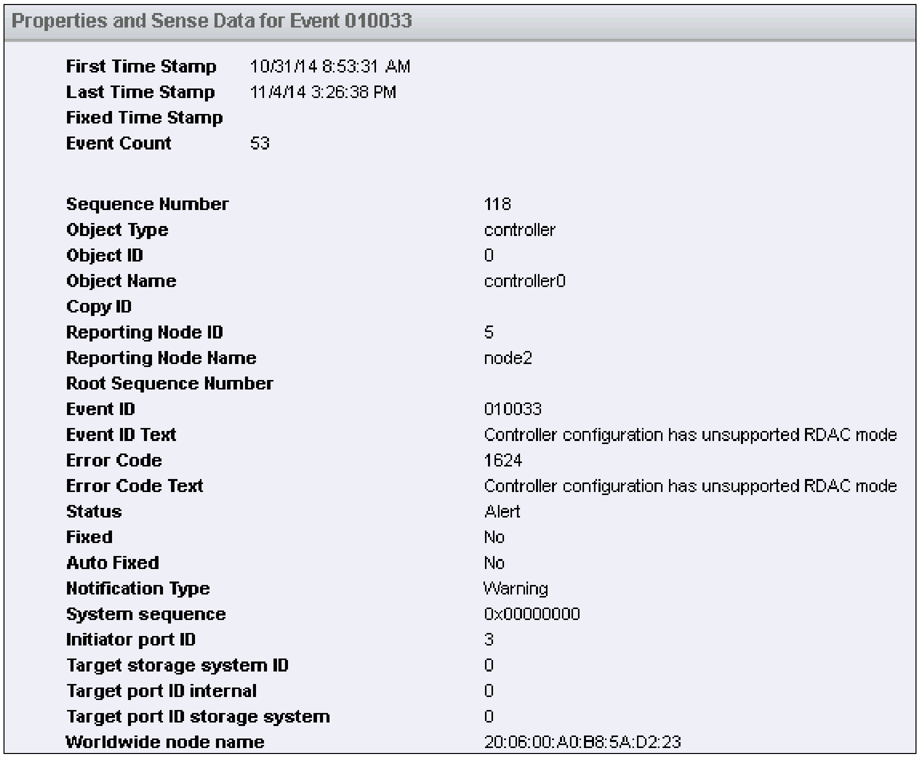
*Figure 13-56 Grid options of the event log*

Select or remove columns as needed. You can then also extend or shrink the width of the column to fit your screen resolution and size. This is the way to manipulate it for the majority of grids in the management GUI of IBM Storwize V7000, not only the events panel.

Every field of the event log is available as a column in the event log grid. Several fields are useful when you work with IBM Support. The preferred method in this case is to use the **Show All** filter, with events sorted by time stamp. All fields have the sequence number, event count, and the fixed state. Using **Restore Default View** sets the grid back to the factory defaults.

You might want to see more details about each critical event. Some details are not shown in the main grid. To access properties and sense data of a specific event, right-click the specific event (anywhere in its row) and choose **Properties** from the menu.

The properties window opens (Figure 13-57) with all the relevant sense data, such as first and last time of an event occurrence, worldwide port name (WWPN), and worldwide node name (WWNN), enabled or disabled automatic fix, and more.



Properties and Sense Data for Event 010033

| | |
|---|---|
| First Time Stamp | 10/31/14 8:53:31 AM |
| Last Time Stamp | 11/4/14 3:26:38 PM |
| Fixed Time Stamp | |
| Event Count | 53 |

| | |
|---|---|
| Sequence Number | 118 |
| Object Type | controller |
| Object ID | 0 |
| Object Name | controller0 |
| Copy ID | |
| Reporting Node ID | 5 |
| Reporting Node Name | node2 |
| Root Sequence Number | |
| Event ID | 010033 |
| Event ID Text | Controller configuration has unsupported RDAC mode |
| Error Code | 1624 |
| Error Code Text | Controller configuration has unsupported RDAC mode |
| Status | Alert |
| Fixed | No |
| Auto Fixed | No |
| Notification Type | Warning |
| System sequence | 0x00000000 |
| Initiator port ID | 3 |
| Target storage system ID | 0 |
| Target port ID internal | 0 |
| Target port ID storage system | 0 |
| Worldwide node name | 20:06:00:A0:B8:5A:D2:23 |

*Figure 13-57 Event sense data and properties*

For more details about troubleshooting options, see IBM Storwize V7000 *Troubleshooting, Recovery, and Maintenance Guide*, GC27-2291, which is available at the following location after you choose your version:

http://www.ibm.com/support/knowledgecenter/ST3FR7

## 13.7 Monitoring

An important step is to correct any issues that are reported by your IBM Storwize V7000 system as soon as possible. Configure your system to send automatic notifications when a new event is reported. To avoid monitoring for new events that use the management GUI, select the type of event for which you want to be notified. For example, restrict notifications to just events that require immediate action. Several event notification mechanisms exist:

**Email** An event notification can be sent to one or more email addresses. This mechanism notifies individuals of problems. Individuals can receive notifications wherever they have email access, including mobile devices.

**SNMP** A Simple Network Management Protocol (SNMP) traps report can be sent to a data center management system, such as IBM Systems Director, that consolidates SNMP reports from multiple systems. With this mechanism, you can monitor your data center from a single workstation.

**Syslog** A syslog report can be sent to a data center management system that consolidates syslog reports from multiple systems. With this option, you can monitor your data center from a single location.

If your system is within warranty, or you have a hardware maintenance agreement, configure your IBM Storwize V7000 system to send email events directly to IBM if an issue that requires hardware replacement is detected. This mechanism is known as *Call Home*. When this event is received, IBM automatically opens a problem report and, if appropriate, contacts you to verify whether replacement parts are required.

> **Important:** If you set up Call Home to IBM, ensure that the contact details that you configure are correct and kept up to date, because personnel can change.

## 13.7.1 Email notifications and the Call Home function

The Call Home function of IBM Storwize V7000 uses the email notification being sent to the specific IBM support center, therefore the configuration is similar as in case of sending emails to the specific person or system owner. The following procedure summarizes how to configure email notifications and emphasizes what is specific to Call Home:

1. Prepare your contact information that you want to use for the email notification and verify the accuracy of the data. From the dynamic menu, select **Settings** → **Notifications** (Figure 13-58).

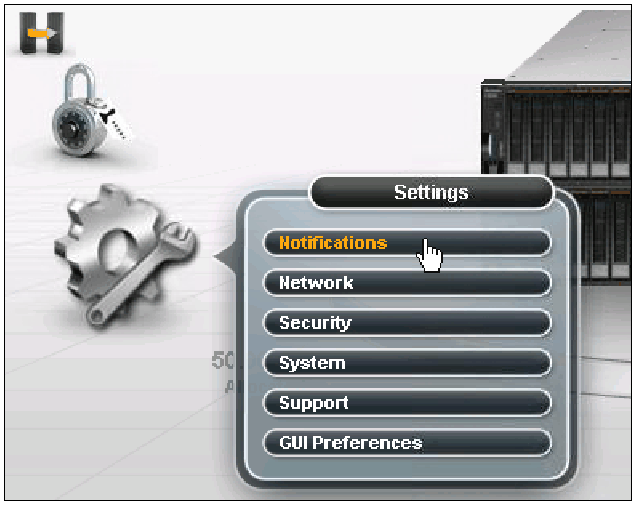
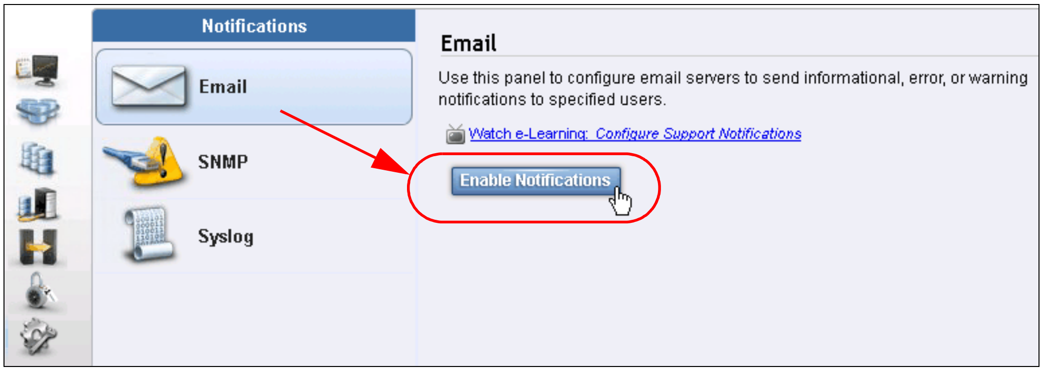


*Figure 13-58 Configuration of event notifications*

2. Select **Email** and then click **Enable Notifications** (Figure 13-59). You can also access the IBM eLearning movie for more technical details:

   `http://www.ibm.com/support/knowledgecenter/ST3FR7/welcome?cp=ST3FR7%2F0&lang=en`

   For the correct functionality of email notifications, ask your network administrator if Simple Mail Transfer Protocol (SMTP) is enabled on the management network and is not, for example, blocked by firewalls. Be sure to test the accessibility to the SMTP server using the **`telnet`** command (port 25 for a non-secured connection, port 465 for Secure Sockets Layer (SSL)-encrypted communication) using any server in the same network segment.



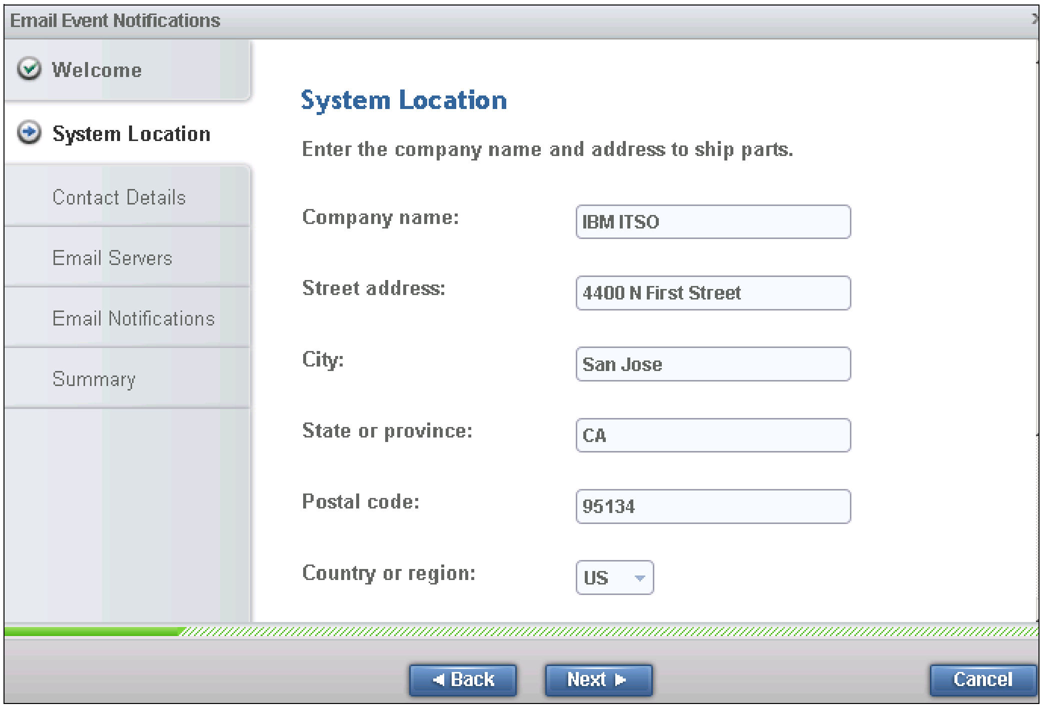
Email

Use this panel to configure email servers to send informational, error, or warning notifications to specified users.

Watch e-Learning: Configure Support Notifications

*Figure 13-59 Configuration of email notifications*

3. Provide the information about the location of the system (Figure 13-60) and contact information of IBM Storwize V7000 owner (Figure 13-61 on page 647) in order to be reachable by IBM Support. *Always* keep this information current.



System Location

Enter the company name and address to ship parts.

Company name: IBM ITSO

Street address: 4400 N First Street

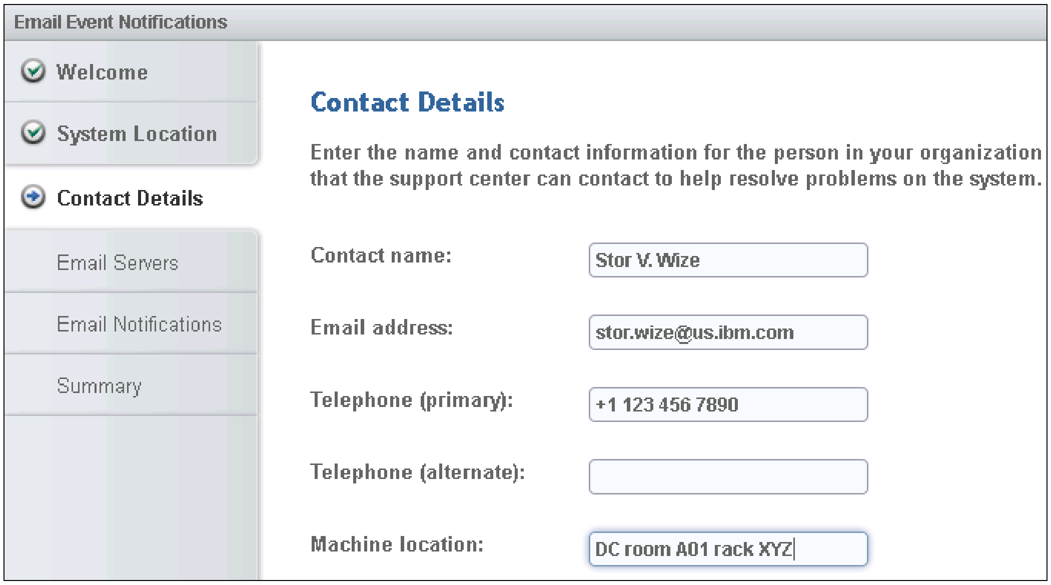
City: San Jose

State or province: CA

Postal code: 95134

Country or region: US

*Figure 13-60 Location of the device*



Email Event Notifications

Contact Details

Enter the name and contact information for the person in your organization that the support center can contact to help resolve problems on the system.

Contact name: Stor V. Wize

Email address: stor.wize@us.ibm.com

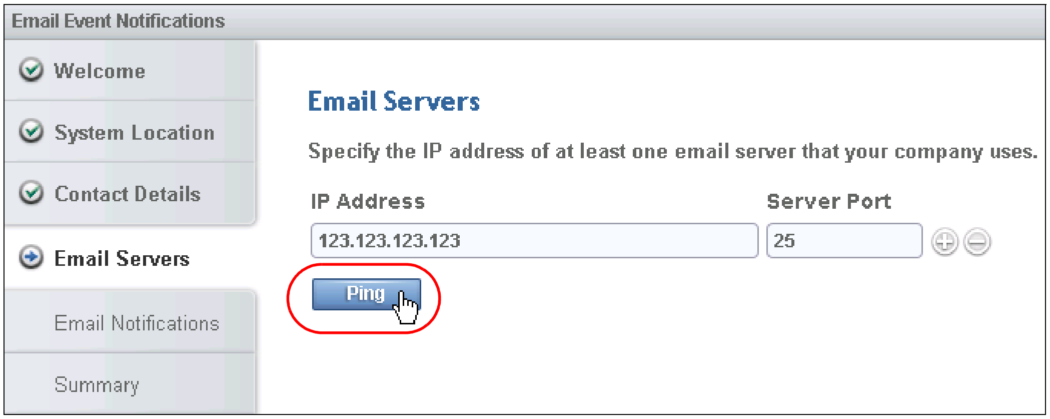
Telephone (primary): +1 123 456 7890

Telephone (alternate):

Machine location: DC room A01 rack XYZ

*Figure 13-61 Contact information*

4. Configure the SMTP server according to the instruction in Figure 13-62. When the correct SMTP server is provided, you can test the connectivity using **Ping** to its IP address.



Email Event Notifications

Email Servers

Specify the IP address of at least one email server that your company uses.

| IP Address | Server Port |
|---|---|
| 123.123.123.123 | 25 |

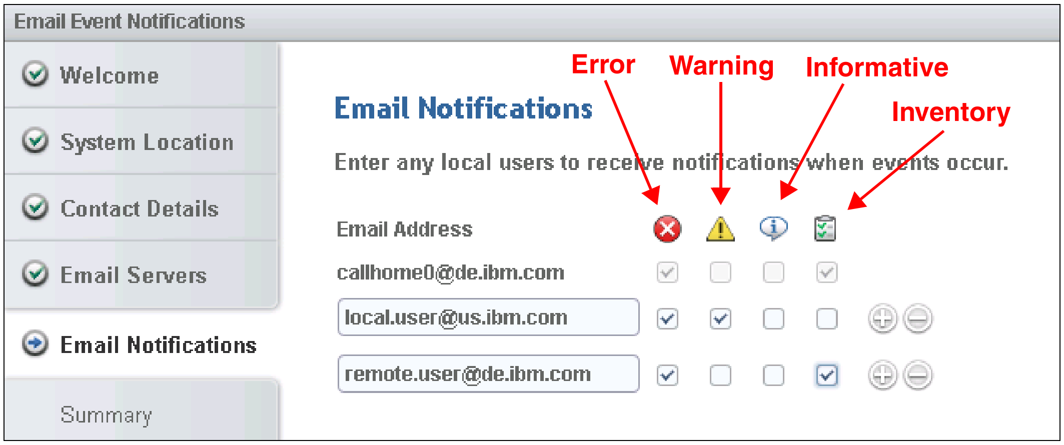
*Figure 13-62 Configure email servers and inventory reporting*

5. In the next step, verify email addresses of IBM Support (`callhome0@de.ibm.com`) and optionally local users who also need to receive notifications. See Figure 13-63 on page 648 for details. In this wizard, you can set only one email address. Other recipients can be added later.

   The default support email address `callhome0@de.ibm.com` is predefined by the system during initial configuration of email notifications, and at this stage cannot be altered. You can modify it in **Edit** mode after the initial configuration is saved (see step 6 on page 648).

   For additional users, you can also enable the **Inventory Reporting** function that is enabled by default for Call Home. Rather than reporting a problem, an email is sent to IBM that describes your system hardware and critical configuration information. Object names and other information, such as IP addresses, are not sent. The inventory email is sent regularly. Based on the information that is received, IBM can inform you if the hardware or software that you are using requires an update because of a known issue.

Figure 13-63 shows options for configuring Call Home.



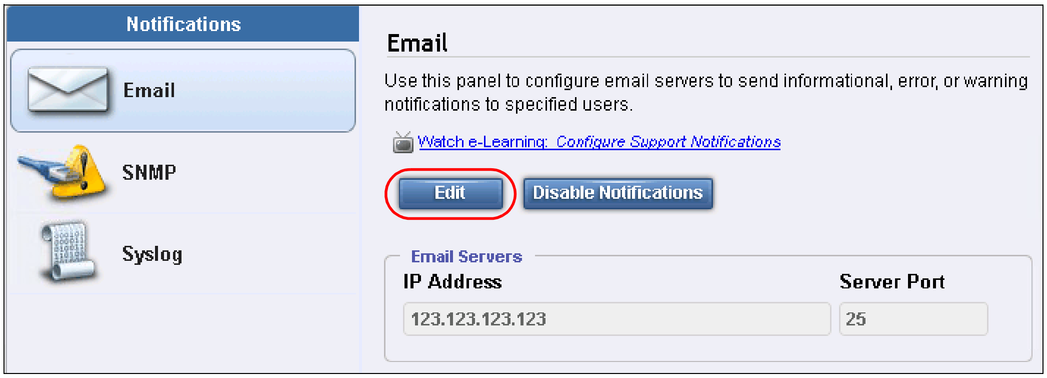
Email Notifications

Enter any local users to receive notifications when events occur.

| Email Address | Error | Warning | Informative | Inventory |
|---|---|---|---|---|
| callhome0@de.ibm.com | ☑ | □ | □ | ☑ |
| local.user@us.ibm.com | ☑ | ☑ | □ | □ |
| remote.user@de.ibm.com | ☑ | □ | □ | ☑ |

*Figure 13-63 Setting email recipients and alert types*

6. Complete the configuration wizard and test the email function. To do so, you have to enter **Edit** mode, as illustrated in Figure 13-64. In the same window, you can define the additional email recipient, either from IBM Support or local users.



Email

Use this panel to configure email servers to send informational, error, or warning notifications to specified users.

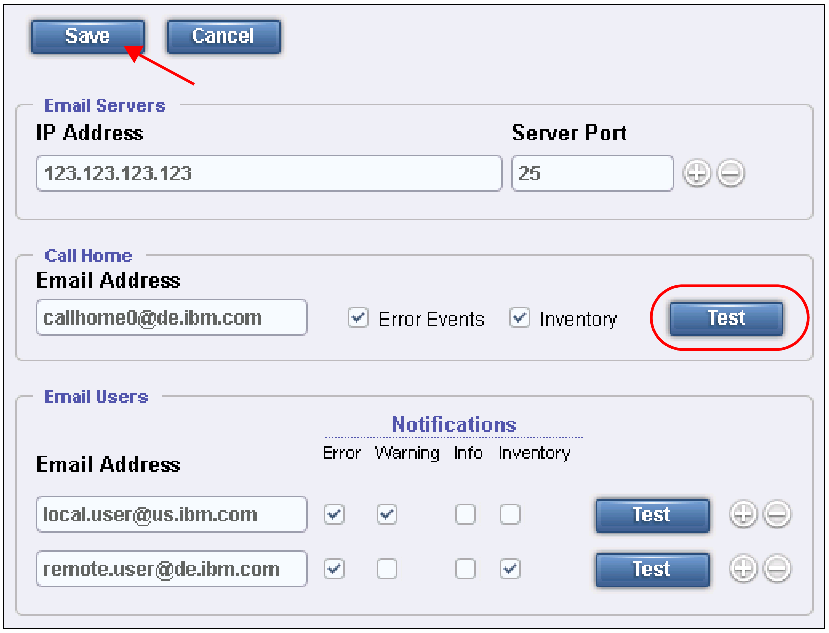
Email Servers

| IP Address | Server Port |
|---|---|
| 123.123.123.123 | 25 |

*Figure 13-64 Entering edit mode*

We strongly suggest that you keep the sending inventory option enabled to at least IBM support. However, it might be beneficial to do the same for local users. The email output can serve as a basis for the client's inventory and asset management, to keep track of all hardware devices installed in the environment.

7. In Edit mode, you are allowed to change any of the previously configured settings. When you are finished editing these parameters, have added more recipients, or just tested the connection, you can save the configuration to make the changes take effect (Figure 13-65).



IP Address: 123.123.123.123
Server Port: 25

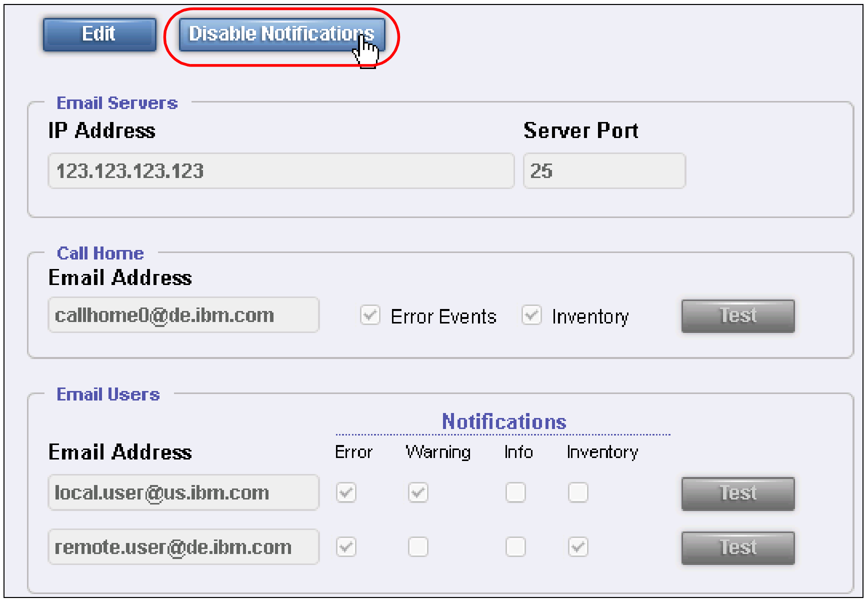
Email Address: callhome0@de.ibm.com ☑ Error Events ☑ Inventory

| Email Address | Error | Warning | Info | Inventory |
|---|---|---|---|---|
| local.user@us.ibm.com | ☑ | ☑ | □ | □ |
| remote.user@de.ibm.com | ☑ | □ | □ | ☑ |

*Figure 13-65 Saving modified configuration*

## Disabling and enabling notifications

At any time, you can temporarily or permanently disable email notifications, as shown in Figure 13-66. This is good practice when running maintenance tasks on your IBM Storwize V7000, such as code upgrade or replacement of malfunctioning parts. After the maintenance operation, remember to re-enable the email notification function. The same results can be achieved with the CLI **`svctask stopmail`** and **`svctask startmail`** commands.



IP Address: 123.123.123.123
Server Port: 25

Email Address: callhome0@de.ibm.com ☑ Error Events ☑ Inventory

| Email Address | Error | Warning | Info | Inventory |
|---|---|---|---|---|
| local.user@us.ibm.com | ☑ | ☑ | □ | □ |
| remote.user@de.ibm.com | ☑ | □ | □ | ☑ |

*Figure 13-66 Disabling or enabling email notifications*

## 13.7.2 SNMP Configuration

Simple Network Management Protocol (SNMP) is a standard protocol for managing networks and exchanging messages. The system can send SNMP messages that notify personnel about an event. You can use an SNMP manager to view the SNMP messages that are sent by the SVC.

You can configure an SNMP server to receive various informational, error, or warning notifications by entering the following information (Figure 13-67 on page 650):

- IP Address

  The address for the SNMP server.

- Server Port

  The remote port number for the SNMP server. The remote port number must be a value of 1 - 65535.

- Community

  The SNMP community is the name of the group to which devices and management stations that run SNMP belong.

- Event Notifications:

  Consider the following points about event notifications:

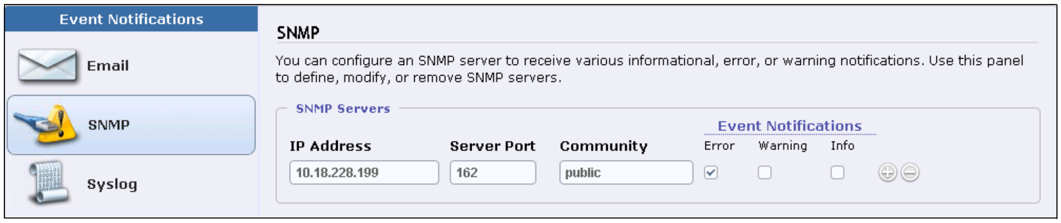
  – Select Error if you want the user to receive messages about problems, such as hardware failures, that must be resolved immediately.

    > **Important:** Browse to Recommended Actions to run the fix procedures on these notifications.

  – Select Warning if you want the user to receive messages about problems and unexpected conditions. Investigate the cause immediately to determine any corrective action.

    > **Important:** Browse to Recommended Actions to run the fix procedures on these notifications.

  – Select Info (Figure 13-67) if you want the user to receive messages about expected events. No action is required for these events.



SNMP

You can configure an SNMP server to receive various informational, error, or warning notifications. Use this panel to define, modify, or remove SNMP servers.

| IP Address | Server Port | Community | Error | Warning | Info |
|---|---|---|---|---|---|
| 10.18.228.199 | 162 | public | ☑ | □ | □ |

*Figure 13-67 SNMP configuration*

To remove an SNMP server, click the Minus sign (**-**). To add another SNMP server, click the Plus sign (**+**).

### 13.7.3 Syslog notifications

The syslog protocol is a standard protocol for forwarding log messages from a sender to a receiver on an IP network. The IP network can be IPv4 or IPv6. The system can send syslog messages that notify personnel about an event.

You can configure a syslog server to receive log messages from various systems and store them in a central repository by entering the following information (Figure 13-68 on page 651):

- IP Address

  The IP address for the syslog server.

- Facility

  The facility determines the format for the syslog messages. The facility can be used to determine the source of the message.

- Message Format

  The message format depends on the facility. The system can transmit syslog messages in the following formats:

  - The concise message format provides standard detail about the event.
  - The expanded format provides more details about the event.

- Event Notifications:

  Consider the following points about event notifications:

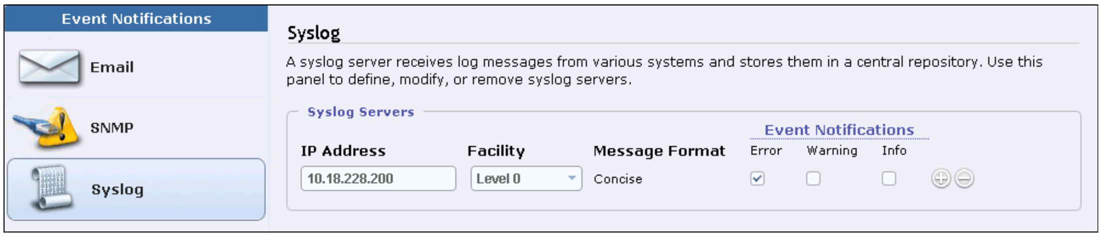
  - Select Error if you want the user to receive messages about problems, such as hardware failures, that must be resolved immediately.

    > **Important:** Browse to Recommended Actions to run the fix procedures on these notifications.

  - Select Warning if you want the user to receive messages about problems and unexpected conditions. Investigate the cause immediately to determine whether any corrective action is necessary.

    > **Important:** Browse to Recommended Actions to run the fix procedures on these notifications.

  - Select Info (Figure 13-68) if you want the user to receive messages about expected events. No action is required for these events.



Syslog

A syslog server receives log messages from various systems and stores them in a central repository. Use this panel to define, modify, or remove syslog servers.

Syslog Servers

| IP Address | Facility | Message Format | Error | Warning | Info | |
|---|---|---|---|---|---|---|
| | | | Event Notifications | | | |
| 10.18.228.200 | Level 0 | Concise | ☑ | □ | □ | ⊕⊖ |

Figure 13-68 Syslog configuration

To remove a syslog server, click the Minus sign (**-**).

To add another syslog server, click the Plus sign (**+**).

The syslog messages can be sent in concise message format or expanded message format.

Example 13-8 shows a compact format syslog message.

*Example 13-8 Compact syslog message example*

```
IBM2145 #NotificationType=Error #ErrorID=077001 #ErrorCode=1070 #Description=Node
CPU fan failed #ClusterName=SVCCluster1 #Timestamp=Wed Jul 02 08:00:00 2014 BST
#ObjectType=Node #ObjectName=Node1 #CopyID=0 #ErrorSequenceNumber=100
```

Example 13-9 shows an expanded format syslog message.

*Example 13-9 Full format syslog message example*

```
IBM2145 #NotificationType=Error #ErrorID=077001 #ErrorCode=1070 #Description=Node
CPU fan failed #ClusterName=SVCCluster1 #Timestamp=Wed Jul 02 08:00:00 2014 BST
#ObjectType=Node #ObjectName=Node1 #CopyID=0 #ErrorSequenceNumber=100 #ObjectID=2
#NodeID=2 #MachineType=21454F2#SerialNumber=1234567 #SoftwareVersion=5.1.0.0
(build 8.14.0805280000)#FRU=fan 24P1118, system board 24P1234
#AdditionalData(0->63)=0000000021000000000000000000000000000000000000000000000000
00000000000000000000000000000000000000000000000000000000000000000000#Additional
Data(64-127)=0000000000000000000000000000000000000000000000000000000000000000000
0000000000000000000000000000000000000000000000000000000000
```

## 13.8 Audit log

The audit log is useful when analyzing past configuration events, especially when trying to determine, for example, how a volume ended up being shared by two hosts, or why the volume was overwritten. The audit log is also included in the *svc_snap* support data to aid in problem determination.

An audit log tracks action commands that are issued through a Secure Shell (SSH) session or through the management GUI. It provides the following entries:

- Identity of the user who issued the action command
- Name of the actionable command
- Time stamp of when the actionable command was issued on the configuration node
- Parameters that were issued with the actionable command

The following items are not documented in the audit log:

- Commands that fail are not logged.
- A result code of 0 (success) or 1 (success in progress) is not logged.
- Result object ID of node type (for the **addnode** command) is not logged.
- Views are not logged.

Several specific service commands are not included in the audit log:

- **dumpconfig**
- **cpdumps**
- **cleardumps**
- **finderr**
- **dumperrlog**
- **dumpintervallog**
- **svcservicetak dumperrlog**
- **svcservicetask finderr**

Figure 13-69 shows the access to the audit log. Click **Audit Log** in the dynamic menu to see which configuration CLI commands have been run on IBM Storwize V7000 system.

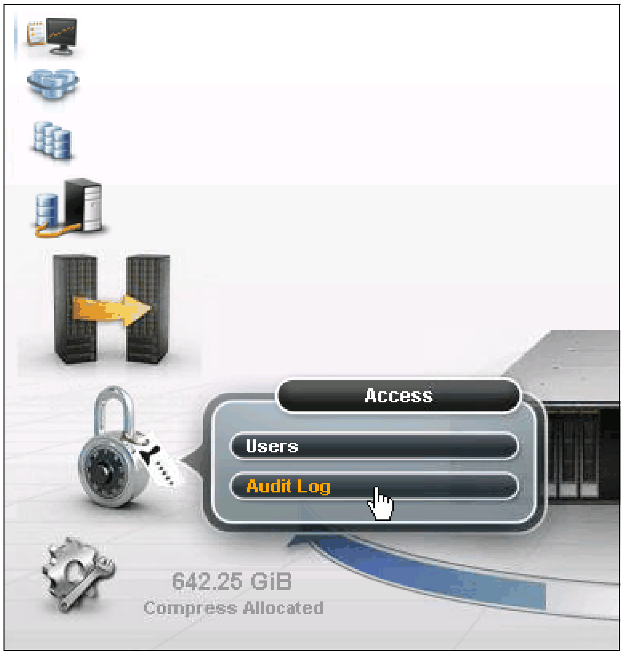
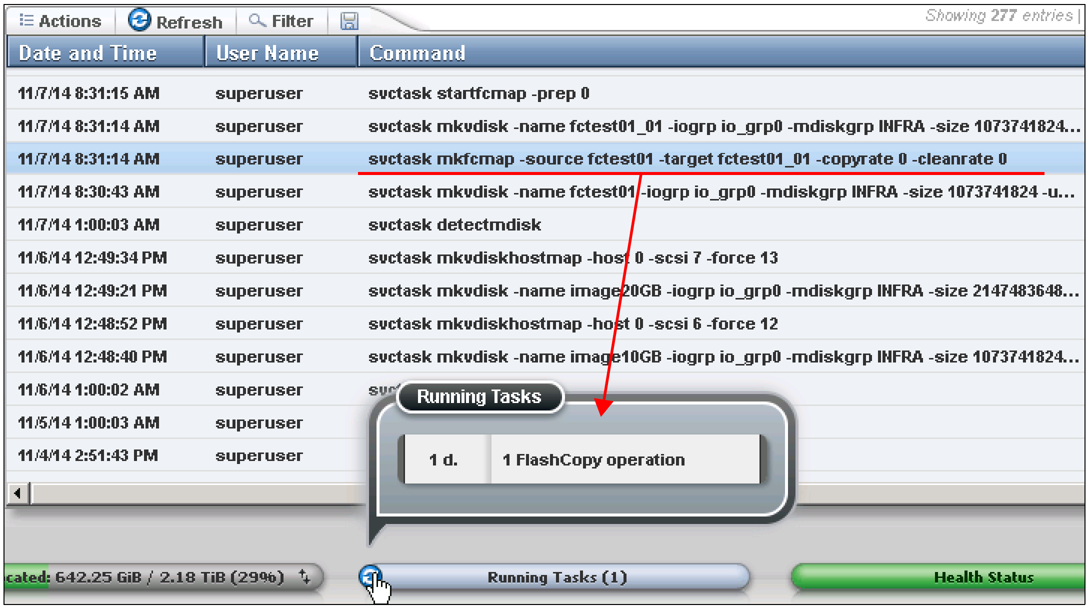


*Figure 13-69 Audit Log from Access Management window*

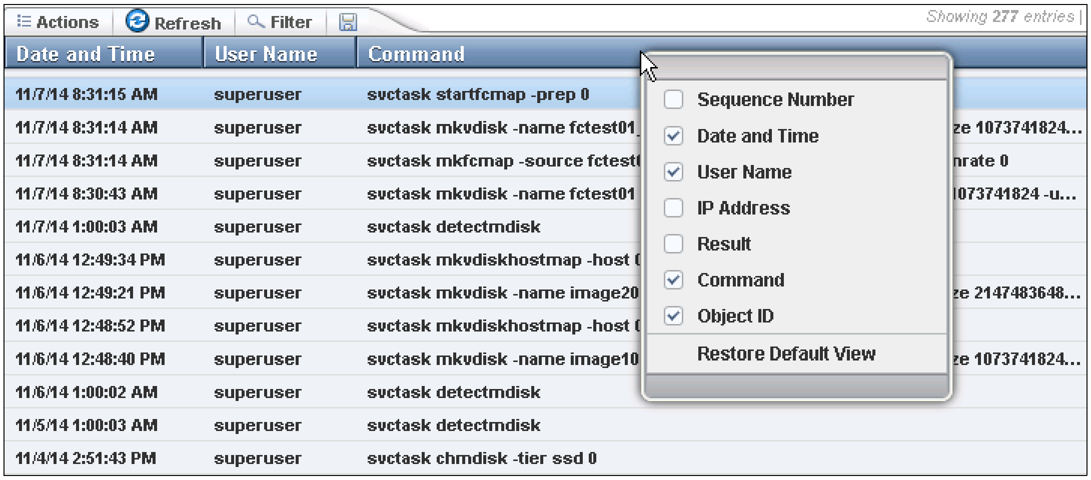
Figure 13-70 shows an example of the audit log after creating a FlashCopy volume and mapping it to hosts, with a command highlighted. The **Running Tasks** button is available at the bottom of the window in the status panel. If you click that button, the progress of the currently running tasks is displayed.



| Date and Time | User Name | Command |
|---|---|---|
| 11/7/14 8:31:15 AM | superuser | svctask startfcmap -prep 0 |
| 11/7/14 8:31:14 AM | superuser | svctask mkvdisk -name fctest01_01 -iogrp io_grp0 -mdiskgrp INFRA -size 1073741824... |
| 11/7/14 8:31:14 AM | superuser | svctask mkfcmap -source fctest01 -target fctest01_01 -copyrate 0 -cleanrate 0 |
| 11/7/14 8:30:43 AM | superuser | svctask mkvdisk -name fctest01 -iogrp io_grp0 -mdiskgrp INFRA -size 1073741824 -u... |
| 11/7/14 1:00:03 AM | superuser | svctask detectmdisk |
| 11/6/14 12:49:34 PM | superuser | svctask mkvdiskhostmap -host 0 -scsi 7 -force 13 |
| 11/6/14 12:49:21 PM | superuser | svctask mkvdisk -name image20GB -iogrp io_grp0 -mdiskgrp INFRA -size 2147483648... |
| 11/6/14 12:48:52 PM | superuser | svctask mkvdiskhostmap -host 0 -scsi 6 -force 12 |
| 11/6/14 12:48:40 PM | superuser | svctask mkvdisk -name image10GB -iogrp io_grp0 -mdiskgrp INFRA -size 1073741824... |
| 11/6/14 1:00:02 AM | superuser | svc |
| 11/5/14 1:00:03 AM | superuser | |
| 11/4/14 2:51:43 PM | superuser | |

*Figure 13-70 Audit log*

Changing the view of the Audit Log grid is also possible by right-clicking column headings (Figure 13-71). The grid layout and sorting is completely under the user's control, so you can view everything in the audit log, sort different columns, or reset the default grid preferences.



| Date and Time | User Name | Command | |
|---|---|---|---|
| 11/7/14 8:31:15 AM | superuser | svctask startfcmap -prep 0 | |
| 11/7/14 8:31:14 AM | superuser | svctask mkvdisk -name fctest01_ | ze 1073741824... |
| 11/7/14 8:31:14 AM | superuser | svctask mkfcmap -source fctest( | nrate 0 |
| 11/7/14 8:30:43 AM | superuser | svctask mkvdisk -name fctest01 | 073741824 -u... |
| 11/7/14 1:00:03 AM | superuser | svctask detectmdisk | |
| 11/6/14 12:49:34 PM | superuser | svctask mkvdiskhostmap -host ( | |
| 11/6/14 12:49:21 PM | superuser | svctask mkvdisk -name image20 | ze 2147483648... |
| 11/6/14 12:48:52 PM | superuser | svctask mkvdiskhostmap -host ( | |
| 11/6/14 12:48:40 PM | superuser | svctask mkvdisk -name image10 | ze 1073741824... |
| 11/6/14 1:00:02 AM | superuser | svctask detectmdisk | |
| 11/5/14 1:00:03 AM | superuser | svctask detectmdisk | |
| 11/4/14 2:51:43 PM | superuser | svctask chmdisk -tier ssd 0 | |

- □ Sequence Number
- ☑ Date and Time
- ☑ User Name
- □ IP Address
- □ Result
- ☑ Command
- ☑ Object ID
- Restore Default View

*Figure 13-71 Right-click audit log column headings*

# 13.9 Collecting support information using the GUI and the CLI

Occasionally, if you have a problem and call the IBM Support Center, they might ask you to provide support data. You can find this data under the Support tab of the Troubleshooting navigation window.

## 13.9.1 Collecting information using the GUI

To collect information using the GUI, complete the following steps:

1. Click **Settings** and then the **Support** tab to begin the procedure of collecting support data (Figure 13-72 on page 655).
2. Assuming that the node restarts, use the menu shown in Figure 13-73 on page 655 to collect the default logs plus all of the existing statesaves to capture the maximum data for support.

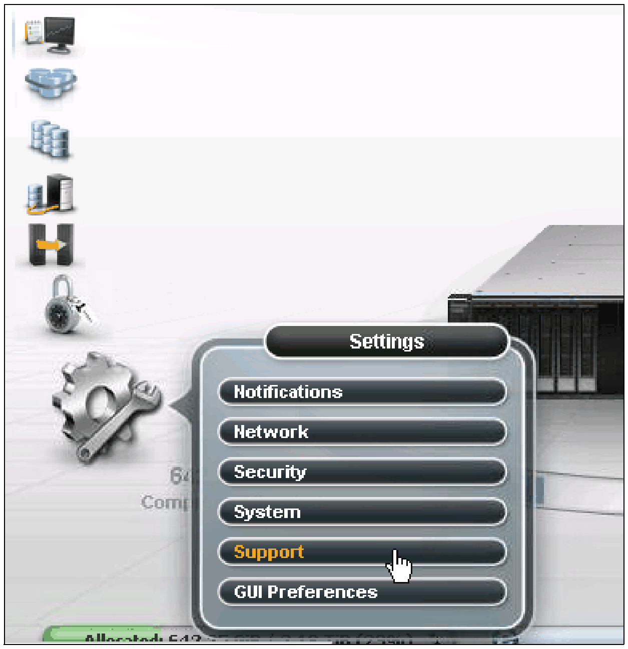
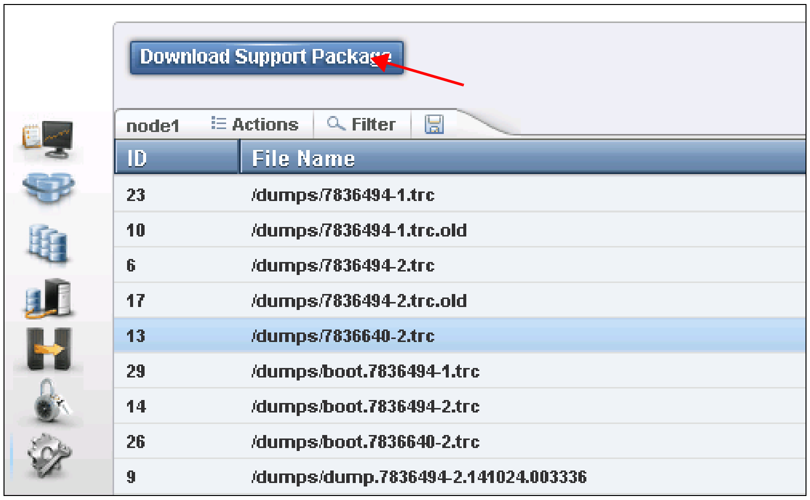


*Figure 13-72 Support option Data Collection*

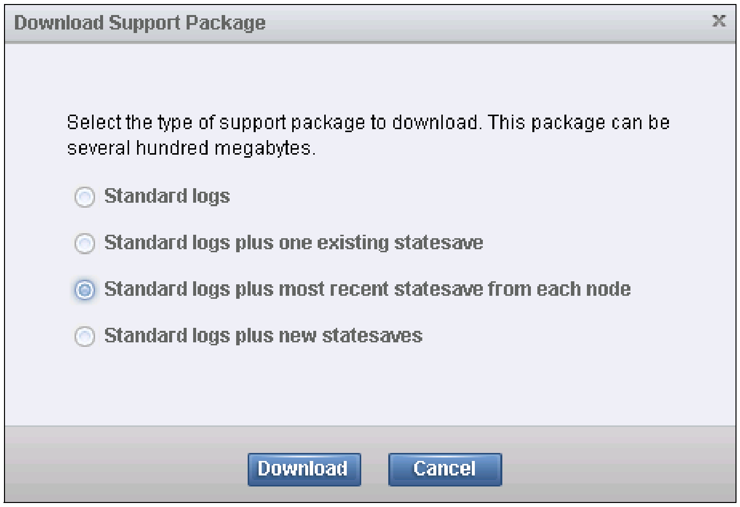
3. Click **Download Support Package** (Figure 13-73). To list all individual log files, click the **Show All Individual Log Files** menu.



| ID | File Name |
|---|---|
| 23 | /dumps/7836494-1.trc |
| 10 | /dumps/7836494-1.trc.old |
| 6 | /dumps/7836494-2.trc |
| 17 | /dumps/7836494-2.trc.old |
| 13 | /dumps/7836640-2.trc |
| 29 | /dumps/boot.7836494-1.trc |
| 14 | /dumps/boot.7836494-2.trc |
| 26 | /dumps/boot.7836640-2.trc |
| 9 | /dumps/dump.7836494-2.141024.003336 |

*Figure 13-73 Download Support Package window*

The window for collecting various versions of `svc_snap` opens (Figure 13-74). The version you download depends on the event that is investigated. For example, if you notice that a node was restarted in the event log, capture the snap with the latest existing statesave.



Download Support Package

Select the type of support package to download. This package can be several hundred megabytes.

○ Standard logs

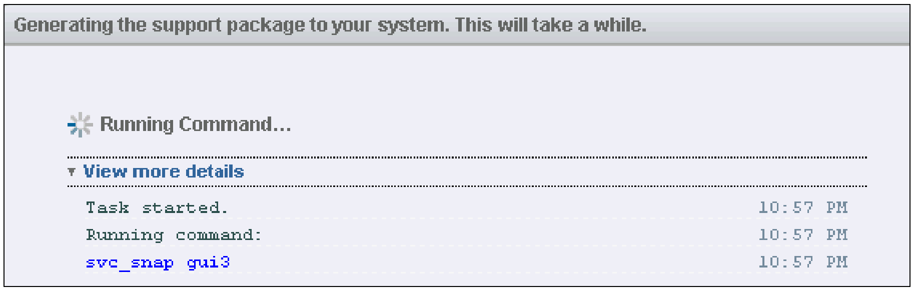
○ Standard logs plus one existing statesave

● Standard logs plus most recent statesave from each node

○ Standard logs plus new statesaves

*Figure 13-74 Download support package choices*

The procedure to create the snap on an IBM Storwize V7000 system, including the latest statesave from each node canister, starts. This might take a few minutes (Figure 13-75).



Generating the support package to your system. This will take a while.

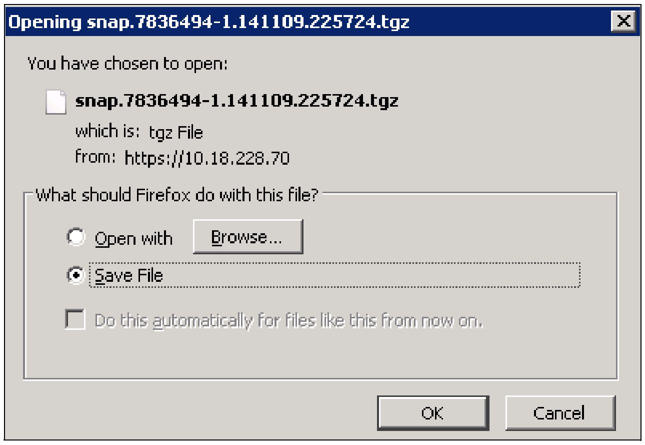
Running Command...

View more details

```
Task started.                                   10:57 PM
Running command:                                10:57 PM
svc_snap gui3                                   10:57 PM
```

*Figure 13-75 Task detail window*

4. A window opens that gives you the choice to save the file on your local Windows system (Figure 13-76).



Opening snap.7836494-1.141109.225724.tgz

You have chosen to open:

snap.7836494-1.141109.225724.tgz

which is: tgz File

from: https://10.18.228.70

What should Firefox do with this file?

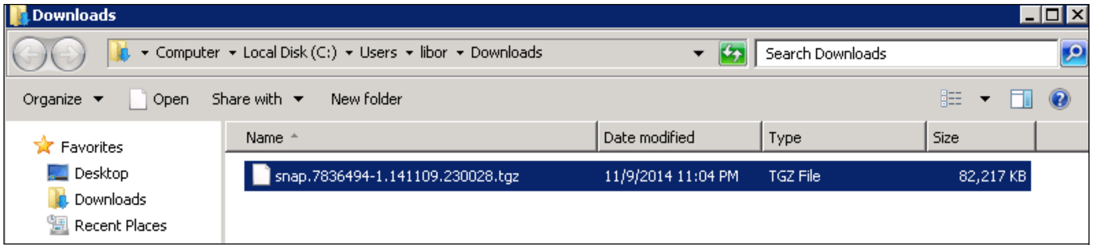
○ Open with

● Save File

□ Do this automatically for files like this from now on.

*Figure 13-76 Save file to the local Windows system*

5. Save the resulting snap file in a directory (Figure 13-77).



| Name | Date modified | Type | Size |
|---|---|---|---|
| snap.7836494-1.141109.230028.tgz | 11/9/2014 11:04 PM | TGZ File | 82,217 KB |

*Figure 13-77 Save the resulting snap file to a local directory*

Before you open a call with IBM Support, be prepared to upload the resulting snap file to the IBM Support portal at the following address:

`http://www.ecurep.ibm.com/app/upload`

You are ready to call the IBM Support Line or use the IBM Support Portal to open a call. If you use the latter option, go to the following address:

`http://www.ibm.com/support/entry/portal/Open_service_request?brandind=Hardware`

## 13.9.2 Collecting logs using the CLI

The CLI can be used to collect logs by perform the following steps:

1. Log in to the CLI and to run the **`svc_snap -c`** command, as shown in Example 13-10.

   *Example 13-10 The svc_snap command*

   ```
   login as: superuser
   superuser@10.18.228.185's password:
   IBM_Storwize:V7000 Gen 1 EXTSTG2:superuser>svc_snap -c
   Collecting data
   ```

2. To generate snap and livedump, run the following commands in this order:

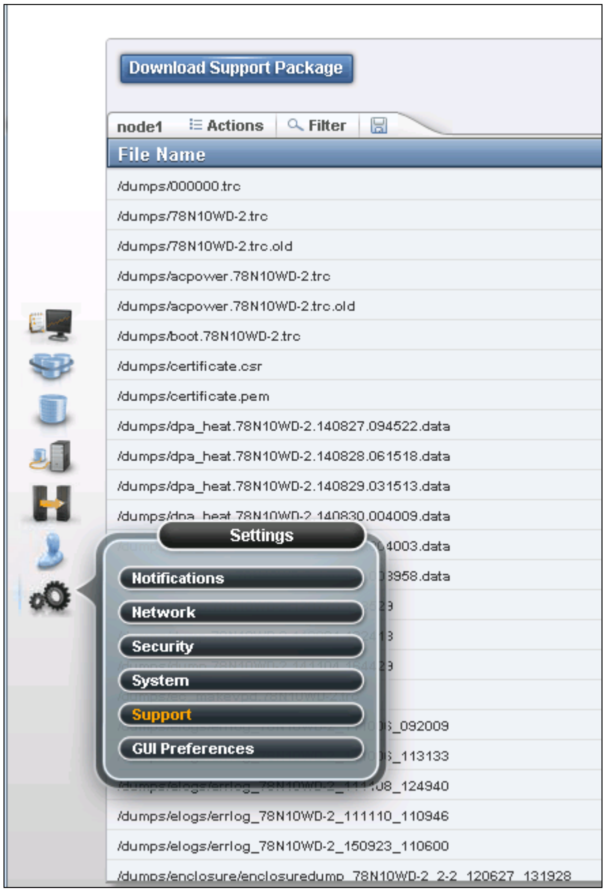
   a. The **`svc_livedump`** command
   b. The **`svc_snap gui3`** command

3. The commands, when completed, create the following file:

   `/dumps/snap.<panel_id>.YYMMDD.hhmmss.tgz`

   It takes a few minutes for the snap file to complete, and longer if pulling statesaves.

4. The generated file can be retrieved from the GUI under **Support** → **show full log listing**, as shown in Figure 13-78.



| File Name |
|---|
| /dumps/000000.trc |
| /dumps/78N10WD-2.trc |
| /dumps/78N10WD-2.trc.old |
| /dumps/acpower.78N10WD-2.trc |
| /dumps/acpower.78N10WD-2.trc.old |
| /dumps/boot.78N10WD-2.trc |
| /dumps/certificate.csr |
| /dumps/certificate.pem |
| /dumps/dpa_heat.78N10WD-2.140827.094522.data |
| /dumps/dpa_heat.78N10WD-2.140828.061518.data |
| /dumps/dpa_heat.78N10WD-2.140829.031513.data |
| /dumps/dpa_heat.78N10WD-2.140830.004009.data |
| 4003.data |
| 3958.data |
| 29 |
| 13 |
| 29 |
| |
| _092009 |
| _113133 |
| 08_124940 |
| /dumps/elogs/errlog_78N10WD-2_111110_110946 |
| /dumps/elogs/errlog_78N10WD-2_150923_110600 |
| /dumps/enclosure/enclosuredump_78N10WD-2_2-2_120627_131928 |

*Figure 13-78 Downloaded log listing*

5. Scroll and locate the file (`dumps/snap.78N10WD_2.151028.113043.tgz`), right-click and download the file.
6. Save the file to your desktop or a temporary folder.

## 13.9.3 Uploading files to the Support Center

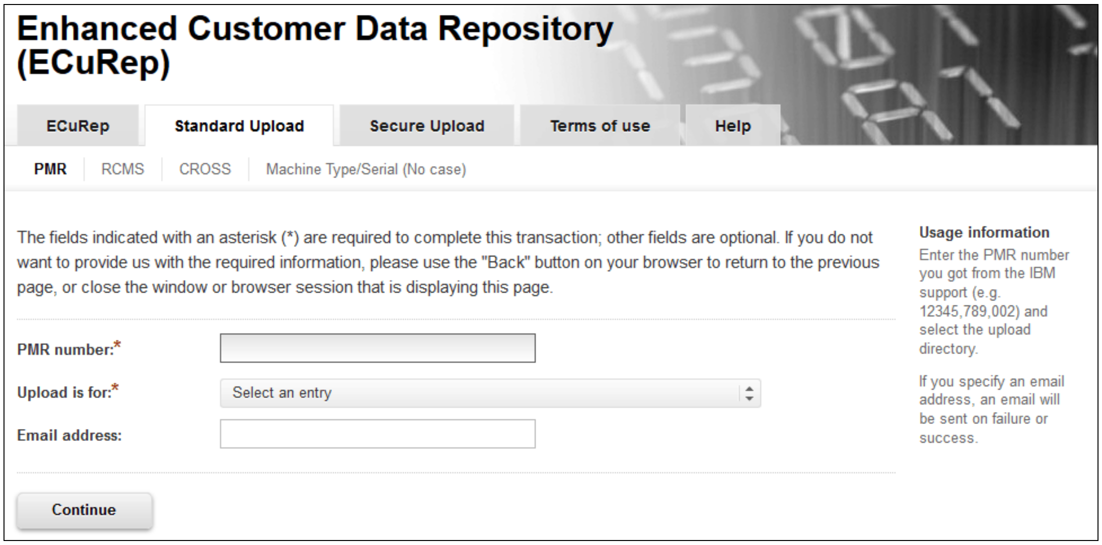
Client information can be uploaded to IBM Support for analysis via the Enhanced Customer Data Repository (ECuRep). Any uploads should be associated with a specific problem management report (PMR). The PMR is also known as a *service request*. In fact, this is a mandatory requirement when uploading.

To upload information, use the following procedure:

1. Using a browser navigate to ECuRep:

   `http://www.ecurep.ibm.com/app/upload`

   This link takes you to the Standard Upload page (Figure 13-79 on page 659).



Enhanced Customer Data Repository (ECuRep)

The fields indicated with an asterisk (*) are required to complete this transaction; other fields are optional. If you do not want to provide us with the required information, please use the "Back" button on your browser to return to the previous page, or close the window or browser session that is displaying this page.

PMR number:*

Upload is for:* Select an entry

Email address:

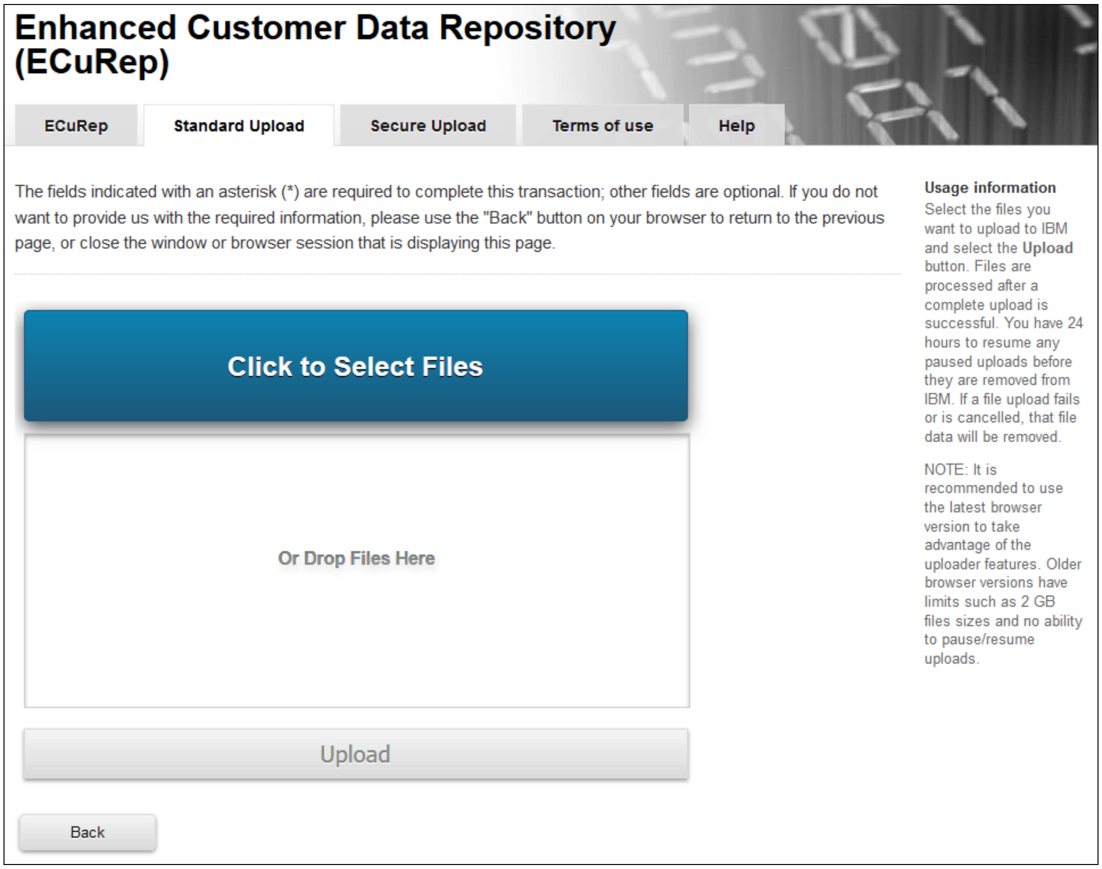
Usage information

Enter the PMR number you got from the IBM support (e.g. 12345,789,002) and select the upload directory.

If you specify an email address, an email will be sent on failure or success.

*Figure 13-79 ECuREP*

2. Fill in the required fields:
   - PMR number (mandatory). This should be in the format of `XXXXX,YYY,ZZZ`, for example, 43885,487,000 using a comma (,) as a separator.
   - Upload is for (mandatory). Select **Hardware** from the drop-down menu.
   - Email address (not mandatory). Client is advised to input their contact email address in this field.
3. When completed, click **Continue**. This opens the input window (Figure 13-80).



The fields indicated with an asterisk (*) are required to complete this transaction; other fields are optional. If you do not want to provide us with the required information, please use the "Back" button on your browser to return to the previous page, or close the window or browser session that is displaying this page.

**Usage information**
Select the files you want to upload to IBM and select the **Upload** button. Files are processed after a complete upload is successful. You have 24 hours to resume any paused uploads before they are removed from IBM. If a file upload fails or is cancelled, that file data will be removed.

NOTE: It is recommended to use the latest browser version to take advantage of the uploader features. Older browser versions have limits such as 2 GB files sizes and no ability to pause/resume uploads.

*Figure 13-80 Upload*

4. After the files have been selected, click **Upload** to continue, and then follow the directions.

## 13.10 Service Assistant Tool

The SA (Service Assistant) is a web-based GUI that is used to service individual node canisters, primarily when a node has a fault and is in a service state. A node cannot be active as part of a clustered system while it is in a service state.

Typically, the IBM Storwize V7000 is initially configured with the following IP addresses:

- One service IP address for each of control canisters.
- One management IP address, which is set when the cluster is started.

The SA is available even when the management GUI is not accessible. The following information and tasks can be accomplished with the Service Assistance Tool.

- Status information about the connections and the node canister
- Basic configuration information, such as configuring IP addresses
- Service tasks, such as restarting the Common Information Model (CIM) object manager (CIMOM) and updating the worldwide node name (WWNN)
- Details about node error codes
- Details about the hardware such as IP address, MAC (Machine Access Control) addresses.

The SA GUI is available using a service assistant IP address that is configured on each node and can also be accessed through the cluster IP addresses by appending service to the cluster management Uniform Resource Locator (URL).

If the system is down, the only other method of communicating with the node canisters is through the SA IP address directly. Each node can have a single IP address on Ethernet port 1. It is recommended that these IP addresses are configured on all nodes of the cluster.

To open the SA GUI, enter one of the following URLs into any web browser:

- http(s)://<cluster IP address of your cluster>/service
- http(s)://<service IP address of a node>/service

Complete the following steps to access the SA:

1. If you are accessing SA by using <cluster IP address>/service, the configuration node canister SA GUI login window opens. Enter the Superuser Password, as shown in Figure 13-81.

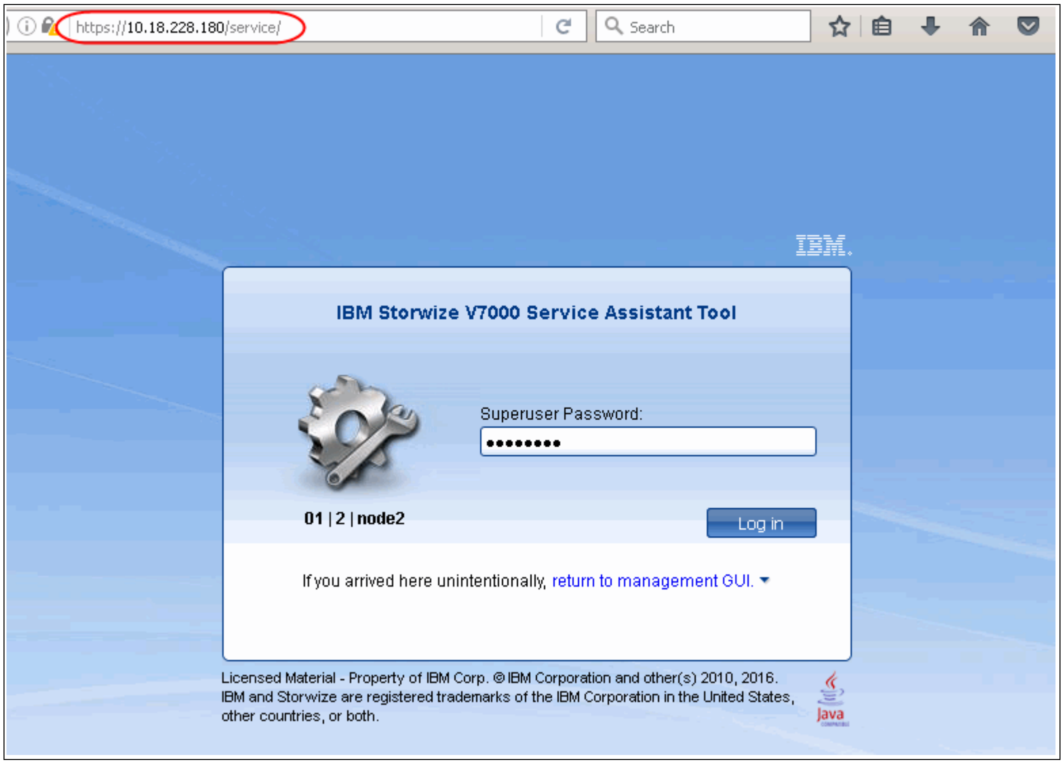
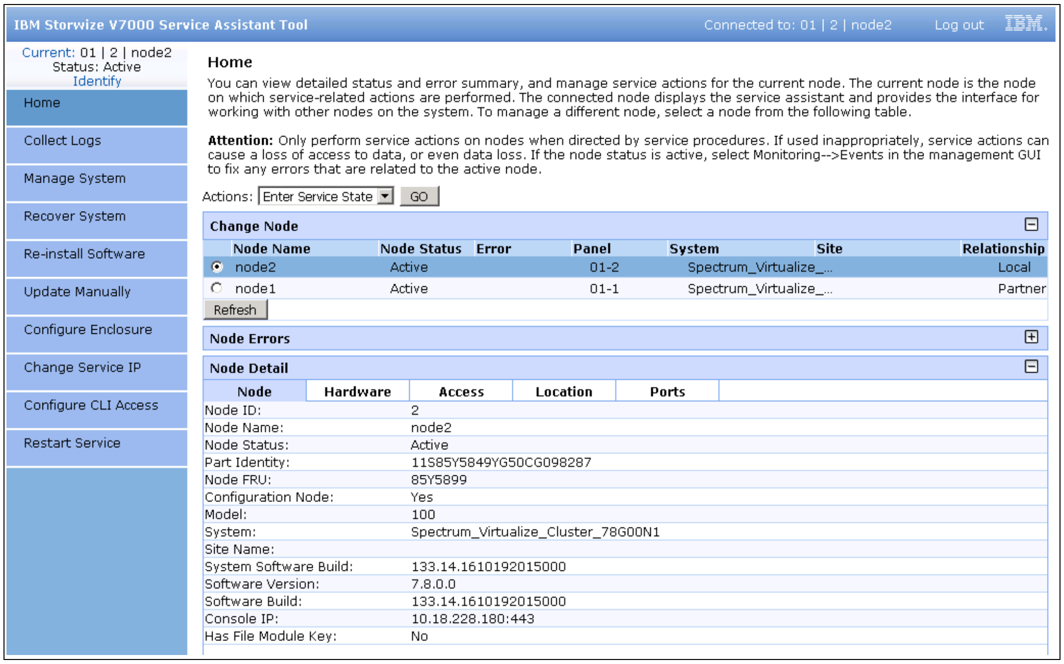

https://10.18.228.180/service/
Search
IBM.
IBM Storwize V7000 Service Assistant Tool
Superuser Password:
••••••••
01 | 2 | node2
Log in
If you arrived here unintentionally, return to management GUI.
Licensed Material - Property of IBM Corp. © IBM Corporation and other(s) 2010, 2016. IBM and Storwize are registered trademarks of the IBM Corporation in the United States, other countries, or both.
Java

*Figure 13-81 Service Assistant Tool Login GUI*

2. After you are logged in, you see the Service Assistant Home window, as shown in Figure 13-82 on page 662. The SA interfaces can view the status of, and run service actions on, other nodes, in addition to the node where the user is connected.



Home

You can view detailed status and error summary, and manage service actions for the current node. The current node is the node on which service-related actions are performed. The connected node displays the service assistant and provides the interface for working with other nodes on the system. To manage a different node, select a node from the following table.

**Attention:** Only perform service actions on nodes when directed by service procedures. If used inappropriately, service actions can cause a loss of access to data, or even data loss. If the node status is active, select Monitoring-->Events in the management GUI to fix any errors that are related to the active node.

Change Node

| | Node Name | Node Status | Error | Panel | System | Site | Relationship |
|---|---|---|---|---|---|---|---|
| ● | node2 | Active | | 01-2 | Spectrum_Virtualize_... | | Local |
| ○ | node1 | Active | | 01-1 | Spectrum_Virtualize_... | | Partner |

| | |
|---|---|
| Node ID: | 2 |
| Node Name: | node2 |
| Node Status: | Active |
| Part Identity: | 11S85Y5849YG50CG098287 |
| Node FRU: | 85Y5899 |
| Configuration Node: | Yes |
| Model: | 100 |
| System: | Spectrum_Virtualize_Cluster_78G00N1 |
| Site Name: | |
| System Software Build: | 133.14.1610192015000 |
| Software Version: | 7.8.0.0 |
| Software Build: | 133.14.1610192015000 |
| Console IP: | 10.18.228.180:443 |
| Has File Module Key: | No |

*Figure 13-82 Service Assistant Tool GUI*

3. The current node canister is displayed in the upper left corner of the GUI. As shown in this is node 2. To change the canister, select the relevant node in the Change Node section of the window. You see the details in the upper left change to reflect the new canister.

> **Note:** The SA GUI provides access to service procedures and shows the status of the node canisters. It is advised that these procedures should only be carried out if directed to do so by IBM Support.

For more information about how to use the SA tool, see the following website:

https://ibm.biz/BdsEGf

A

# Performance data and statistics gathering

In this appendix, we provide a brief overview of the performance analysis capabilities of the IBM Storwize V7000 and IBM Spectrum Virtualize V7.8. We also describe a method that you can use to collect and process IBM Spectrum Virtualize performance statistics.

It is beyond the intended scope of this book to provide an in-depth understanding of performance statistics, or explain how to interpret them. For more information about the performance of the Storwize V7000, see *IBM System Storage SAN Volume Controller and Storwize V7000 Best Practices and Performance Guidelines*, SG24-7521, which is available at the following website:

http://www.redbooks.ibm.com/abstracts/sg247521.html

This appendix describes the following topics:

- Storwize V7000 performance overview
- Performance monitoring

# Storwize V7000 performance overview

Although storage virtualization with the IBM Spectrum Virtualize provides many administrative benefits, it can also provide a substantial increase in performance for various workloads. The caching capability of the IBM Spectrum Virtualize and its ability to stripe volumes across multiple disk arrays can provide a significant performance improvement over what can otherwise be achieved when midrange disk subsystems are used.

To ensure that the performance levels of your system that you want are maintained, monitor performance periodically to provide visibility to potential problems that exist or are developing so that they can be addressed in a timely manner.

## Performance considerations

When you are designing the IBM Spectrum Virtualize infrastructure or maintaining an existing infrastructure, you must consider many factors in terms of their potential effect on performance. These factors include, but are not limited to dissimilar workloads competing for the same resources, overloaded resources, insufficient available resources, poor performing resources, and similar performance constraints.

Remember the following high-level rules when you are designing your storage area network (SAN) and IBM Spectrum Virtualize layout:

- Host-to-Storwize V7000 inter-switch link (ISL) oversubscription

  This area is the most significant input/output (I/O) load across ISLs. The recommendation is to maintain a maximum of 7-to-1 oversubscription. A higher ratio is possible, but it tends to lead to I/O bottlenecks. This suggestion also assumes a core-edge design, where the hosts are on the edges and the Storwize V7000 is the core.

- Storage-to-Storwize V7000 ISL oversubscription

  This area is the second most significant I/O load across ISLs. The maximum oversubscription is 7-to-1. A higher ratio is not supported. Again, this suggestion assumes a multiple-switch SAN fabric design.

- ISL trunking/port channeling

  For the best performance and availability, we strongly advise you to use ISL trunking or port channeling. Independent ISL links can easily become overloaded and turn into performance bottlenecks. Bonded or trunked ISLs automatically share load and provide better redundancy in a failure.

- Number of paths per host multipath device

  The maximum supported number of paths per multipath device that is visible on the host is eight. Although the IBM Subsystem Device Driver Path Control Module (SDDPCM), related products, and most vendor multipathing software can support more paths, the Storwize V7000 expects a maximum of eight paths. In general, you see only an effect on performance from more paths than eight. Although the IBM Spectrum Virtualize can work with more than eight paths, this design is technically unsupported.

- Do not intermix dissimilar array types or sizes

  Although the IBM Spectrum Virtualize supports an intermix of differing storage within storage pools, it is best to always use the same array model, Redundant Array of Independent Disks (RAID) mode, RAID size (RAID 5 6+P+S does not mix well with RAID 6 14+2), and drive speeds.

Rules and guidelines are no substitution for monitoring performance. Monitoring performance can provide a validation that design expectations are met, and identify opportunities for improvement.

## IBM Spectrum Virtualize performance perspectives

The software was developed by the IBM Research Group and was designed to run on commodity hardware (mass-produced Intel-based processors (CPUs) with mass-produced expansion cards) and to provide distributed cache and a scalable cluster architecture.

The performance is near linear when nodes are added into the cluster until performance eventually becomes limited by the attached components. Also, although virtualization provides significant flexibility in terms of the components that are used, it does not diminish the necessity of designing the system around the components so that it can deliver the level of performance that you want.

The key item for planning is your SAN layout. Switch vendors have slightly different planning requirements, but the end goal is that you always want to maximize the bandwidth that is available to the Storwize V7000 ports. The Storwize V7000 is one of the few devices that can drive ports to their limits on average, so it is imperative that you put significant thought into planning the SAN layout.

Essentially, performance improvements are gained by spreading the workload across a greater number of back-end resources, and by more caching, which capabilities are provided by the Storwize V7000 cluster. However, the performance of individual resources eventually becomes the limiting factor.

# Performance monitoring

In this section, we highlight several performance monitoring techniques.

## Collecting performance statistics

IBM Spectrum Virtualize is constantly collecting performance statistics. The default frequency by which files are created is 5-minute intervals. The collection interval can be changed by using the **`startstats`** command.

The statistics files (Volume, managed disk (MDisk), and Node) are saved at the end of the sampling interval, and a maximum of 16 files (each) are stored before they are overlaid in a rotating log fashion. This design provides statistics for the most recent 80-minute period if the default 5-minute sampling interval is used. IBM Spectrum Virtualize supports user-defined sampling intervals of 1 - 60 minutes.

The maximum space that is required for a performance statistics file is 1,153,482 bytes. Up to 128 (16 per each of the three types across eight nodes) different files can exist across eight Storwize V7000 nodes. This design makes the total space requirement a maximum of 147,645,694 bytes for all performance statistics from all nodes in a Storwize V7000 cluster.

**Note:** Remember this maximum of 147,645,694 bytes for all performance statistics from all nodes in a Storwize V7000 cluster when you are in time-critical situations. The required size is not otherwise important, because Storwize V7000 node hardware can map the space.

You can define the sampling interval by using the **startstats -interval 2** command to collect statistics at, for example, 2-minute intervals.

Statistics are collected at the end of each sampling period (as specified by the **-interval** parameter). These statistics are written to a file. A file is created at the end of each sampling period. Separate files are created for MDisks, volumes, and node statistics.

Use the **startstats** command to start the collection of statistics, as shown in Example A-1.

*Example A-1 The startstats command*

```
IBM_2076:ITSO_V7000:superuser>startstats -interval 5
```

This command starts statistics collection and gathers data at 5-minute intervals.

**Statistics collection:** To verify that the statistics collection is set, display the system properties again, as shown in Example A-2.

*Example A-2 Statistics collection status and frequency*

```
IBM_2076:ITSO_V7000:superuser>lssystem
statistics_status on
statistics_frequency 5
-- The output has been shortened for easier reading. --
```

The statistics collection is now started on the clustered system.

**Collection intervals:** Although more frequent collection intervals provide a more detailed view of what happens within IBM Spectrum Virtualize and Storwize V7000, they shorten the amount of time that the historical data is available on the IBM Spectrum Virtualize. For example, rather than an 80-minute period of data with the default five-minute interval, if you adjust to 2-minute intervals, you have a 32-minute period instead.

Per node statistics are collected and the sampling of the internal performance counters is coordinated across the cluster so that when a sample is taken, all nodes sample their internal counters at the same time. It is important to collect all files from all nodes for a complete analysis. Tools, such as IBM Tivoli Storage Productivity Center, perform this intensive data collection for you.

## Statistics file naming

The statistics files that are generated are written to the `/dumps/iostats/` directory. The file name is in the following formats:

- `Nm_stats_`*`<control_enclosure_id>`*`_`*`<date>`*`_`*`<time>`* for managed disk (MDisk) statistics
- `Nv_stats_`*`<control_enclosure_id>`*`_`*`<date>`*`_`*`<time>`* for virtual disks (Volumes) statistics
- `Nn_stats_`*`<control_enclosure_id>`*`_`*`<date>`*`_`*`<time>`* for node statistics
- `Nd_stats_`*`<control_enclosure_id>`*`_`*`<date>`*`_`*`<time>`* for disk drive statistics

The *`control_enclosure_id`* is of the canister on which the statistics were collected. The date is in the form *<yymmdd>* and the time is in the form *<hhmmss>*. The following example shows an MDisk statistics file name:

```
Nm_stats_78K00TM-1_161031_151832
```

Example A-3 shows typical MDisk, volume, node, and disk drive statistics file names.

*Example A-3 File names of per node statistics*

```
IBM_2076:ITSO_V7000:superuser>svcinfo lsiostatsdumps
id iostat_filename
1  Nd_stats_78K00TM-1_141031_151832
2  Nv_stats_78K00TM-1_141031_151832
3  Nv_stats_78K00TM-1_141031_151832
4  Nd_stats_78K00TM-1_141031_151832
5  Nd_stats_78K00TM-1_141031_151932
6  Nv_stats_78K00TM-1_141031_151932
7  Nv_stats_78K00TM-1_141031_151932
8  Nd_stats_78K00TM-1_141031_152032
9  Nv_stats_78K00TM-1_141031_152232
10 Nd_stats_78K00TM-1_141031_152232
11 Nv_stats_78K00TM-1_141031_152432
12 Nd_stats_78K00TM-1_141031_152432
13 Nm_stats_78K00TM-1_141031_152432
```

**Tip:** The performance statistics files can be copied from the Storwize V7000 nodes to a local drive on your workstation by using the `pscp.exe` (included with PuTTY) from an MS-DOS command line, as shown in this example:

```
C:\Program Files\PuTTY>pscp -unsafe -load ITSO_V7000
admin@10.18.227.71:/dumps/iostats/* c:\statsfiles
```

Use the **`-load`** parameter to specify the session that is defined in PuTTY.

Specify the **`-unsafe`** parameter when you use wildcards.

You can obtain PuTTY from the following website:

http://www.chiark.greenend.org.uk/~sgtatham/putty/download.html

## Real-time performance monitoring

Storwize V7000 supports real-time performance monitoring. Real-time performance statistics provide short-term status information for the Storwize V7000. The statistics are shown as graphs in the management GUI, or can be viewed from the CLI.

With system-level statistics, you can quickly view the CPU usage and the bandwidth of volumes, interfaces, and MDisks. Each graph displays the current bandwidth in megabytes per second (MBps) or I/Os operations per second (IOPS), and a view of bandwidth over time.

Each node collects various performance statistics, mostly at 5-second intervals, and the statistics that are available from the config node in a clustered environment. This information can help you determine the performance effect of a specific node. As with system statistics, node statistics help you to evaluate whether the node is operating within normal performance metrics.

Real-time performance monitoring gathers the following system-level performance statistics:

- CPU utilization
- Port utilization and I/O rates
- Volume and MDisk I/O rates
- Bandwidth

- Latency

Real-time statistics are not a configurable option and cannot be disabled.

## Real-time performance monitoring with the CLI

The **lsnodecanisterstats** and **lssystemstats** commands are available for monitoring the statistics through the CLI. Next, we show you examples of how to use them.

The **lsnodecanisterstats** command provides performance statistics for the nodes that are part of a clustered system, as shown in Example A-4 (the output is truncated and shows only part of the available statistics). You can also specify a node name in the command to limit the output for a specific node.

*Example A-4 The lsnodecanisterstats command output*

```
IBM_2076:ITSO_V7000:admin>lsnodecanisterstats
node_id node_name stat_name          stat_current stat_peak stat_peak_time
1       node1     compression_cpu_pc 0            0         161031123930
1       node1     cpu_pc             4            7         161031123925
1       node1     fc_mb              0            0         161031123930
1       node1     fc_io              56           56        161031123930
1       node1     sas_mb             0            0         161031123930
1       node1     sas_io             0            0         161031123930
1       node1     iscsi_mb           0            0         161031123930
1       node1     iscsi_io           0            0         161031123930
1       node1     write_cache_pc     0            0         161031123930
1       node1     total_cache_pc     1            1         161031123930
1       node1     vdisk_mb           0            0         161031123930
1       node1     vdisk_io           0            0         161031123930
1       node1     vdisk_ms           0            0         161031123930
1       node1     mdisk_mb           0            0         161031123930
1       node1     mdisk_io           0            0         161031123930
1       node1     mdisk_ms           0            0         161031123930
1       node1     drive_mb           0            0         161031123930
1       node1     drive_io           0            0         161031123930
1       node1     drive_ms           0            0         161031123930
1       node1     vdisk_r_mb         0            0         161031123930
.....
2       node2     compression_cpu_pc 0            0         161031123511
2       node2     cpu_pc             0            1         161031123501
2       node2     fc_mb              0            0         161031123511
2       node2     fc_io              0            0         161031123511
2       node2     sas_mb             28           60        161031123056
2       node2     sas_io             115          579       161031123106
2       node2     iscsi_mb           0            0         161031123511
2       node2     iscsi_io           0            0         161031123511
2       node2     write_cache_pc     0            0         161031123511
2       node2     total_cache_pc     79           79        161031123511
2       node2     vdisk_mb           0            0         161031123511
2       node2     vdisk_io           0            0         161031123511
2       node2     vdisk_ms           0            0         161031123511
2       node2     mdisk_mb           0            0         161031123511
2       node2     mdisk_io           0            0         161031123511
2       node2     mdisk_ms           0            0         161031123511
2       node2     drive_mb           28           65        161031123056
2       node2     drive_io           115          240       161031123056
```

```
2       node2     drive_ms           11           35         161031123401

...
```

The previous example shows statistics for the two node members of cluster ITSO_V7000. For each node, the following columns are displayed:

- stat_name: Provides the name of the statistic field
- stat_current: The current value of the statistic field
- stat_peak: The peak value of the statistic field in the last 5 minutes
- stat_peak_time: The time that the peak occurred

On the other side, the **lssystemstats** command lists the same set of statistics that is listed with the **lsnodecanisterstats** command, but representing all nodes in the cluster. The values for these statistics are calculated from the node statistics values in the following way:

- Bandwidth: Sum of bandwidth of all nodes
- Latency: Average latency for the cluster, which is calculated by using data from the whole cluster, not an average of the single node values
- IOPS: Total IOPS of all nodes
- CPU percentage: Average CPU percentage of all nodes

Example A-5 shows the resulting output of the **lssystemstats** command.

*Example A-5 The lssystemstats command output*

```
IBM_2076:ITSO_V7000:admin>lssystemstats
stat_name          stat_current stat_peak stat_peak_time
compression_cpu_pc 0            0         161031124250
cpu_pc             0            5         161031123925
fc_mb              0            0         161031124250
fc_io              60           11235     161031123856
sas_mb             39           74        161031124155
sas_io             160          542       161031124105
iscsi_mb           0            0         161031124250
iscsi_io           0            0         161031124250
write_cache_pc     0            0         161031124250
total_cache_pc     40           79        161031123921
vdisk_mb           0            0         161031124250
vdisk_io           0            0         161031124250
vdisk_ms           0            0         161031124250
mdisk_mb           0            0         161031124250
mdisk_io           0            0         161031124250
mdisk_ms           0            14        161031123930
drive_mb           39           74        161031124155
drive_io           160          297       161031124155
drive_ms           37           37        161031124250
...
```

Table A-1 has a brief description of each of the statistics that are presented by the **lssystemstats** and **lsnodecanisterstats** commands.

*Table A-1 The lssystemstats and lsnodestats statistics field name descriptions*

| Field name | Unit | Description |
|---|---|---|
| cpu_pc | Percentage | Utilization of node CPUs. |
| fc_mb | MBps | Fibre Channel (FC) bandwidth. |
| fc_io | IOPS | Fibre Channel throughput. |
| sas_mb | MBps | Serial-attached SCSI (SAS) bandwidth. |
| sas_io | IOPS | SAS throughput. |
| iscsi_mb | MBps | Internet Small Computer System Interface (iSCSI) bandwidth. |
| iscsi_io | IOPS | iSCSI throughput. |
| write_cache_pc | Percentage | Write cache fullness. Updated every 10 seconds. |
| total_cache_pc | Percentage | Total cache fullness. Updated every 10 seconds. |
| vdisk_mb | MBps | Total Volume bandwidth. |
| vdisk_io | IOPS | Total Volume throughput. |
| vdisk_ms | Milliseconds | Average Volume latency. |
| mdisk_mb | MBps | MDisk (SAN and RAID) bandwidth. |
| mdisk_io | IOPS | MDisk (SAN and RAID) throughput. |
| mdisk_ms | Milliseconds | Average MDisk latency. |
| drive_mb | MBps | Drive bandwidth. |
| drive_io | IOPS | Drive throughput. |
| drive_ms | Milliseconds (ms) | Average drive latency. |
| vdisk_w_mb | MBps | Volume write bandwidth. |
| vdisk_w_io | IOPS | Volume write throughput. |
| vdisk_w_ms | Milliseconds | Average Volume write latency. |
| mdisk_w_mb | MBps | MDisk (SAN and RAID) write bandwidth. |
| mdisk_w_io | IOPS | MDisk (SAN and RAID) write throughput. |
| mdisk_w_ms | Milliseconds | Average MDisk write latency. |
| drive_w_mb | MBps | Drive write bandwidth. |
| drive_w_io | IOPS | Drive write throughput. |
| drive_w_ms | Milliseconds | Average drive write latency. |
| vdisk_r_mb | MBps | Volume read bandwidth. |
| vdisk_r_io | IOPS | Volume read throughput. |
| vdisk_r_ms | Milliseconds | Average Volume read latency. |
| mdisk_r_mb | MBps | MDisk (SAN and RAID) read bandwidth. |
| mdisk_r_io | IOPS | MDisk (SAN and RAID) read throughput. |

| Field name | Unit | Description |
|---|---|---|
| mdisk_r_ms | Milliseconds | Average MDisk read latency. |
| drive_r_mb | MBps | Drive read bandwidth. |
| drive_r_io | IOPS | Drive read throughput. |
| drive_r_ms | Milliseconds | Average drive read latency. |

## Real-time performance statistics monitoring with the GUI

Use real-time statistics to monitor CPU utilization, volume, interface, and MDisk bandwidth of your system and nodes. Each graph represents five minutes of collected statistics and provides a means of assessing the overall performance of your system.

The real-time statistics are available from the IBM Spectrum Virtualize GUI. Click **Monitoring** → **Performance** (as shown in Figure A-1) to open the Performance Monitoring window.

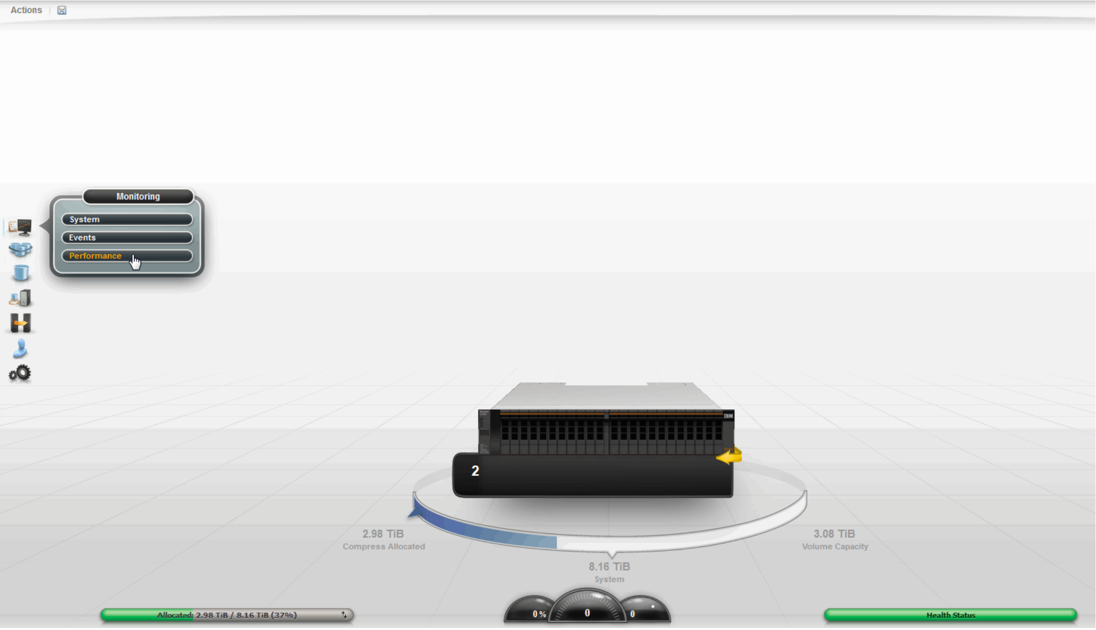


*Figure A-1 Storwize V7000 Monitoring menu*

As shown in Figure A-2 on page 672, the Performance monitoring window is divided into the following sections that provide utilization views for the following resources:

- CPU Utilization: The CPU utilization graph shows the current percentage of CPU usage and peaks in utilization. It can also display compression CPU usage for systems with compressed volumes.
- Volumes: Shows four metrics on the overall volume utilization graphics:
  - Read
  - Write
  - Read latency
  - Write latency
- Interfaces: The Interfaces graph displays data points for Fibre Channel (FC), iSCSI, serial-attached SCSI (SAS), and Internet Protocol (IP) Remote Copy interfaces. You can use this information to help determine connectivity issues that might affect performance.

  - Fibre Channel
  - iSCSI
  - SAS
  - IP Remote Copy
- MDisks: Also shows four metrics on the overall MDisks graphics:
  - Read
  - Write
  - Read latency
  - Write latency

You can use these metrics to help determine the overall performance health of the volumes and MDisks on your system. Consistent unexpected results can indicate errors in configuration, system faults, or connectivity issues.

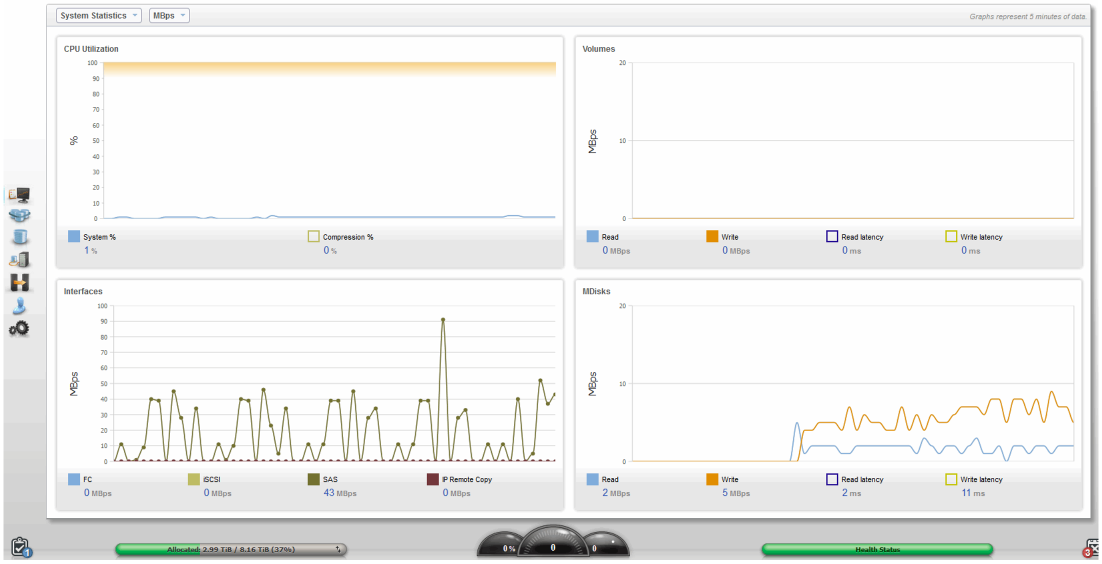


*Figure A-2 Storwize V7000 performance monitoring window*

You can also select to view performance statistics for each of the available nodes of the system, as shown in Figure A-3.

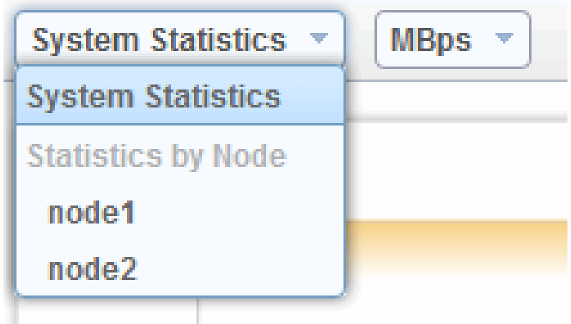


*Figure A-3 Select a system node*

You can also change the metric between MBps or IOPS, as shown in Figure A-4.

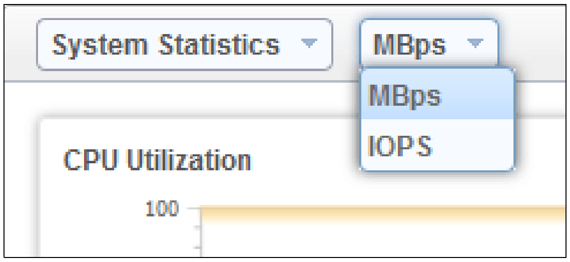


*Figure A-4 Changing to MBps or IOPS*

On any of these views, you can select any point with your cursor to know the exact value and when it occurred. When you place your cursor over the timeline, it becomes a dotted line with the various values gathered, as shown in Figure A-5.

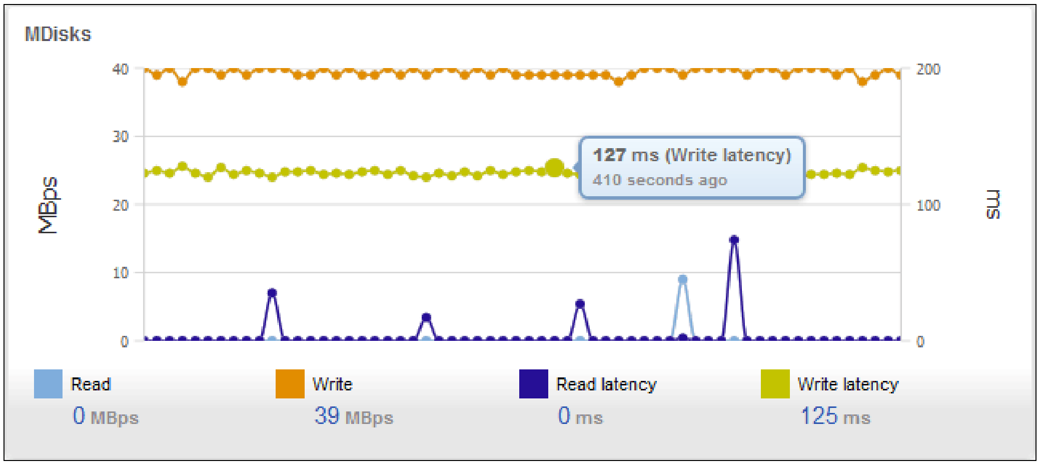


*Figure A-5 Detailed resource utilization*

For each of the resources, various values exist that you can view by selecting the value. For example, as shown in Figure A-6, the four available fields are selected for the MDisks view: Read, Write, Read latency, and Write latency. In our example, Read is not selected.

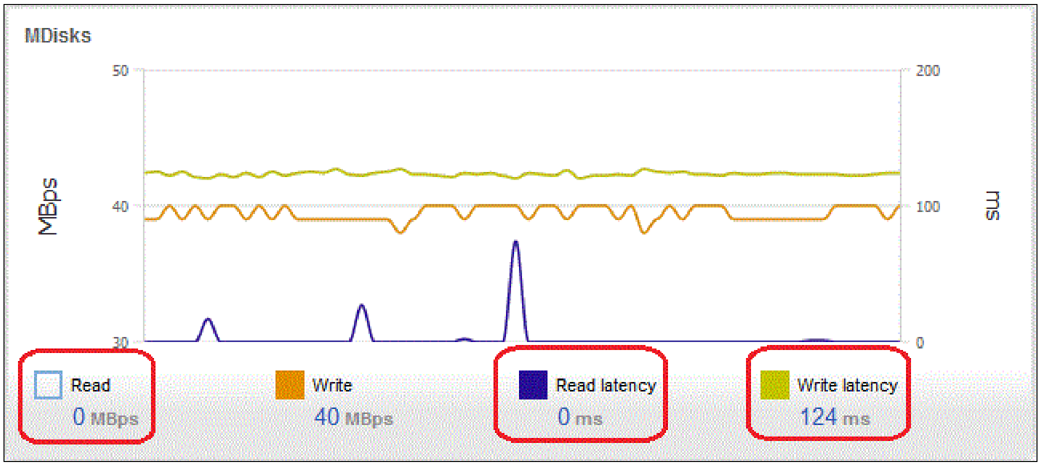


*Figure A-6 Detailed resource utilization*

## Performance data collection and Tivoli Storage Productivity Center for Disk

Although you can obtain performance statistics in standard `.xml` files, the use of `.xml` files is a less practical and more complicated method to analyze the IBM Spectrum Virtualize performance statistics. Tivoli Storage Productivity Center for Disk is the supported IBM tool to collect and analyze Storwize V7000 performance statistics.

Tivoli Storage Productivity Center for Disk is installed separately on a dedicated system, and it is not part of the IBM Spectrum Virtualize bundle.

For more information about the use of Tivoli Storage Productivity Center to monitor your storage subsystem, see:

http://www.ibm.com/systems/uk/storage/software/center/

B

# CLI setup and SAN boot

In this chapter we describe CLI setup and SAN boot.

# CLI setup

IBM Spectrum Virtualize system has a powerful CLI, which offers a few more options and flexibility as compared to the graphical user interface (GUI). This section describes how to configure a management system using secure shell (SSH) protocol to connect to IBM Spectrum Virtualize system for the purpose of issuing the commands using the CLI.

Detailed CLI information is available on the IBM Storwize V7000 Knowledge Center web page under the command-line section, which is at the following address:

hhttps://ibm.biz/BdsKgw

> **Note:** If a task completes in the GUI, the CLI command is always displayed in the details, as shown throughout this book.

## Basic setup

In the IBM Spectrum Virtualize GUI, authentication is performed by supplying a user name and password. The CLI uses a Secure Shell (SSH) to connect from the host to the IBM Spectrum Virtualize system. Either a private and a public key pair or user name and password is necessary. The following steps are required to enable CLI access with SSH keys:

- A public key and a private key are generated together as a pair.
- A public key is uploaded to the IBM Spectrum Virtualize system through the GUI.
- A client SSH tool must be configured to authenticate with the private key.
- A secure connection can be established between the client and IBM Spectrum Virtualize system.

Secure Shell is the communication vehicle between the management workstation and the IBM Spectrum Virtualize system. The SSH client provides a secure environment from which to connect to a remote machine. It uses the principles of public and private keys for authentication.

SSH keys are generated by the SSH client software. The SSH keys include a public key, which is uploaded and maintained by the clustered system, and a private key, which is kept private on the workstation that is running the SSH client. These keys authorize specific users to access the administration and service functions on the system.

Each key pair is associated with a user-defined ID string that can consist of up to 40 characters. Up to 100 keys can be stored on the system. New IDs and keys can be added, and unwanted IDs and keys can be deleted. To use the CLI, an SSH client must be installed on that system, the SSH key pair must be generated on the client system, and the client's SSH public key must be stored on the IBM Spectrum Virtualize systems.

The SSH client used in this book is PuTTY. Also, a PuTTY key generator can be used to generate the private and public key pair. The PuTTY client can be downloaded from the following address at no cost:

http://www.chiark.greenend.org.uk

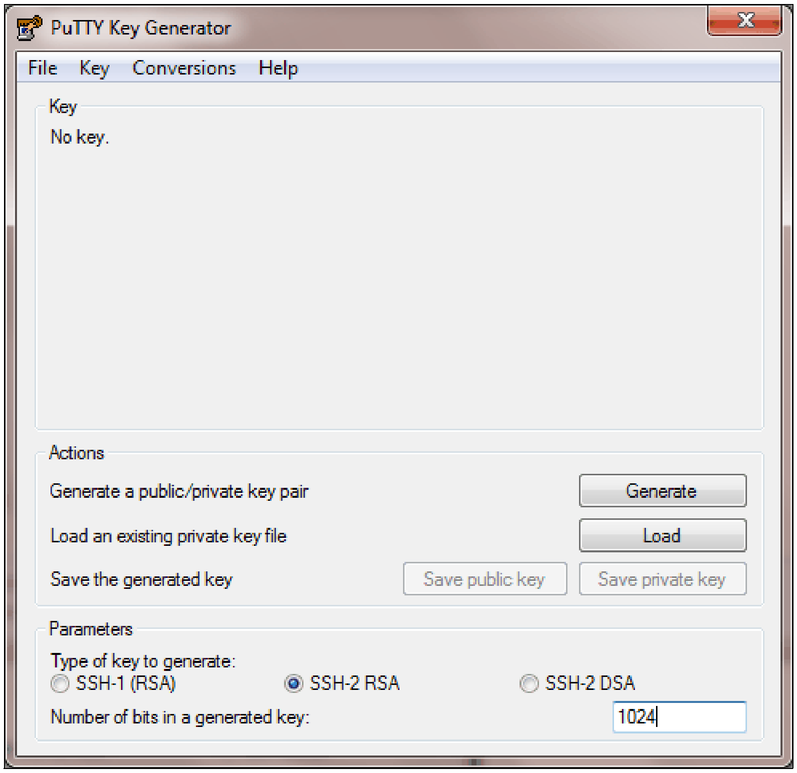
Download the following tools:

- PuTTY SSH client: **putty.exe**
- PuTTY key generator: **puttygen.exe**

## Generating a public and private key pair

To generate a public and private key pair, complete the following steps:

1. Start the PuTTY key generator to generate the public and private key pair (Figure B-1).



Key
No key.

Generate a public/private key pair
Load an existing private key file
Save the generated key

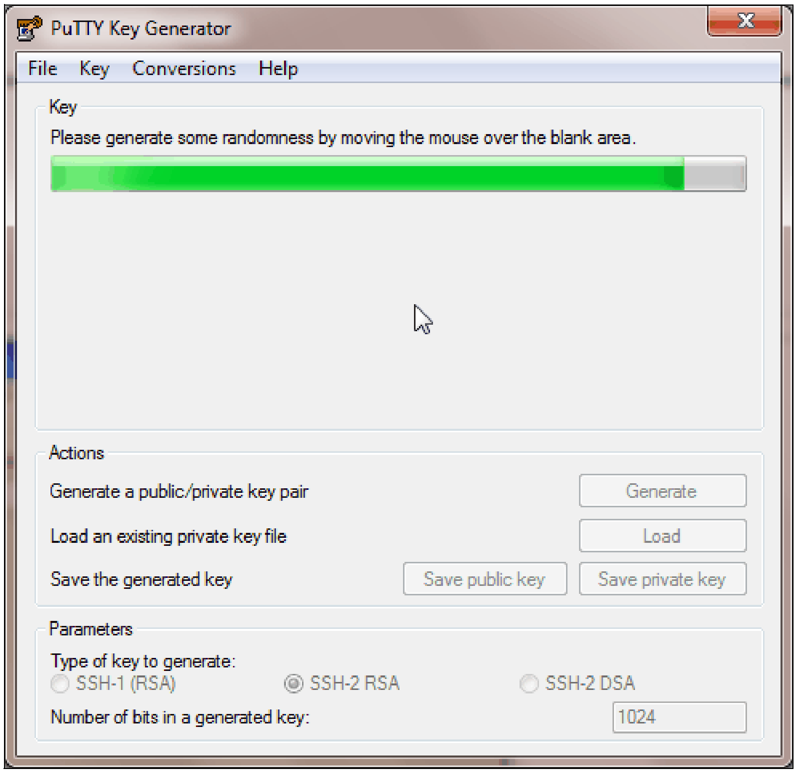
Type of key to generate:
○ SSH-1 (RSA) ● SSH-2 RSA ○ SSH-2 DSA
Number of bits in a generated key: 1024

*Figure B-1 PuTTY key generator*

Ensure that the following options are selected:

- SSH-2 RSA
- Number of bits in a generated key: 1024

2. Click **Generate** and move the cursor over the blank area to generate keys (Figure B-2).



Please generate some randomness by moving the mouse over the blank area.

Generate a public/private key pair

Load an existing private key file

Save the generated key

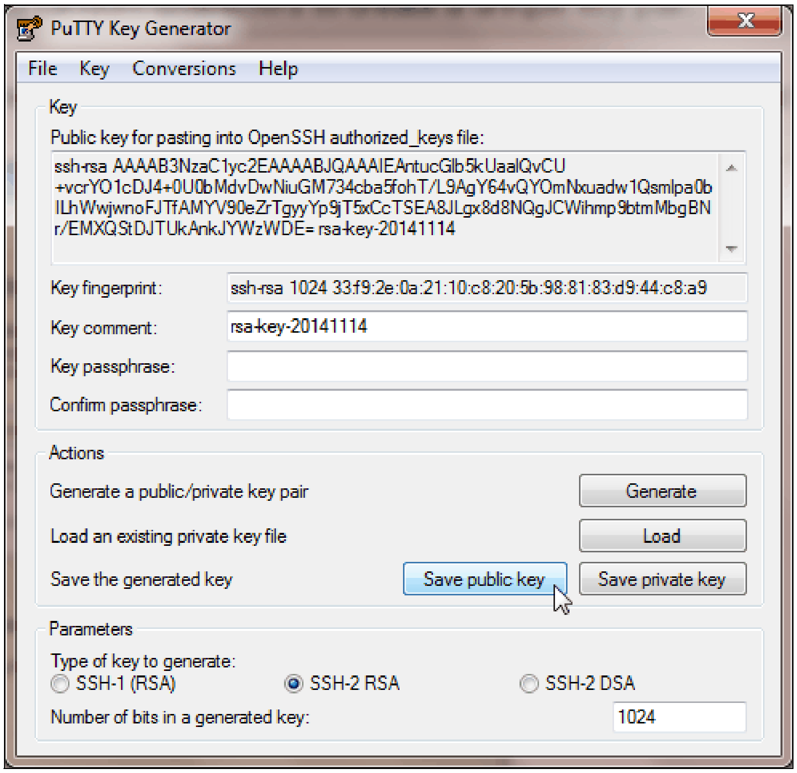
Type of key to generate:
○ SSH-1 (RSA) ● SSH-2 RSA ○ SSH-2 DSA

Number of bits in a generated key: 1024

*Figure B-2 Generate keys*

> **To generate keys**: The blank area that is indicated by the message is the large blank rectangle on the GUI inside the section of the GUI labeled Key. Continue to move the mouse pointer over the blank area until the progress bar reaches the far right. This action generates random characters to create a unique key pair.

3. After the keys are generated, save them for later use. Click **Save public key** (Figure B-3).



Public key for pasting into OpenSSH authorized_keys file:

```
ssh-rsa AAAAB3NzaC1yc2EAAAABJQAAAIEAntucGlb5kUaalQvCU
+vcrYO1cDJ4+0U0bMdvDwNiuGM734cba5fohT/L9AgY64vQYOmNxuadw1Qsmlpa0b
ILhWwjwnoFJTfAMYV90eZrTgyyYp9jT5xCcTSEA8JLgx8d8NQgJCWihmp9btmMbgBN
r/EMXQStDJTUkAnkJYWzWDE= rsa-key-20141114
```

Key fingerprint: ssh-rsa 1024 33:f9:2e:0a:21:10:c8:20:5b:98:81:83:d9:44:c8:a9

Key comment: rsa-key-20141114

Key passphrase:

Confirm passphrase:

Generate a public/private key pair

Load an existing private key file

Save the generated key

Type of key to generate:
○ SSH-1 (RSA) ● SSH-2 RSA ○ SSH-2 DSA

Number of bits in a generated key: 1024

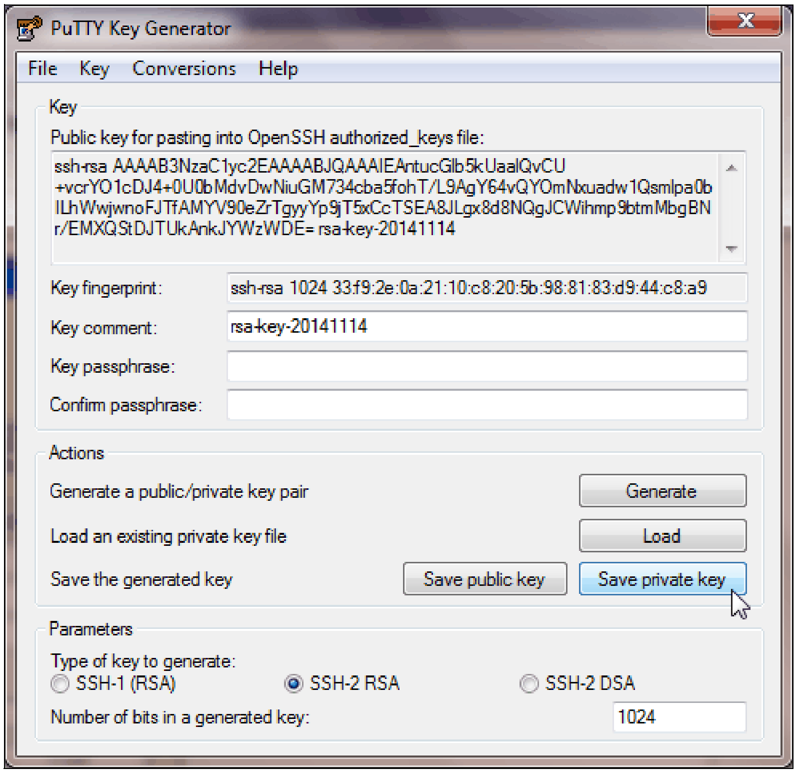
*Figure B-3 Save public key*

4. You are prompted for a name (for example, `pubkey`) and a location for the public key (for example, `C:\Support Utils\PuTTY`). Click **Save**.

   Ensure that you record the name and location, because the name and location of this SSH public key must be specified later.

   > **Public key extension:** By default, the PuTTY key generator saves the public key with no extension. Use the string `pub` for naming the public key, for example, `pubkey`, to easily differentiate the SSH public key from the SSH private key.

5. Click **Save private key** (Figure B-4).



Public key for pasting into OpenSSH authorized_keys file:

```
ssh-rsa AAAAB3NzaC1yc2EAAAABJQAAAIEAntucGlb5kUaalQvCU
+vcrYO1cDJ4+0U0bMdvDwNiuGM734cba5fohT/L9AgY64vQYOmNxuadw1Qsmlpa0b
ILhWwjwnoFJTfAMYV90eZrTgyyYp9jT5xCcTSEA8JLgx8d8NQgJCWihmp9btmMbgBN
r/EMXQStDJTUkAnkJYWzWDE= rsa-key-20141114
```

Key fingerprint: ssh-rsa 1024 33:f9:2e:0a:21:10:c8:20:5b:98:81:83:d9:44:c8:a9

Key comment: rsa-key-20141114

Key passphrase:

Confirm passphrase:

Generate a public/private key pair

Load an existing private key file

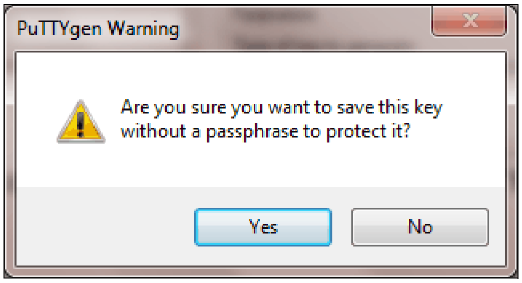
Save the generated key

Type of key to generate:
○ SSH-1 (RSA) ● SSH-2 RSA ○ SSH-2 DSA

Number of bits in a generated key: 1024

*Figure B-4 Save private key*

6. You are prompted with a warning message (Figure B-5). Click **Yes** to save the private key without a passphrase.



Are you sure you want to save this key without a passphrase to protect it?

*Figure B-5 Confirm the security warning*

7. When prompted, enter a name (for example, `icat`), select a secure place as the location, and click **Save**.

> **Key generator:** The PuTTY key generator saves the PuTTY private key (PPK) with the `.ppk` extension.

8. Close the PuTTY key generator.

### B.0.0.1 Uploading the SSH public key to the IBM Spectrum Virtualize system

After you create your SSH key pair, upload your SSH public key onto the IBM Spectrum Virtualize system. Complete the following steps:

1. Open the user section in the GUI (Figure B-6).

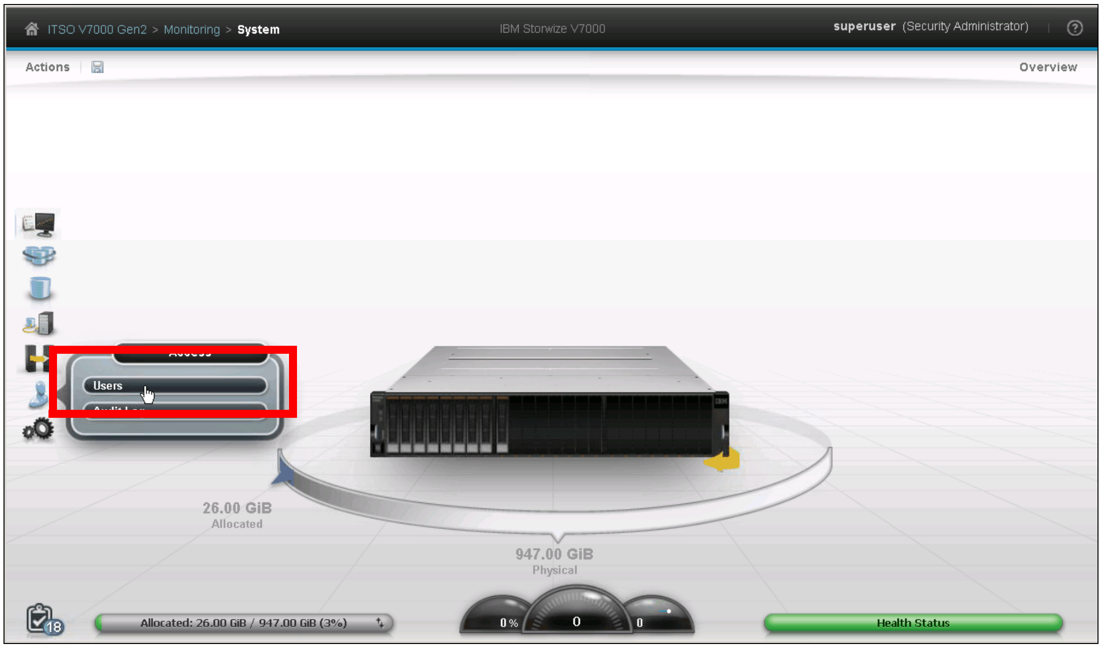
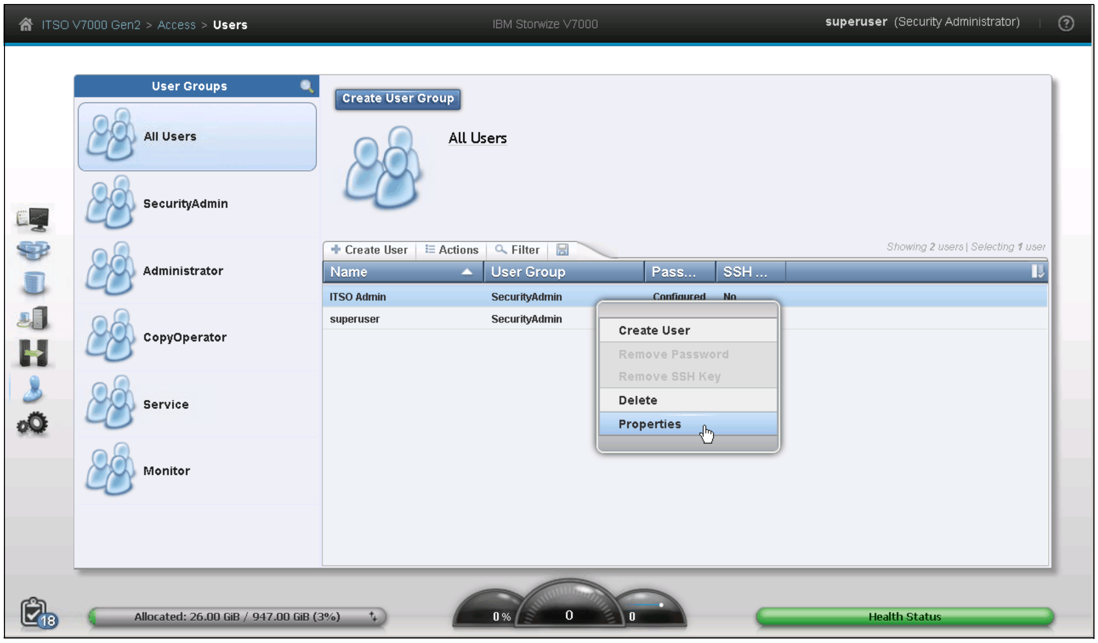

ITSO V7000 Gen2 > Monitoring > System
IBM Storwize V7000
superuser (Security Administrator)
Actions
Overview
Access
Users
26.00 GiB
Allocated
947.00 GiB
Physical
18
Allocated: 26.00 GiB / 947.00 GiB (3%)
0 %
0
0
Health Status

Figure B-6 Open user section

2. Right-click the user name for which you want to upload the key and click **Properties** (Figure B-7).



| Name | User Group | Pass... | SSH ... |
|---|---|---|---|
| ITSO Admin | SecurityAdmin | Configured | No |
| superuser | SecurityAdmin | | |

Figure B-7 Superuser properties

3. To upload the public key, click **Browse**, and select the folder where you stored the public SSH key (Figure B-8).

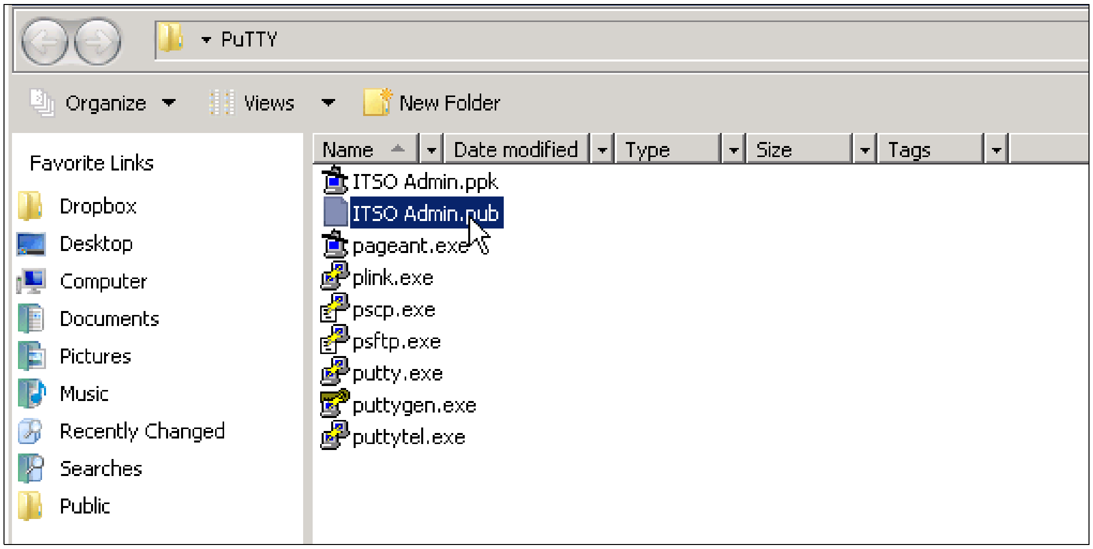
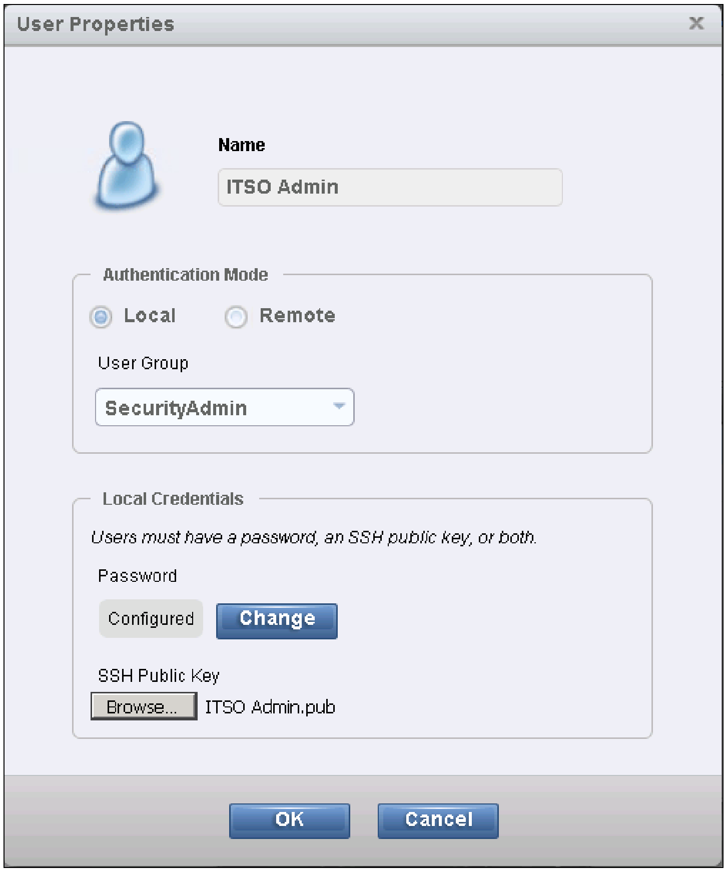


*Figure B-8 Selection of the public SSH key*

4. Select your public key, and click **OK** (Figure B-9).



Name: ITSO Admin

● Local ○ Remote

User Group: SecurityAdmin

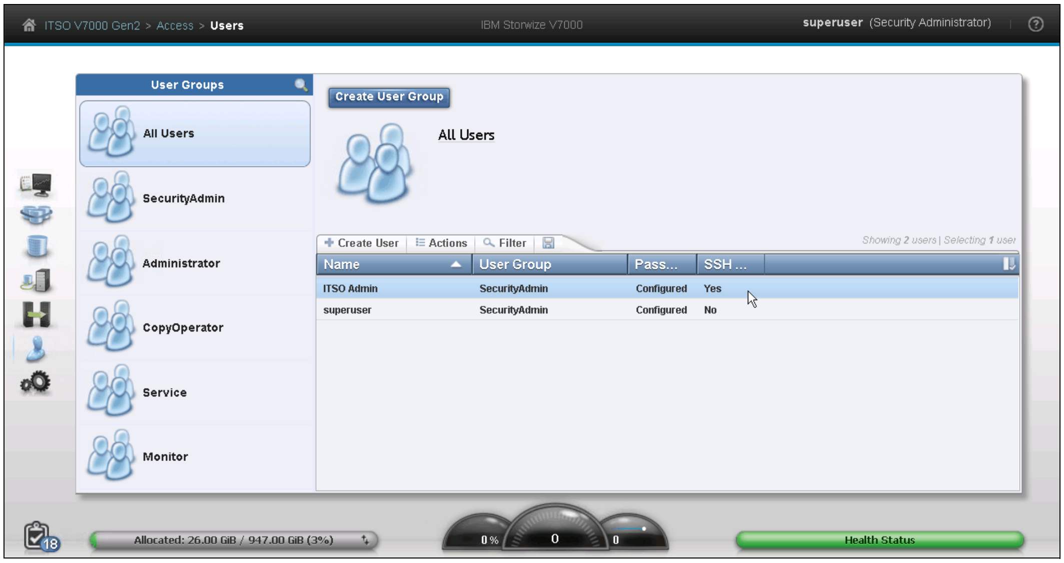
*Users must have a password, an SSH public key, or both.*

Password: Configured

SSH Public Key: Browse... ITSO Admin.pub

*Figure B-9 Select public key*

5. Click **OK** and the key is uploaded.
6. Check in the GUI if the SSH key is successfully imported. See Figure B-10.



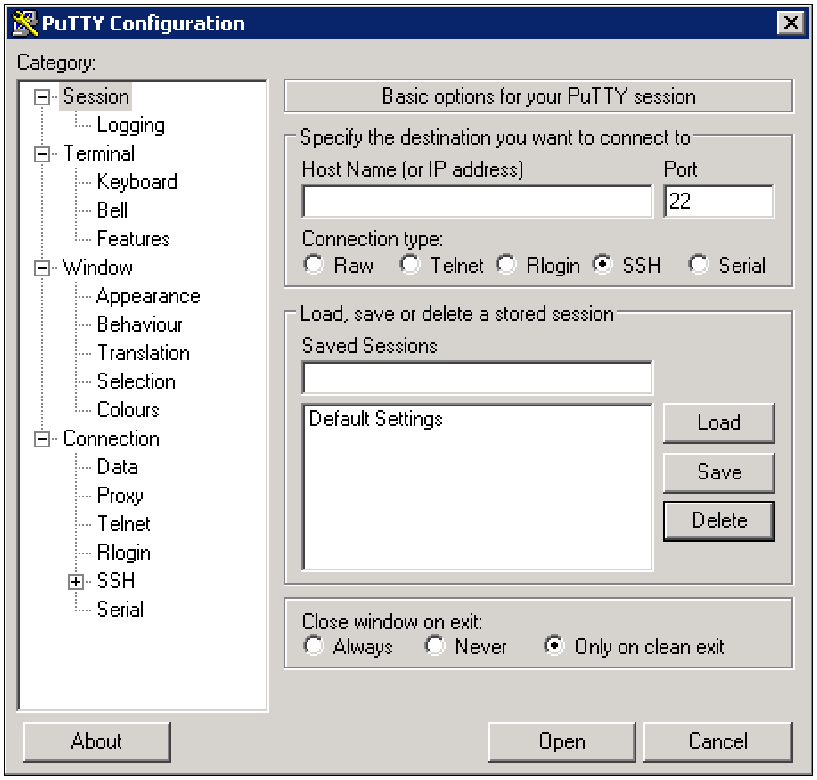
| Name | User Group | Pass... | SSH ... |
|---|---|---|---|
| ITSO Admin | SecurityAdmin | Configured | Yes |
| superuser | SecurityAdmin | Configured | No |

*Figure B-10 SSH Key successful imported*

## Configuring the SSH client

Before the CLI can be used, the SSH client must be configured as follows:

1. Start PuTTY. The PuTTY Configuration window opens (Figure B-11).



Specify the destination you want to connect to

Host Name (or IP address)

Port: 22

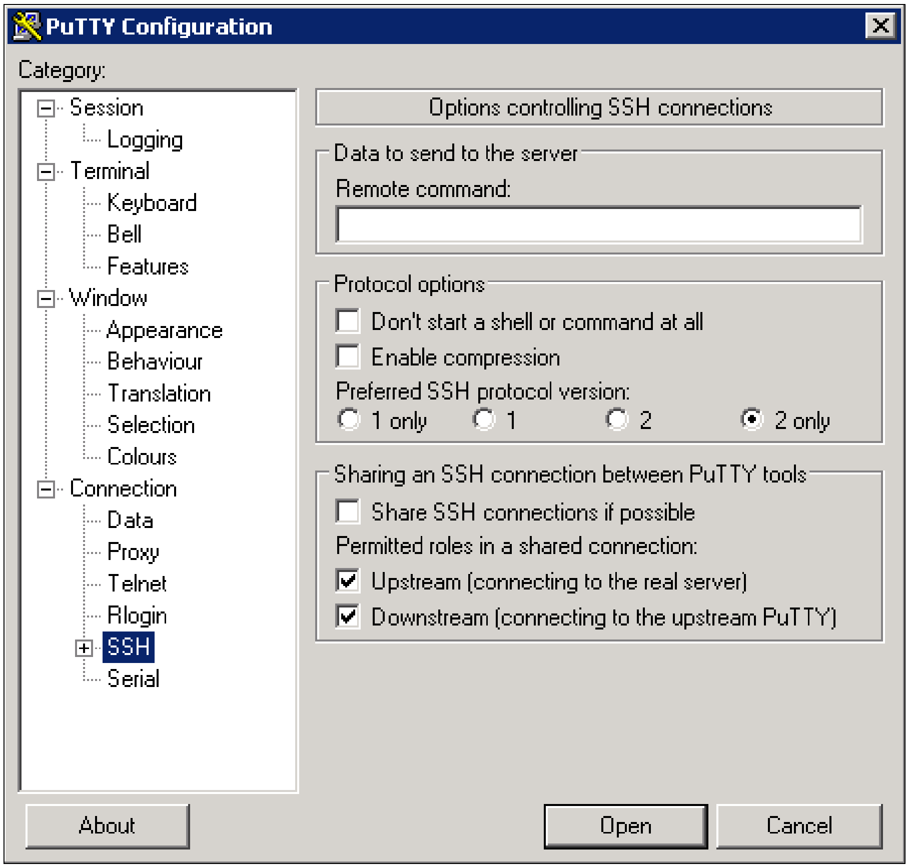
Connection type: ○ Raw ○ Telnet ○ Rlogin ● SSH ○ Serial

Close window on exit: ○ Always ○ Never ● Only on clean exit

*Figure B-11 PuTTY*

2. In the right pane, select **SSH** as the connection type. Under the **Close window on exit** section, select **Only on clean exit**, which ensures that if any connection errors occur, they are displayed on the user's window.

3. In the Category pane, on the left side of the PuTTY Configuration window (Figure B-12), click **Connection** → **SSH** to open the PuTTY SSH Configuration window.



Data to send to the server

Remote command:

Protocol options

☐ Don't start a shell or command at all

☐ Enable compression

Preferred SSH protocol version:

○ 1 only ○ 1 ○ 2 ● 2 only

Sharing an SSH connection between PuTTY tools

☐ Share SSH connections if possible

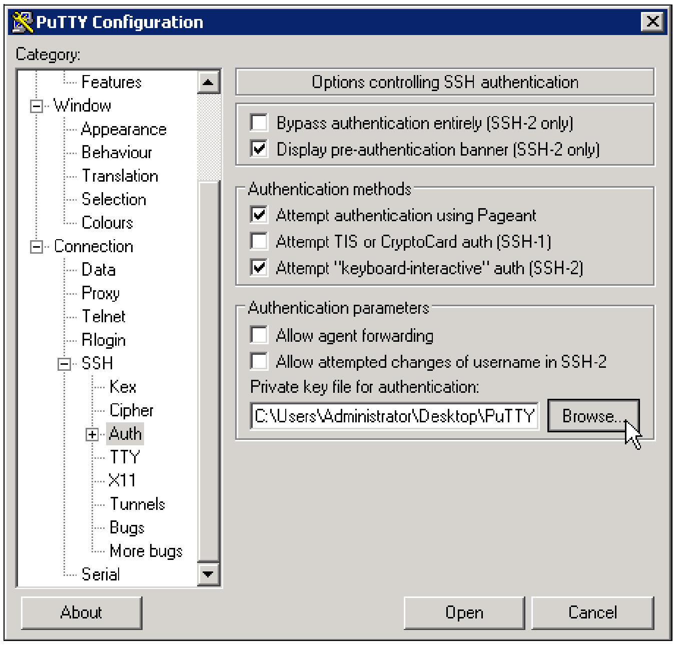
Permitted roles in a shared connection:

☑ Upstream (connecting to the real server)

☑ Downstream (connecting to the upstream PuTTY)

*Figure B-12 SSH protocol version 2*

4. Under the **Preferred SSH protocol version** section, select **2**.
5. In the Category pane on the left, click **Connection** → **SSH** → **Auth**. More options are displayed for controlling SSH authentication.
6. In the **Private key file for authentication** field (Figure B-13), either browse to or type the fully qualified directory path and file name of the SSH client private key file, which was created previously (for example, `C:\Support Utils\PuTTY\icat.PPK`).



Options controlling SSH authentication

☐ Bypass authentication entirely (SSH-2 only)
☑ Display pre-authentication banner (SSH-2 only)

Authentication methods

☑ Attempt authentication using Pageant
☐ Attempt TIS or CryptoCard auth (SSH-1)
☑ Attempt "keyboard-interactive" auth (SSH-2)

Authentication parameters

☐ Allow agent forwarding
☐ Allow attempted changes of username in SSH-2

Private key file for authentication:
C:\Users\Administrator\Desktop\PuTTY

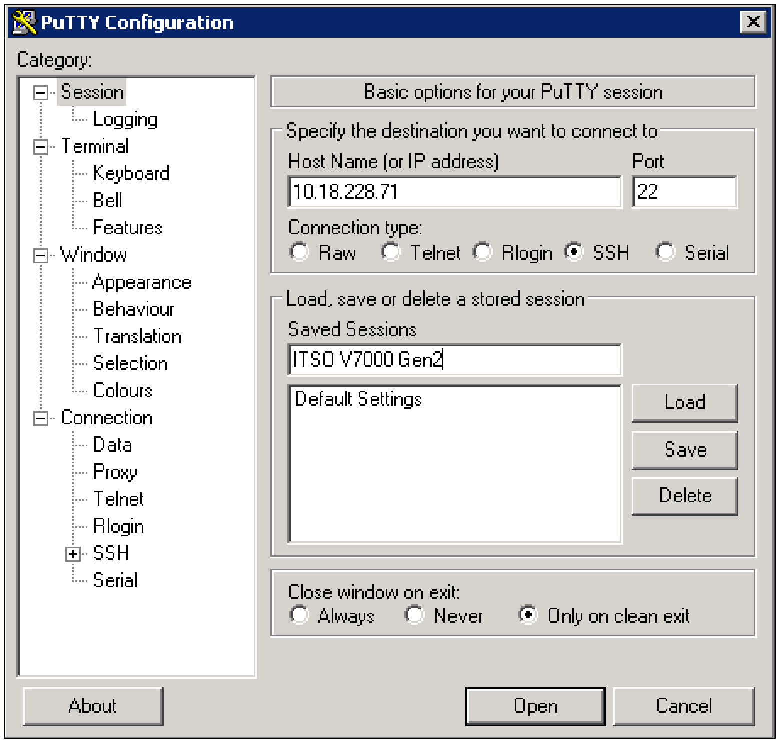
*Figure B-13 SSH authentication*

7. In the Category pane, click **Session** to return to the Basic options for your PuTTY session view (Figure B-14 on page 686).

8. Enter the following information in these fields (Figure B-14) in the right pane:

   – Host Name. Specify the host name or system Internet Protocol (IP) address of the IBM Spectrum Virtualize system.

   – Saved Sessions. Enter a session name.



Specify the destination you want to connect to

Host Name (or IP address): 10.18.228.71

Port: 22

Connection type: ○ Raw ○ Telnet ○ Rlogin ● SSH ○ Serial

Load, save or delete a stored session

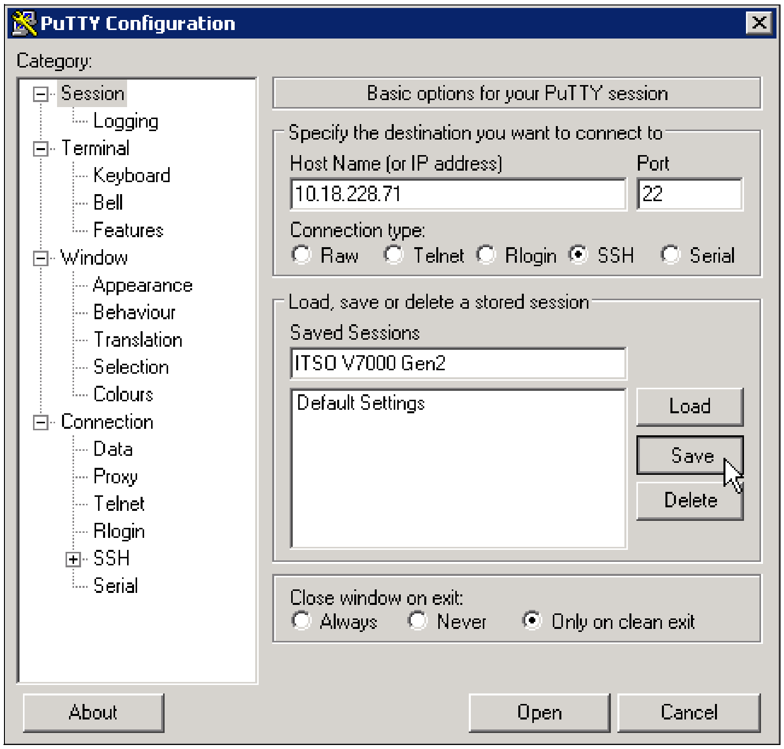
Saved Sessions: ITSO V7000 Gen2

Default Settings

Close window on exit: ○ Always ○ Never ● Only on clean exit

*Figure B-14 Enter session information*

9. Click **Save** to save the new session (Figure B-15).



Specify the destination you want to connect to

Host Name (or IP address): 10.18.228.71

Port: 22

Connection type: ○ Raw ○ Telnet ○ Rlogin ● SSH ○ Serial

Load, save or delete a stored session

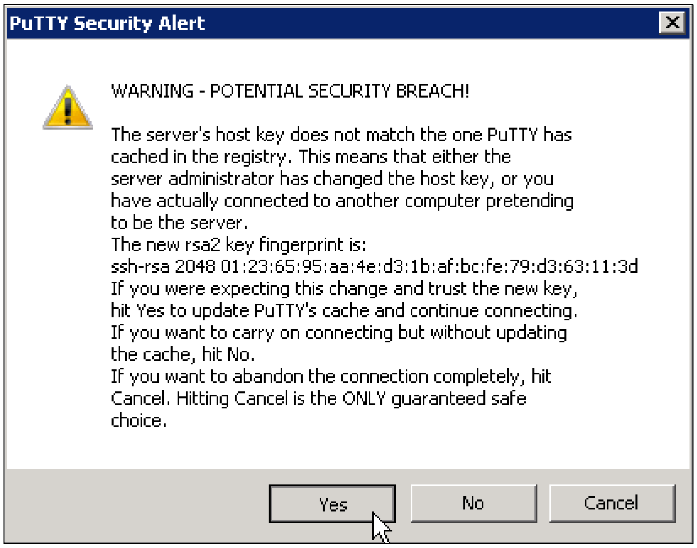
Saved Sessions: ITSO V7000 Gen2

Default Settings

Close window on exit: ○ Always ○ Never ● Only on clean exit

*Figure B-15 Save the new session*

10. Select the new session and click **Open** to connect to the IBM Spectrum Virtualize system. A PuTTY Security Alert opens. Confirm it by clicking **Yes** (Figure B-16).



WARNING - POTENTIAL SECURITY BREACH!

The server's host key does not match the one PuTTY has cached in the registry. This means that either the server administrator has changed the host key, or you have actually connected to another computer pretending to be the server.
The new rsa2 key fingerprint is:
ssh-rsa 2048 01:23:65:95:aa:4e:d3:1b:af:bc:fe:79:d3:63:11:3d
If you were expecting this change and trust the new key, hit Yes to update PuTTY's cache and continue connecting.
If you want to carry on connecting but without updating the cache, hit No.
If you want to abandon the connection completely, hit Cancel. Hitting Cancel is the ONLY guaranteed safe choice.

*Figure B-16 Confirm Security Alert*

11. PuTTY now connects to the system and prompts you for a user name to log in as. Enter `ITSO Admin` as the user name (Example B-1) and click Enter.

*Example: B-1 Enter user name*

```
login as: ITSO Admin
Authenticating with public key "putty public key"
IBM_2076:ITSO V7000Gen2:ITSO Admin>
```

You have now completed the tasks to configure the CLI for IBM Spectrum Virtualize system administration.

# SAN Boot

IBM Spectrum Virtualize system supports SAN Boot for AIX, Windows, VMware, and many other operating systems. This section describes general steps for SAN Boot configuration from the Spectrum Virtualize system for Microsoft Windows and VMware as a reference. For the latest information on SAN Boot support, please refer to the IBM Storwize V7000 interoperability matrix at this address:

http://www.ibm.com/systems/support/storage/ssic/interoperability.wss

The IBM Knowledge Center for IBM Storwize V7000 has further information about SAN Boot in combination with various operating systems. For more information, please refer to:

https://ibm.biz/BdsKgU

If the host operating system is using IBM Subsystem Device Driver (SDD) for multipathing then, to configure SAN Boot using SDD, please refer to the *IBM Multipath Subsystem Device Driver User’s Guide*, which is available at the following address:

https://ibm.biz/BdsxWp

## Enabling SAN Boot for Microsoft Windows

Complete the following procedure if you want to install Windows host using SAN Boot:

1. Configure the IBM Spectrum Virtualize system so that only the boot volume is mapped to the host.
2. Configure the Fibre Channel (FC) SAN so that the host only sees one IBM Spectrum Virtualize system node port. Multiple paths during installation are not supported.
3. Configure and enable the host bus adapter (HBA) basic input/output system (BIOS).
4. Install the operating system using the normal procedure, selecting the volume as the partition on which to install.

> **HBAs:** You might need to load an additional HBA device driver during installation, depending on your Windows version and the HBA type.

5. Install Subsystem Device Driver device-specific module (SDDDSM) after the installation has completed.
6. Modify your SAN zoning to allow multiple paths.
7. Check your host to see if all paths are available.
8. Set redundant boot devices in the HBA BIOS to enable the host to boot when its original path has failed.

## Enabling SAN Boot for VMware

Complete the following steps if you want to install a VMware ESX host using SAN Boot:

1. Configure the IBM Spectrum Virtualize system so that only the boot volume is mapped to the host.
2. Configure the FC SAN so that the host only sees one IBM Spectrum Virtualize system node port. Multiple paths during installation are not supported.
3. Configure and enable the HBA BIOS.
4. Install the operating system using the normal procedure, selecting the volume as the partition on which to install.

> **HBAs:** You might need to load an additional HBA device driver during installation, depending on your ESX level and the HBA type.

5. Modify your SAN zoning to allow multiple paths.
6. Check your host if all paths are available and modify the multipath policy if required.

C

# Terminology

In this appendix, we define the IBM Storwize V7000 terms that are commonly used in this book.

To see the complete set of terms that relate to the V7000, see the Glossary section of the IBM Knowledge Center for IBM Storwize V7000, which is available at the following website:

http://www.ibm.com/support/knowledgecenter/en/ST3FR7

# Commonly encountered terms

This appendix includes the following Storwize V7000 terminology.

## Array

An ordered collection, or group, of physical devices (disk drive modules) that are used to define logical volumes or devices. An array is a group of drives designated to be managed with a Redundant Array of Independent Disks (RAID).

## Asymmetric virtualization

Asymmetric virtualization is a virtualization technique in which the virtualization engine is outside the data path and performs a metadata-style service. The metadata server contains all the mapping and locking tables, and the storage devices contain only data. See also “Symmetric virtualization” on page 703.

## Asynchronous replication

Asynchronous replication is a type of replication in which control is given back to the application as soon as the write operation is made to the source volume. Later, the write operation is made to the target volume. See also “Synchronous replication” on page 703.

## Automatic data placement mode

Automatic data placement mode is an Easy Tier operating mode in which the host activity on all the volume extents in a pool are “measured”, a migration plan is created, and then automatic extent migration is performed.

## Back end

See “Front end and back end” on page 695.

## Caching I/O Group

The caching I/O Group is the I/O Group in the system that performs the cache function for a volume.

## Call home

Call home is a communication link that is established between a product and a service provider. The product can use this link to call IBM or another service provider when the product requires service. With access to the machine, service personnel can perform service tasks, such as viewing error and problem logs or initiating trace and dump retrievals.

## Canister

A canister is a single processing unit within a storage system.

## Capacity licensing

Capacity licensing is a licensing model that licenses features with a price-per-terabyte model. Licensed features are FlashCopy, Metro Mirror, Global Mirror, and virtualization. See also “FlashCopy” on page 694, “Metro Mirror” on page 698, and “Virtualization” on page 704.

## Chain

A set of enclosures that are attached to provide redundant access to the drives inside the enclosures. Each control enclosure can have one or more chains.

## Challenge Handshake Authentication Protocol

Challenge Handshake Authentication Protocol (CHAP) is an authentication protocol that protects against eavesdropping by encrypting the user name and password.

## Channel extender

A channel extender is a device that is used for long-distance communication that connects other storage area network (SAN) fabric components. Generally, channel extenders can involve protocol conversion to asynchronous transfer mode (ATM), Internet Protocol (IP), or another long-distance communication protocol.

## Child pool

Administrators can use child pools to control capacity allocation for volumes that are used for specific purposes. Rather than being created directly from managed disks (MDisks), child pools are created from existing capacity that is allocated to a parent pool. As with parent pools, volumes can be created that specifically use the capacity that is allocated to the child pool. Child pools are similar to parent pools with similar properties. Child pools can be used for volume copy operation. Also, see “Parent pool” on page 699.

## Cloud Container

Cloud Container is a virtual object that includes all of the elements, components or data that are common to a specific application or data.

## Cloud Service Provider

Cloud Service Provider (CSP) is the company or organisation that provides off- and on-premises cloud services such as storage, server, network, etc. IBM Spectrum Virtualize has built in software capabilities to interact with Cloud Providers such as IBM Softlayer, Amazon S3 and deployments of OpenStack Swift.

## Cloud Tenant

Cloud Tenant is a group or an instance that provides common access with the specific privileges to a object, software or data source.

## Clustered system (IBM Storwize V7000)

A clustered system, which was known as a cluster, is a group of up to eight Storwize V7000 canisters that presents a single configuration, management, and service interface to the user.

## Cold extent

A cold extent is an extent of a volume that does not get any performance benefit if it is moved from a hard disk drive (HDD) to a Flash disk. A cold extent also refers to an extent that needs to be migrated onto an HDD if it is on a Flash disk drive.

## Compression

Compression is a function that removes repetitive characters, spaces, strings of characters, or binary data from the data that is being processed and replaces characters with control characters. Compression reduces the amount of storage space that is required for data.

## Compression accelerator

A compression accelerator is hardware onto which the work of compression is offloaded from the microprocessor.

## Configuration node

While the cluster is operational, a single node in the cluster is appointed to provide configuration and service functions over the network interface. This node is termed the configuration node. This configuration node manages the data that describes the clustered-system configuration and provides a focal point for configuration commands. If the configuration node fails, another node in the cluster transparently assumes that role.

## Consistency Group

A Consistency Group is a group of copy relationships between virtual volumes or data sets that are maintained with the same time reference so that all copies are consistent in time. A Consistency Group can be managed as a single entity.

## Container

A container is a software object that holds or organizes other software objects or entities.

## Contingency capacity

For thin-provisioned volumes that are configured to automatically expand, the unused real capacity that is maintained. For thin-provisioned volumes that are not configured to automatically expand, the difference between the used capacity and the new real capacity.

## Copied state

Copied is a FlashCopy state that indicates that a copy was triggered after the copy relationship was created. The Copied state indicates that the copy process is complete and the target disk has no further dependency on the source disk. The time of the last trigger event is normally displayed with this status.

## Counterpart SAN

A counterpart SAN is a non-redundant portion of a redundant SAN. A counterpart SAN provides all of the connectivity of the redundant SAN, but without the 100% redundancy. Storwize V7000 canisters are typically connected to a “redundant SAN” that is made up of two counterpart SANs. A counterpart SAN is often called a SAN fabric.

## Cross-volume consistency

A consistency group property that ensures consistency between volumes when an application issues dependent write operations that span multiple volumes.

## Data consistency

Data consistency is a characteristic of the data at the target site where the dependent write order is maintained to ensure the recoverability of applications.

## Data encryption key

The data encryption key is used to encrypt data and it is created automatically when an encrypted object, such as an array, a pool, or a child pool, is created. It is stored in secure memory and it cannot be viewed or changed. The data encryption key is encrypted using the master access key.

## Data migration

Data migration is the movement of data from one physical location to another physical location without the disruption of application I/O operations.

## Dependent write operation

A write operation that must be applied in the correct order to maintain cross-volume consistency.

## Directed Maintenance Procedure

The fix procedures, which are also known as Directed Maintenance Procedures (DMPs), ensure that you fix any outstanding errors in the error log. To fix errors, from the Monitoring panel, click **Events**. The Next Recommended Action is displayed at the top of the Events window. Select **Run This Fix Procedure** and follow the instructions.

## Discovery

The automatic detection of a network topology change, for example, new and deleted nodes or links.

## Disk tier

MDisks (logical unit numbers (LUNs)) that are presented to the Storwize V7000 cluster likely have different performance attributes because of the type of disk or RAID array on which they are installed. The MDisks can be on 15,000 revolutions per minute (RPM) Fibre Channel (FC) or serial-attached SCSI (SAS) disk, Nearline SAS, or Serial Advanced Technology Attachment (SATA), or even Flash Disks. Therefore, a storage tier attribute is assigned to each MDisk and the default is generic_hdd.

## Distributed RAID

An alternative RAID scheme where the number of drives that are used to store the array can be greater than the equivalent, typical RAID scheme. The same data stripes are distributed across a greater number of drives, which increases the opportunity for parallel I/O and

## Easy Tier

Easy Tier is a volume performance function within the Storwize family that provides automatic data placement of a volume’s extents in a multitiered storage pool. The pool normally contains a mix of Flash Disks and HDDs. Easy Tier measures host I/O activity on the volume’s extents and migrates hot extents onto the Flash Disks to ensure the maximum performance.

## Encryption key

The encryption key, also known as master access key, is created and stored on USB flash drives or on a key server when encryption is enabled. The master access key is used to decrypt the data encryption key.

## Encryption key server

An internal or external system that receives and then serves existing encryption keys or certificates to a storage system.

## Encryption of data at rest

Encryption of data at rest is the inactive encryption data that is stored physically on the storage system.

## Evaluation mode

The evaluation mode is an Easy Tier operating mode in which the host activity on all the volume extents in a pool are “measured” only. No automatic extent migration is performed.

## Event (error)

An event is an occurrence of significance to a task or system. Events can include the completion or failure of an operation, user action, or a change in the state of a process.

## Event code

An event code is a value that is used to identify an event condition to a user. This value might map to one or more event IDs or to values that are presented on the service panel. This value is used to report error conditions to IBM and to provide an entry point into the service guide.

## Event ID

An event ID is a value that is used to identify a unique error condition that was detected by the Storwize V7000. An event ID is used internally in the cluster to identify the error.

## Excluded condition

The excluded condition is a status condition. It describes an MDisk that the Storwize V7000 decided is no longer sufficiently reliable to be managed by the cluster. The user must issue a command to include the MDisk in the cluster-managed storage.

## Extent

An extent is a fixed-size unit of data that is used to manage the mapping of data between MDisks and volumes. The size of the extent can range 16 MB - 8 GB in size.

## External storage

External storage refers to managed disks (MDisks) that are SCSI logical units that are presented by storage systems that are attached to and managed by the clustered system.

## Failback

Failback is the restoration of an appliance to its initial configuration after the detection and repair of a failed network or component.

## Failover

Failover is an automatic operation that switches to a redundant or standby system or node in a software, hardware, or network interruption. See also “Failback”.

## Feature activation code

An alphanumeric code that activates a licensed function on a product.

## Fibre Channel port logins

Fibre Channel (FC) port logins refer to the number of hosts that can see any one V7000 port. The Storwize V7000 has a maximum limit per node port of FC logins that are allowed.

## Field-replaceable unit

Field-replaceable units (FRUs) are individual parts that are replaced entirely when any one of the unit’s components fails. They are held as spares by the IBM service organization.

## FlashCopy

FlashCopy refers to a point-in-time copy where a virtual copy of a volume is created. The target volume maintains the contents of the volume at the point in time when the copy was established. Any subsequent write operations to the source volume are not reflected on the target volume.

## FlashCopy mapping

A FlashCopy mapping is a continuous space on a direct-access storage volume, which is occupied by or reserved for a particular data set, data space, or file.

## FlashCopy relationship

See “FlashCopy mapping”.

## FlashCopy service

FlashCopy service is a copy service that duplicates the contents of a source volume on a target volume. In the process, the original contents of the target volume are lost. See also “Point-in-time copy” on page 699.

## Flash drive

A data storage device that uses solid-state memory to store persistent data.

## Flash module

A modular hardware unit containing flash memory, one or more flash controllers, and associated electronics.

## Front end and back end

The Storwize V7000 takes MDisks to create pools of capacity from which volumes are created and presented to application servers (hosts). The MDisks are in the controllers at the back end of the Storwize V7000 and in the Storwize V7000 to the back-end controller zones. The volumes that are presented to the hosts are in the front end of the Storwize V7000.

## Global Mirror

Global Mirror (GM) is a method of asynchronous replication that maintains data consistency across multiple volumes within or across multiple systems. Global Mirror is generally used where distances between the source site and target site cause increased latency beyond what the application can accept.

## Global Mirror with change volumes

Change volumes are used to record changes to the primary and secondary volumes of a remote copy relationship. A FlashCopy mapping exists between a primary and its change volume and a secondary and its change volume.

## Grain

A grain is the unit of data that is represented by a single bit in a FlashCopy bitmap (64 KiB or 256 KiB) in the Storwize V7000. A grain is also the unit to extend the real size of a thin-provisioned volume (32 KiB, 64 KiB, 128 KiB, or 256 KiB).

## Hop

One segment of a transmission path between adjacent nodes in a routed network.

## Host bus adapter

A host bus adapter (HBA) is an interface card that connects a server to the SAN environment through its internal bus system, for example, PCI Express.

## Host ID

A host ID is a numeric identifier that is assigned to a group of host FC ports or Internet Small Computer System Interface (iSCSI) host names for LUN mapping. For each host ID, SCSI IDs are mapped to volumes separately. The intent is to have a one-to-one relationship between hosts and host IDs, although this relationship cannot be policed.

## Host mapping

Host mapping refers to the process of controlling which hosts have access to specific volumes within a cluster (host mapping is equivalent to LUN masking).

## Hot extent

A hot extent is a frequently accessed volume extent that gets a performance benefit if it is moved from an HDD onto a Flash Disk.

## HyperSwap

Pertaining to a function that provides continuous, transparent availability against storage errors and site failures, and is based on synchronous replication.

## Image mode

Image mode is an access mode that establishes a one-to-one mapping of extents in the storage pool (existing LUN or (image mode) MDisk) with the extents in the volume.

## Image volume

An image volume is a volume in which a direct block-for-block translation exists from the managed disk (MDisk) to the volume.

## I/O Group

Each pair of Storwize V7000 canisters is known as an input/output (I/O) Group. An I/O Group has a set of volumes that are associated with it that are presented to host systems. Each Storwize V7000 canister is associated with exactly one I/O Group. The canister in an I/O Group provide a failover and failback function for each other.

## Internal storage

Internal storage refers to an array of managed disks (MDisks) and drives that are held in Storwize V7000 enclosures.

## Internet Small Computer System Interface qualified name

Internet Small Computer System Interface (iSCSI) qualified name (IQN) refers to special names that identify both iSCSI initiators and targets. IQN is one of the three name formats that is provided by iSCSI. The IQN format is `iqn.<yyyy-mm>.<reversed domain name>`. For example, the default for a Storwize V7000 canister can be in the following format:

```
iqn.1986-03.com.ibm:2076.<clustername>.<nodename>
```

## Internet storage name service

The Internet storage name service (iSNS) protocol that is used by a host system to manage iSCSI targets and the automated iSCSI discovery, management, and configuration of iSCSI and FC devices. It was defined in Request for Comments (RFC) 4171.

## Inter-switch link hop

An inter-switch link (ISL) is a connection between two switches and counted as one ISL hop. The number of hops is always counted on the shortest route between two N-ports (device connections). In a Storwize V7000 environment, the number of ISL hops is counted on the shortest route between the pair of canister that are farthest apart. The Storwize V7000 supports a maximum of three ISL hops.

## Input/output group

A collection of volumes and canister relationships that present a common interface to host systems. Each pair of canister is known as an input/output (I/O) group.

## iSCSI initiator

An initiator functions as an iSCSI client. An initiator typically serves the same purpose to a computer as a SCSI bus adapter would, except that, instead of physically cabling SCSI devices (like hard drives and tape changers), an iSCSI initiator sends SCSI commands over an IP network.

## iSCSI session

An iSCSI Initiator and an iSCSI Target talk with each other and this conversation called an iSCSI Session.

## iSCSI target

An iSCSI target is a storage resource located on an Internet Small Computer System Interface (iSCSI) server.

## Latency

The time interval between the initiation of a send operation by a source task and the completion of the matching receive operation by the target task. More generally, latency is the time between a task initiating data transfer and the time that transfer is recognized as complete at the data destination.

## Least recently used

Least recently used (LRU) pertains to an algorithm used to identify and make available the cache space that contains the data that was least recently used.

## Licensed capacity

The amount of capacity on a storage system that a user is entitled to configure.

## License key

An alphanumeric code that activates a licensed function on a product.

## License key file

A file that contains one or more licensed keys.

## Lightweight Directory Access Protocol

Lightweight Directory Access Protocol (LDAP) is an open protocol that uses TCP/IP to provide access to directories that support an X.500 model and that does not incur the resource requirements of the more complex X.500 Directory Access Protocol (DAP). For example, LDAP can be used to locate people, organizations, and other resources in an Internet or intranet directory.

## Local and remote fabric interconnect

The local fabric interconnect and the remote fabric interconnect are the SAN components that are used to connect the local and remote fabrics. Depending on the distance between the two fabrics, they can be single-mode optical fibers that are driven by long wave (LW) gigabit interface converters (GBICs) or small form-factor pluggables (SFPs), or more sophisticated components, such as channel extenders or special SFP modules that are used to extend the distance between SAN components.

## Local fabric

The local fabric is composed of SAN components (switches, cables, and so on) that connect the components (nodes, hosts, and switches) of the local cluster together.

## Logical unit and logical unit number

The logical unit (LU) is defined by the SCSI standards as a logical unit number (LUN). LUN is an abbreviation for an entity that exhibits disk-like behavior, for example, a volume or an MDisk.

## Machine signature

A string of characters that identifies a system. A machine signature might be required to obtain a license key.

## Managed disk

A managed disk (MDisk) is a SCSI disk that is presented by a RAID controller and managed by the Storwize V7000. The MDisk is not visible to host systems on the SAN.

## Managed disk group (storage pool)

See “Storage pool (managed disk group)” on page 703.

## Metro Global Mirror

Metro Mirror Global is a cascaded solution where Metro Mirror synchronously copies data to the target site. This Metro Mirror target is the source volume for Global Mirror that asynchronously copies data to a third site. This solution has the potential to provide disaster recovery with no data loss at Global Mirror distances when the intermediate site does not participate in the disaster that occurs at the production site.

## Metro Mirror

Metro Mirror (MM) is a method of synchronous replication that maintains data consistency across multiple volumes within the system. Metro Mirror is generally used when the write latency that is caused by the distance between the source site and target site is acceptable to application performance.

## Mirrored volume

A mirrored volume is a single virtual volume that has two physical volume copies. The primary physical copy is known within the Storwize V7000 as `copy 0` and the secondary copy is known within the Storwize V7000 as `copy 1`.

## Node canister

A node canister is a hardware unit that includes the node hardware, fabric and service interfaces, and serial-attached SCSI (SAS) expansion ports.

## Node rescue

The process by which a node that has no valid software installed on its hard disk drive can copy software from another node connected to the same Fibre Channel fabric.

## NPIV

NPIV or N_Port ID Virtualization is a Fibre Channel feature whereby multiple Fibre Channel node port (N_Port) IDs can share a single physical N_Port.

## Object Storage

Object storage is a general term that refers to the entity in which an Cloud Object Storage (COS) organize, manage and store with units of storage, or just "objects".

## Oversubscription

Oversubscription refers to the ratio of the sum of the traffic on the initiator N-port connections to the traffic on the most heavily loaded ISLs, where more than one connection is used between these switches. Oversubscription assumes a symmetrical network, and a specific workload that is applied equally from all initiators and sent equally to all targets. A symmetrical network means that all the initiators are connected at the same level, and all the controllers are connected at the same level.

## Parent pool

Parent pools receive their capacity from MDisks. All MDisks in a pool are split into extents of the same size. Volumes are created from the extents that are available in the pool. You can add MDisks to a pool at any time either to increase the number of extents that are available for new volume copies or to expand existing volume copies. The system automatically balances volume extents between the MDisks to provide the best performance to the volumes.

## Partnership

In Metro Mirror or Global Mirror operations, the relationship between two clustered systems. In a clustered-system partnership, one system is defined as the local system and the other system as the remote system.

## Point-in-time copy

A point-in-time copy is the instantaneous copy that the FlashCopy service makes of the source volume. See also “FlashCopy service” on page 695.

## Preparing phase

Before you start the FlashCopy process, you must prepare a FlashCopy mapping. The preparing phase flushes a volume’s data from cache in preparation for the FlashCopy operation.

## Primary volume

In a stand-alone Metro Mirror or Global Mirror relationship, the target of write operations issued by the host application.

## Private fabric

Configure one SAN per fabric so that it is dedicated for node-to-node communication. This SAN is referred to as a private SAN.

## Public fabric

Configure one SAN per fabric so that it is dedicated for host attachment, storage system attachment, and remote copy operations. This SAN is referred to as a public SAN. You can configure the public SAN to allow Storwize V7000 node-to-node communication also. You can optionally use the `-localportfcmask` parameter of the **`chsystem`** command to constrain the node-to-node communication to use only the private SAN.

## Quorum disk

A disk that contains a reserved area that is used exclusively for system management. The quorum disk is accessed when it is necessary to determine which half of the clustered system continues to read and write data. Quorum disks can either be MDisks or drives.

## Quorum index

The quorum index is the pointer that indicates the order that is used to resolve a tie. Nodes attempt to lock the first quorum disk (index 0), followed by the next disk (index 1), and finally the last disk (index 2). The tie is broken by the node that locks them first.

## RACE engine

The RACE engine compresses data on volumes in real time with minimal effect on performance. See “Compression” on page 691 or “Real-time Compression” on page 700.

## Real capacity

Real capacity is the amount of storage that is allocated to a volume copy from a storage pool.

## Real-time Compression

Real-time Compression (RtC) is an IBM integrated software function for storage space efficiency. The RACE engine compresses data on volumes in real time with minimal effect on performance.

## Redundant Array of Independent Disks

Redundant Array of Independent Disks (RAID) refers to two or more physical disk drives that are combined in an array in a certain way, which incorporates a RAID level for failure protection or better performance. The most common RAID levels are 0, 1, 5, 6, and 10.

## RAID 0

RAID 0 is a data striping technique that is used across an array and no data protection is provided.

## RAID 1

RAID 1 is a mirroring technique that is used on a storage array in which two or more identical copies of data are maintained on separate mirrored disks.

## RAID 10

RAID 10 is a combination of a RAID 0 stripe that is mirrored (RAID 1). Therefore, two identical copies of striped data exist; no parity exists.

## RAID 5

RAID 5 is an array that has a data stripe, which includes a single logical parity drive. The parity check data is distributed across all the disks of the array.

## RAID 6

RAID 6 is a RAID level that has two logical parity drives per stripe, which are calculated with different algorithms. Therefore, this level can continue to process read and write requests to all of the array's virtual disks in the presence of two concurrent disk failures.

## Read intensive drives

The Read Intensive (RI) solid state drives (SSDs) that are available on Storwize V7000 Gen2+, Storwize V5000 Gen2, and IBM SAN Volume Controller 2145-DH8/24F are one Drive Write Per Day (DWPD) Read Intensive drives.

## Rebuild area

Reserved capacity that is distributed across all drives in a redundant array of drives. If a drive in the array fails, the lost array data is systematically restored into the reserved capacity, returning redundancy to the array. The duration of the restoration process is minimized because all drive members simultaneously participate in restoring the data. See also distributed RAID.

## Redundant storage area network

A redundant SAN is a SAN configuration in which there is no single point of failure (SPoF); therefore, data traffic continues no matter what component fails. Connectivity between the devices within the SAN is maintained (although possibly with degraded performance) when an error occurs. A redundant SAN design is normally achieved by splitting the SAN into two independent counterpart SANs (two SAN fabrics), so that if one path of the counterpart SAN is destroyed, the other counterpart SAN path keeps functioning.

## Relationship

In Metro Mirror or Global Mirror, a relationship is the association between a master volume and an auxiliary volume. These volumes also have the attributes of a primary or secondary volume.

## Reliability, availability, and serviceability

Reliability, availability, and serviceability (RAS) are a combination of design methodologies, system policies, and intrinsic capabilities that, when taken together, balance improved hardware availability with the costs that are required to achieve it.

Reliability is the degree to which the hardware remains free of faults. Availability is the ability of the system to continue operating despite predicted or experienced faults. Serviceability is how efficiently and nondisruptively broken hardware can be fixed.

## Remote fabric

The remote fabric is composed of SAN components (switches, cables, and so on) that connect the components (nodes, hosts, and switches) of the remote cluster together. Significant distances can exist between the components in the local cluster and those components in the remote cluster.

## SAN Volume Controller

The IBM System Storage SAN Volume Controller (SVC) is an appliance that is designed for attachment to various host computer systems. The SVC performs block-level virtualization of disk storage.

## Secondary volume

Pertinent to remote copy, the volume in a relationship that contains a copy of data written by the host application to the primary volume.

## Secure Sockets Layer (SSL) certificate

SSL (Secure Sockets Layer) is the standard security technology for establishing an encrypted link between a web server and a browser. This link ensures that all data passed between the web server and browsers remain private and to be able to create an SSL connection a web server requires an SSL Certificate.

## Security Key Lifecycle Manager (SKLM)

SKLM centralizes, simplifies and automates the encryption key management process to help minimize risk and reduce operational costs of encryption key management.

## Serial-attached SCSI

Serial-attached Small Computer System Interface (SAS) is a method that is used in accessing computer peripheral devices that employs a serial (one bit at a time) means of digital data transfer over thin cables. The method is specified in the American National Standard Institute standard called SAS. In the business enterprise, SAS is useful for access to mass storage devices, particularly external hard disk drives.

## Service Location Protocol

The Service Location Protocol (SLP) is an Internet service discovery protocol that enables computers and other devices to find services in a local area network (LAN) without prior configuration. It was defined in the request for change (RFC) 2608.

## Small Computer System Interface (SCSI)

Small Computer System Interface (SCSI) is an ANSI-standard electronic interface with which personal computers can communicate with peripheral hardware, such as disk drives, tape drives, CD-ROM drives, printers, and scanners, faster and more flexibly than with previous interfaces.

## Snapshot

A snapshot is an image backup type that consists of a point-in-time view of a volume.

## Solid-state disk

A solid-state disk (SSD) or Flash Disk is a disk that is made from solid-state memory and therefore has no moving parts. Most SSDs use NAND-based flash memory technology. It is defined to the Storwize V7000 as a disk tier generic_ssd.

## Space efficient

See “Thin provisioning” on page 703.

## Spare

An extra storage component, such as a drive or tape, that is predesignated for use as a replacement for a failed component.

## Spare goal

The optimal number of spares that are needed to protect the drives in the array from failures. The system logs a warning event when the number of spares that protect the array drops below this number.

## Space-efficient volume

For more information about a space-efficient volume, see “Thin-provisioned volume” on page 703.

## Stand-alone relationship

In FlashCopy, Metro Mirror, and Global Mirror, relationships that do not belong to a consistency group and that have a null consistency-group attribute.

## Statesave

Binary data collection that is used in problem determination.

## Storage area network

A storage area network (SAN) is a dedicated storage network that is tailored to a specific environment, which combines servers, systems, storage products, networking products, software, and services.

## Storage pool (managed disk group)

A storage pool is a collection of storage capacity, which is made up of managed disks (MDisks), that provides the pool of storage capacity for a specific set of volumes. A storage pool can contain more than one tier of disk, which is known as a multitier storage pool and a prerequisite of Easy Tier automatic data placement.

## Striped

Pertaining to a volume that is created from multiple managed disks (MDisks) that are in the storage pool. Extents are allocated on the MDisks in the order specified.

## Symmetric virtualization

Symmetric virtualization is a virtualization technique in which the physical storage, in the form of a Redundant Array of Independent Disks (RAID), is split into smaller chunks of storage known as extents. These extents are then concatenated, by using various policies, to make volumes. See also “Asymmetric virtualization” on page 690.

## Synchronous replication

Synchronous replication is a type of replication in which the application write operation is made to both the source volume and target volume before control is given back to the application. See also “Asynchronous replication” on page 690.

## Thin-provisioned volume

A thin-provisioned volume is a volume that allocates storage when data is written to it.

## Thin provisioning

Thin provisioning refers to the ability to define storage, usually a storage pool or volume, with a “logical” capacity size that is larger than the actual physical capacity that is assigned to that pool or volume. Therefore, a thin-provisioned volume is a volume with a virtual capacity that differs from its real capacity.

## Transparent Cloud Tiering (TCT)

Transparent Cloud Tiering is a separately installable feature of IBM Spectrum Scale that provides a native cloud storage tier.

## T10 DIF

T10 DIF is a *Data Integrity Field* (DIF) extension to SCSI to enable end-to-end protection of data from host application to physical media.

## Unique identifier

A unique identifier (UID) is an identifier that is assigned to storage-system logical units when they are created. It is used to identify the logical unit regardless of the logical unit number (LUN), the status of the logical unit, or whether alternate paths exist to the same device. Typically, a UID is used only once.

## Virtualization

In the storage industry, virtualization is a concept in which a pool of storage is created that contains several storage systems. Storage systems from various vendors can be used. The pool can be split into volumes that are visible to the host systems that use them. See also “Capacity licensing” on page 690.

## Virtualized storage

Virtualized storage is physical storage that has virtualization techniques applied to it by a virtualization engine.

## Virtual local area network

Virtual local area network (VLAN) tagging separates network traffic at the layer 2 level for Ethernet transport. The system supports VLAN configuration on both IPv4 and IPv6 connections.

## Virtual storage area network

A virtual storage area network (VSAN) is a fabric within the storage area network (SAN).

## Vital product data

Vital product data (VPD or VDP) is information that uniquely defines system, hardware, software, and microcode elements of a processing system.

## Volume

A volume is a Storwize V7000 logical device that appears to host systems that are attached to the SAN as a SCSI disk. Each volume is associated with exactly one I/O Group. A volume has a preferred node within the I/O Group.

## Volume copy

A volume copy is a physical copy of the data that is stored on a volume. Mirrored volumes have two copies. Non-mirrored volumes have one copy.

## Volume protection

To prevent active volumes or host mappings from inadvertent deletion, the system supports a global setting that prevents these objects from being deleted if the system detects that they have recent I/O activity. When you delete a volume, the system checks to verify whether it is part of a host mapping, FlashCopy mapping, or remote-copy relationship. In these cases, the system fails to delete the volume, unless the **`-force`** parameter is specified. Using the **`-force`** parameter can lead to unintentional deletions of volumes that are still active. Active means that the system detected recent I/O activity to the volume from any host.

### Write-through mode

Write-through mode is a process in which data is written to a storage device at the same time that the data is cached.

# Related publications

The publications listed in this section are considered particularly suitable for a more detailed discussion of the topics covered in this book.

## IBM Redbooks

The following IBM Redbooks publications provide additional information about the topic in this document (note that some publications referenced in this list might be available in softcopy only):

- *IBM b-type Gen 5 16 Gbps Switches and Network Advisor*, SG24-8186
- *IBM FlashSystem 820 Running in an IBM StorwizeV7000 Environment*, TIPS1101
- *IBM FlashSystem 840 Product Guide*, TIPS1079
- *IBM FlashSystem in IBM PureFlex System Environments*, TIPS1042
- *IBM FlashSystem V840*, TIPS1158
- *IBM Midrange System Storage Implementation and Best Practices Guide*, SG24-6363
- *IBM SAN Volume Controller and IBM FlashSystem 820: Best Practices and Performance Capabilities*, REDP-5027
- *IBM System Storage b-type Multiprotocol Routing: An Introduction and Implementation*, SG24-7544
- *IBM Tivoli Storage Area Network Manager: A Practical Introduction*, SG24-6848
- *IBM Tivoli Storage Productivity Center V5.2 Release Guide*, SG24-8204
- *Implementing an IBM b-type SAN with 8 Gbps Directors and Switches*, SG24-6116
- *Implementing FlashSystem 840 with SAN Volume Controller*, TIPS1137
- *Implementing IBM FlashSystem 840*, SG24-8189
- *Implementing the IBM SAN Volume Controller and FlashSystem 820*, SG24-8172
- *Implementing the IBM Storwize V7000 V7.4*, SG24-7938
- *Implementing the IBM System Storage SAN Volume Controller V7.4*, SG24-7933
- *Introduction to Storage Area Networks*, SG24-5470
- *Tivoli Storage Productivity Center for Replication for Open Systems*, SG24-8149

You can search for, view, download, or order these documents and other Redbooks, Redpapers, Web Docs, draft and additional materials, at the following website:

**ibm.com**/redbooks

The following is a list of useful Redbooks domains related to this book:

IBM Storage Networking Redbooks:

http://www.redbooks.ibm.com/Redbooks.nsf/domains/san?Open

IBM Flash Storage Redbooks

http://www.redbooks.ibm.com/Redbooks.nsf/domains/flash?Open

IBM Software Defined Storage Redbooks

http://www.redbooks.ibm.com/Redbooks.nsf/domains/sds?Open

IBM Disk Storage Redbooks

http://www.redbooks.ibm.com/Redbooks.nsf/domains/disk?Open

IBM Storage Solutions Redbooks

http://www.redbooks.ibm.com/Redbooks.nsf/domains/storagesolutions?Open

IBM Tape Storage Redbooks

http://www.redbooks.ibm.com/Redbooks.nsf/domains/tape?Open

## Other resources

These publications are also relevant as further information sources:

- *IBM System Storage Master Console: Installation and User’s Guide*, GC30-4090
- *IBM System Storage Open Software Family SAN Volume Controller: CIM Agent Developers Reference*, SC26-7545
- *IBM System Storage Open Software Family SAN Volume Controller: Command-Line Interface User's Guide*, SC26-7544
- *IBM System Storage Open Software Family SAN Volume Controller: Configuration Guide*, SC26-7543
- *IBM System Storage Open Software Family SAN Volume Controller: Host Attachment Guide*, SC26-7563
- *IBM System Storage Open Software Family SAN Volume Controller: Installation Guide*, SC26-7541
- *IBM System Storage Open Software Family SAN Volume Controller: Planning Guide*, GA22-1052
- *IBM System Storage Open Software Family SAN Volume Controller: Service Guide*, SC26-7542
- *IBM System Storage SAN Volume Controller - Software Installation and Configuration Guide,* SC23-6628
- *IBM System Storage SAN Volume Controller V6.2.0 - Software Installation and Configuration Guide,* GC27-2286
- *IBM System Storage SAN Volume Controller 6.2.0 Configuration Limits and Restrictions*, S1003799
- *IBM TotalStorage Multipath Subsystem Device Driver User’s Guide*, SC30-4096
- *IBM XIV and SVC Best Practices Implementation Guide*

  http://ibm.co/1bk64gW
- *Considerations and Comparisons between IBM SDD for Linux and DM-MPIO*

  http://ibm.co/1CD1gxG

# Referenced websites

These websites are also relevant as further information sources:

- IBM Storage home page

  http://www.ibm.com/systems/storage
- SAN Volume Controller supported platform

  http://ibm.co/1FNjddm
- SAN Volume Controller IBM Knowledge Center

  http://www.ibm.com/support/knowledgecenter/STPVGU/welcome
- Cygwin Linux-like environment for Windows

  http://www.cygwin.com
- Open source site for SSH for Windows and Mac

  http://www.openssh.com/windows.html
- Sysinternals home page

  http://www.sysinternals.com
- Download site for Windows SSH freeware

  http://www.chiark.greenend.org.uk/~sgtatham/putty

# Help from IBM

IBM Support and downloads

**ibm.com**/support

IBM Global Services

**ibm.com**/services

To determine the spine width of a book, you divide the paper PPI into the number of pages in the book. An example is a 250 page book using Plainfield opaque 50# smooth which has a PPI of 526. Divided 250 by 526 which equals a spine width of .4752". In this case, you would use the .5" spine. Now select the Spine width for the book and hide the others: **Special>Conditional Text>Show/Hide>SpineSize(-->Hide:)>Set** . Move the changed Conditional text settings to all files in your book by opening the book file with the spine.fm still open and **File>Import>Formats** the Conditional Text Settings (ONLY!) to the book files.

Redbooks

# Implementing the IBM Storwize V7000 and IBM Spectrum

SG24-7938-05

ISBN DocISBN

(1.5" spine)
1.5"<-> 1.998"
789 <->1051 pages

Redbooks

# Implementing the IBM Storwize V7000 and IBM Spectrum

SG24-7938-05

ISBN DocISBN

(1.0" spine)
0.875"<->1.498"
460 <-> 788 pages

Redbooks

**Implementing the IBM Storwize V7000 and IBM Spectrum**

SG24-7938-05

ISBN DocISBN

(0.5" spine)
0.475"<->0.873"
250 <-> 459 pages

Redbooks

**Implementing the IBM Storwize V7000 and IBM Spectrum Virtualize V7.8**

(0.2"spine)
0.17"<->0.473"
90<->249 pages

(0.1"spine)
0.1"<->0.169"
53<->89 pages

To determine the spine width of a book, you divide the paper PPI into the number of pages in the book. An example is a 250 page book using Plainfield opaque 50# smooth which has a PPI of 526. Divided 250 by 526 which equals a spine width of .4752". In this case, you would use the .5" spine. Now select the Spine width for the book and hide the others: **Special>Conditional Text>Show/Hide>SpineSize(-->Hide:)>Set** . Move the changed Conditional text settings to all files in your book by opening the book file with the spine.fm still open and **File>Import>Formats** the Conditional Text Settings (ONLY!) to the book files.

Redbooks

# Implementing the IBM Storwize V7000 and IBM

SG24-7938-05

ISBN DocISBN

IBM

(2.5" spine)
2.5"<->nnn.n"
1315<-> nnnn pages

Redbooks

# Implementing the IBM Storwize V7000 and IBM Spectrum Virtualize V7.8

SG24-7938-05

ISBN DocISBN

IBM

(2.0" spine)
2.0" <-> 2.498"
1052 <-> 1314 pages

SG24-7938-05

ISBN DocISBN

Printed in U.S.A.