IBM DS8880 Architecture and Implementation (Release 8.2.1)

Kerstin Blum

Sherry Brunson

Andre Coelho

Jeffery Cook

Bert Dufrasne

Peter Kimmel

Bjoern Wesselbaum

Storage

IBM

International Technical Support Organization

IBM DS8880 Architecture and Implementation (Release 8.2.1)

January 2017

Third Edition (January 2017)

Contents

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®	IBM z Systems®	Redpaper™
BigFix®	IBM z13®	Redbooks (logo) ®
DB2®	InfoSphere®	RMF™
DS8000®	Parallel Sysplex®	Storwize®
Easy Tier®	POWER®	System i®
Enterprise Storage Server®	Power Architecture®	System Storage®
Enterprise Workload Manager™	POWER Hypervisor™	Tivoli®
FICON®	Power Systems™	WebSphere®
FlashCopy®	POWER6+™	XIV®
GDPS®	POWER7®	z Systems®
Geographically Dispersed Parallel	POWER7+™	z/OS®
Sysplex™	POWER8®	z/VM®
HyperSwap®	PowerPC®	z13™
IBM®	ProtecTIER®	z13s™
IBM Spectrum™	RACF®	zEnterprise®
IBM Spectrum Control™	Redbooks®	

The following terms are trademarks of other companies:

Resilient, and Resilient SystemsÆ are trademarks or registered trademarks of Resilient Systems, an IBM Company.

Intel, 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

This IBM® Redbooks® publication describes the concepts, architecture, and implementation of the IBM DS8880 family. The book provides reference information to assist readers who need to plan for, install, and configure the DS8880 systems.

The IBM DS8000® family is a high-performance, high-capacity, highly secure, and resilient series of disk storage systems. The DS8880 family is the latest and most advanced of the DS8000 offerings to date. The high availability, multiplatform support, including IBM z Systems®, and simplified management tools help provide a cost-effective path to an on-demand world.

The IBM DS8880 family now offers business-critical, all-flash, and hybrid data systems that span a wide range of price points:

► DS8884 — Business Class
► DS8886 — Enterprise Class
► DS8888 — Analytics Class

The DS8884 and DS8886 are available as either hybrid models, or can be configured as *all-flash*. Each model represents the most recent in this series of high-performance, high-capacity, flexible, and resilient storage systems. These systems are intended to address the needs of the most demanding clients.

Two powerful IBM POWER8® processor-based servers manage the cache to streamline disk I/O, maximizing performance and throughput. These capabilities are further enhanced with the availability of the *second generation* of high-performance flash enclosures (HPFEs Gen-2).

Like its predecessors, the DS8880 supports advanced disaster recovery (DR) solutions, business continuity solutions, and thin provisioning. All disk drives in the DS8880 storage system include the Full Disk Encryption (FDE) feature. The DS8880 can automatically optimize the use of each storage tier, particularly flash drives and flash cards, through the IBM Easy Tier® feature. The DS8880 also includes the Copy Services Manager code and allows for easier integration in a Lightweight Directory Access Protocol (LDAP) infrastructure.

Authors

This book was produced by a team of specialists from around the world:

Kerstin Blum is a Senior Product Service Specialist at the ESCC (European Storage Competence Center) in Kelsterbach, Germany. She joined IBM in 1999 and since then has worked as a Product Field Engineer (last level support) for the DS8000 product line. She was assigned for 18 months as the EMEA field support interface with the ESS/DS8000 development and test teams in Tucson, AZ. She holds a degree in Electrical Engineering, with a specialization in Communications Engineering, from the University of Applied Sciences of Wiesbaden (Germany).

Sherry Brunson joined IBM in March of 1985 and worked as large system IBM service representative before becoming a Top Gun in 1990. Sherry is a Top Gun in the Eastern US for all storage products, power systems, and IBM z Systems®. She has supported and implemented DS8000 and Scaled out Network Appliance storage products globally as well as developing and teaching educational classes. She also has taught z System classes in the United States.

Andre Coelho is a Consulting IT Storage Specialist in Rio de Janeiro, Brazil. He works for Internal Accounts (US and EMEA). He has over 30 years of experience in IBM z/OS® and predecessor operating systems. His areas of expertise include DFSMS, DFSMShsm, DFSMSrmm, DS8000, and Hydra. He has contributed to several IBM Redbooks publications.

Jeffery Cook joined IBM in 1979 as a large systems IBM service representative. Before he joined the DS8000 Product Engineering team in 2011, he worked as a storage systems regional Top Gun for the western United States. He has developed and taught DS8000 and IBM XIV® Storage System courses in the United States and South America.

Bert Dufrasne is an IBM Certified IT Specialist and Project Leader for IBM System Storage® disk products at the ITSO, San Jose Center. He has worked at IBM in various IT areas. He has written many IBM publications, and has developed and taught technical workshops. Before Bert joined the ITSO, he worked for IBM Global Services as an Application Architect. He holds a Master's degree in Electrical Engineering.

Peter Kimmel is an IT Specialist and ATS team lead of the Enterprise Storage Solutions team at the EMEA Storage Competence Center (ESCC) in Kelsterbach, Germany. He joined IBM Storage in 1999 and since then worked with all the various IBM Enterprise Storage Server® and System Storage DS8000 generations, with a focus on architecture and performance. He was involved in the Early Shipment Programs (ESPs) of these early installs, and co-authored several DS8000 IBM publications. Peter holds a Diploma (MSc) degree in Physics from the University of Kaiserslautern.

Bjoern Wesselbaum has been a Senior Product Services Specialist for IBM Germany since 1998. He has worked as a Product Field Engineer for IBM Storage Systems within the EMEA Advanced Technical Support and Product Field Engineering support team at the Storage Competence Center in Kelsterbach. Since 2007, Bjoern has also taught IBM DS8000 Implementation and DS8000 Copy Services classes for IBM Learning Services. His areas of expertise include IBM AIX® and other UNIX-like operating systems, storage attachment, Copy Services Solutions for open systems, and DS8000 management and operations, including Remote Support Connectivity Solutions for the DS8000.

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

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 SG24-8323-02
for *IBM DS8880 Architecture and Implementation (Release 8.2.1)*
as created or updated on March 10, 2017.

January 2017, Third Edition

This revision includes the following new and changed information.

New information
► Introduces DS8884, DS8886, and DS8888 with High-Performance Flash Enclosure Gen-2

► RAID 6 default for all storage types

► Thin provisioning enhancements

Changed information
► Various GUI changes in support of new or enhanced functions
► Syslog support for auditlog
► Virtualization information to reflect use of small extents pools and volumes

Part 1

Concepts and architecture

This part of the book provides an overview of the IBM DS8880 concepts and architecture.

This part contains the following chapters:

Introduction to the IBM DS8880

This chapter introduces the features and functions of the IBM DS8880. More information about functions and features is provided in subsequent chapters.

This chapter covers the following topics:

► Introduction to the IBM DS8880
► DS8880 controller options and frames
► DS8880 architecture and functions overview
► Performance features

The previous DS8870 model is described in *IBM DS8870 Architecture and Implementation (Release 7.5)*, SG24-8085.

1.1 Introduction to the IBM DS8880

IBM has a wide range of product offerings that are based on open standards. These offerings share a common set of tools, interfaces, and innovative features. The IBM DS8000 family is a high-performance, high capacity, highly secure, and resilient series of disk storage systems. The DS8880 family is the latest and most advanced of the DS8000 offerings to date. The high availability, multiplatform support, including IBM z Systems, and simplified management tools help provide a cost-effective path to an on-demand world.

The current IBM DS8880 family includes the following three high-performance models:

► DS8884 – Business Class
► DS8886 – Enterprise Class
► DS8888 – Analytics Class

These additions to the family expand on features that clients expect from a high-end storage system:

► High performance
► High availability
► Cost efficiency
► Energy efficiency
► Scalable
► Business continuity and data protection functions

The DS8880 (Figure 1-1) supports the most demanding business applications with its exceptional all-around performance and data throughput.

Figure 1-1   DS8884 40U Rack and the expanded DS8886/DS8888 46U Rack

A noticeable change with the introduction of the DS8880 is the reduction of the footprint to a 19-inch rack. The DS8880 architecture is server-based. Two powerful POWER8 processor-based servers manage the cache to streamline disk I/Os, maximizing performance and throughput. These capabilities are further enhanced with the availability of high-performance flash enclosures (HPFEs). An HPFE Gen-2 is now available. It is described in *DS8880 High Performance Flash Enclosure Gen2*, REDP-5422.

Combined with world-class business resiliency and encryption features, the DS8880 provides a unique combination of high availability, performance, and security. The DS8880 is equipped with encryption-capable disk drives. Encryption-capable solid-state flash drives (SSDs) and high-performance flash cards are also available. Figure 1-2 shows the DS8888 All-Flash and the 3-phase direct current uninterruptible power supply (DC-UPS).

Figure 1-2   DS8888 all-flash system

The DS8880 includes a power architecture that is based on a DC-UPS. The DC-UPS converts incoming AC line voltage to rectified AC and contains an integrated battery subsystem. The DC-UPS allows the DS8880 to achieve the highest energy efficiency in the DS8000 series. The DS8880 is designed to comply with the emerging ENERGY STAR specifications. The DS8888 All-Flash is available in 3-phase DC-UPS. The DS8886 is available in 3-phase or single-phase DC-UPS, and the DS8884 is available in single-phase DC-UPS.

The DS8880 is delivered in a reduced footprint, 19-inch rack.

The I/O enclosure that was used for the DS8870 is replaced by an I/O enclosure that is attached to the POWER8 servers with Peripheral Component Interconnect Express (PCIe) Generation 3 cables. The I/O enclosure has six PCIe adapter slots and two additional direct-attached, high-performance flash enclosure ports.

High-density storage enclosures offer a considerable reduction in the footprint and energy consumption. The rack also has an integrated keyboard and display pocket in the side wall that can be accessed from the front of the rack. A small form factor Hardware Management Console (HMC) is used on the DS8880. One or two HMCs are supported in the base rack. A unique feature of the new rack is the ability to expand from the height of a 40U configuration to a 46U configuration concurrently in the DS8886.

Figure 1-3 shows the differences in configuration that are available between the two sizes of the base frame with the single-phase DC-UPS. The expansion enables clients to add more capacity to their storage systems in the same footprint.
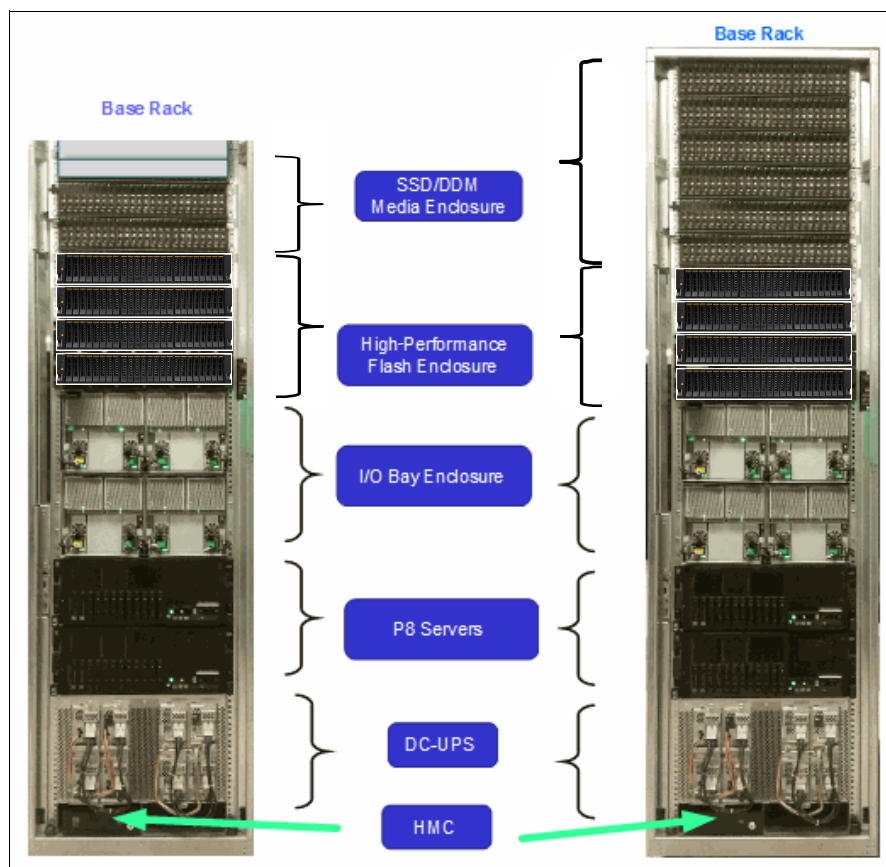
Figure 1-3   DS8886 40U (left) and DS8886 46U (right) fully configured base frames (985 model)

1.1.1 At-a-glance features of the DS8880

The DS8880 offers the following features (more details can be found in subsequent chapters of this book):

► The DS8880 is available in a variety of POWER8 processor options, from DS8888 48-core to a DS8884 single 6-core. Each system covers a wide range of performance and cost options.

► Memory configurations are available ranging from 64 GB up to 2 TB. The DS8880 family introduces a new POWER8 central processor complex (CPC) unit for each member. The DS8888 uses a model E850 processor complex that performs at greater than twice the performance of a DS8870. The DS8884 by using a model S822 CPC also performs better than the DS8870.

► The DS8880 is available with HPFEs in addition to Fibre Channel-attached flash drives (which are also known as SSDs). HPFEs can be installed in the base frame and the first expansion frame. One High-Performance Flash Enclosure (Gen-2) pair can hold 16, 32 or 48 flash cards, with capacities of 400 GB, 800 GB, 1.6 TB or 3.2 TB.

► The POWER8 simultaneous multithreading (SMT) technology allows analytic data processing clients to achieve up to 2.5 million I/O per second (IOPS) in Database for Open environments (70% read/30% write, 4 KB IOs, 50% read cache hit) with the 384 Flash cards, 2 TB cache, and 48-cores in the DS8888 all-flash.

► The z-Synergy Services include IBM z/OS licensed features that are supported on the storage system. The licensed features include IBM FICON® attachment for z Systems, PAV, HyperPAV, SuperPAV, IBM z Systems High-Performance FICON (zHPF), and z/OS Distributed Data Backup, as well as the option for thinly provisioned CKD volumes.

► The DS8880 offers enhanced connectivity with four-port 16 Gbps and four or eight-port 8 Gbps Fibre Channel/IBM FICON host adapters (HAs). Each host adapter port can be configured independently, based on HA speed:

  – 16 Gbps can use Fibre Channel Protocol (FCP) or FICON protocol.

  – 8 Gbps can use Fibre Channel Arbitrated Loop (FC-AL), FCP, or FICON protocol.

  Ports that are configured for FCP can be used for mirroring.

► High-Performance FICON for z Systems (zHPF) is an enhancement to IBM FICON architecture to offload I/O management processing from the z Systems channel subsystem to the DS8880 HA and controller. zHPF is an optional feature of the z Systems server and of the DS8880. Recent enhancements to zHPF include Extended Distance Facility zHPF List Pre-fetch support for IBM DB2® and utility operations, and zHPF support for sequential access methods. All DB2 I/O is now zHPF-capable, and now supports the DB2 castout accelerator function that allows the DS8000 to treat a castout as a single chain of IOs.

  The 16 Gbps host adapters on the DS8880 and IBM z13® and later z Systems server channels increase the write throughput by 50% or better. This improvement allows zHPF Extended Distance II to support heavy write I/Os over an extended distance of up to 100 km (62 miles).

► Peripheral Component Interconnect Express (PCIe Generation 3) I/O enclosures complement the improved I/O bandwidth that is provided by POWER8. The DS8880 includes new PCIe Gen3-based I/O enclosures. Each I/O enclosure connects to both internal servers over a pair of x8 PCIe Gen3 cables, each of which provides 8 GBps connectivity. An additional pair of PCIe connectors on the I/O enclosure provides connections to HPFEs.

  Each pair of I/O enclosures supports up to two pairs of High-Performance Flash Enclosures Gen2. The storage virtualization that is offered by the DS8880 allows

organizations to allocate system resources more effectively and better control application quality of service (QoS).

► The Adaptive Multi-stream Prefetching (AMP) caching algorithm can dramatically improve sequential performance, reducing times for backup, processing for business intelligence and streaming media. Sequential Adaptive Replacement Cache is a caching algorithm that allows you to run different workloads, such as sequential and random workloads, without negatively affecting each other. For example, sequential workload does not fill up the cache and does not affect cache hits for random workloads. Intelligent Write Caching (IWC) improves the Cache Algorithm for random writes.

► The DS8880 provides a graphical management interface to configure the DS8880, query status information, and monitor the performance of the DS8880.

► Easy Tier is a no-charge feature that enables automatic dynamic data relocation capabilities. Easy Tier optimizes the usage of each tier, especially the use of flash storage. No manual tuning is required. The auto-balancing algorithms also provide benefits when homogeneous storage pools are used to eliminate hot spots on disk arrays.

► Storage Tier Advisor Tool (STAT) is used with the Easy Tier feature to help clients understand their current storage system workloads. The tool also provides guidance about how much of their existing data is better suited for the various drive types (spinning drives or flash).

► I/O Priority Manager is a feature that provides application-level QoS for workloads that share a storage pool. This feature provides a way to manage QoS for I/O operations that are associated with critical workloads and gives them priority over other I/O operations that are associated with non-critical workloads. For IBM z/OS, the I/O Priority Manager allows increased interaction with the host side.

► Large volume support that allows DS8880 to support logical unit number (LUN) sizes up to 16 TB. This configuration simplifies storage management tasks. In a z/OS environment, extended address volumes (EAVs) with sizes up to 1 TB are supported.

► Support for IBM System i® variable LUN adds flexibility for volume sizes. It can increase capacity usage for IBM i environments. Before, clients were limited to fixed LUN sizes of 35 GB, 70 GB, 141 GB, and 282 GB.

► The DS8880 has an Active Volume Protection feature that prevents the deletion of volumes that are still in use.

► The American National Standards Institute (ANSI) T10 Data Integrity Field (DIF) standard is supported on IBM z Systems for Small Computer System Interface (SCSI) end-to-end data protection on fixed-block architecture (FB) LUN volumes. This support applies to the IBM DS8000. IBM z Systems support applies to FCP channels only.

► The Dynamic Volume Expansion simplifies management by enabling easier, online volume expansion (for open systems and z Systems) to support application data growth, and to support data center migration and consolidation to larger volumes to ease addressing constraints.

► Thin provisioning allows the creation of over-provisioned devices for more efficient usage of the storage capacity for open systems. Extent space-efficient volumes with small extents are now supported. For more information, see *IBM DS8880 Thin Provisioning*, REDP-5343.

► Quick Initialization provides fast volume initialization for open system LUNs and count key data (CKD) volumes. Quick Initialization allows the creation of devices, making them available when the command completes.

► Full disk encryption (FDE) can protect business-sensitive data by providing disk-based hardware encryption that is combined with external key management software (IBM Security Key Lifecycle Manager or Gemalto SafeNet KeySecure). FDE is available for all

disks and drives, including flash cards and flash drives (SSD). Because encryption is performed by the disk drive, it is transparent to host systems and it can be used in any environment, including z/OS.

► The DS8880 enables clients to comply with the Special Publication (SP) number 800-131a. This is a National Institute of Standards and Technology (NIST) directive that provides guidance for protecting sensitive data by using cryptographic algorithms that have key strengths of 112 bits.

► The following specific features of encryption key management help address Payment Card Industry Data Security Standard (PCI DSS) requirements:

– The storage system supports data encryption with external key servers. In addition to the IBM Security Key Lifecycle Manager key server, DS8880 also supports the Gemalto Safenet KeySecure key servers for customers who need to use the OASIS Key Management Interoperability Protocol (KMIP).

– Encryption deadlock recovery key option, which allows the user to restore access to a DS8880 when the encryption key for the storage is unavailable because of an encryption deadlock scenario.

– Dual platform key server support, which is important if key servers on z/OS share keys with key servers on open systems. The DS8880 requires an isolated key server in encryption configurations. Dual platform key server support allows two server platforms to host the key manager with either platform operating in either clear key or secure key mode.

– Recovery key Enabling/Disabling and Rekey data key option for the FDE feature, which can help clients satisfy PCI security standards.

► Lightweight Directory Access Protocol (LDAP) authentication support, which allows single sign-on (SSO), can simplify user management by allowing the DS8000 to rely on a centralized LDAP directory rather than a local user repository. LDAP can be enabled through the Copy Services Manager (CSM), which is packaged with the HMC code.

► For data protection and availability needs, the DS8880 provides a rich set of point-in-time and remote mirror and copy functions. These functions provide storage mirroring and copying over large distances for disaster recovery or availability purposes. These functions are also fully interoperable with previous models of the DS8000 family.

► New since DS8880 Release 8.1, the CSM software (previously known as IBM Tivoli® Productivity Center for Replication) is preinstalled on the DS8880 HMC.

► Resource groups offer a policy-based resource scope-limiting function that enables the secure use of Copy Services functions by multiple users on a DS8000 series storage system. Resource groups are used to define an aggregation of resources and policies for configuration and management of those resources. The scope of the aggregated resources can be tailored to meet each hosted client's Copy Services requirements for any operating system platform that is supported by the DS8000 series.

► The DS8880 offers enhanced z Systems Synergy items:

– Forward Error Correction (FEC)

FEC is a protocol that captures errors that are generated during data transmission.

Both the z Systems z13 (and later) and the DS8880 extend the use of FEC to complete end-to-end coverage for 16 Gbps links and preserve data integrity with more redundancy.

– Fibre Channel (FC) Read Diagnostic Parameters (RDP)

  FC RDP improves the end-to-end link fault isolation for 16 Gbps links on the z Systems z13 (and later) and the DS8000. RDP data provides the optical signal strength, error counters, and other critical information that is crucial to determine the quality of the link.

– FICON Dynamic Routing (FIDR)

  FIDR is a z Systems z13 (and later) and DS8000 feature that supports the use of dynamic routing policies in the fabric switch to balance load across inter-switch links (ISLs) on a per I/O basis.

– Fabric I/O Priority

  Fabric I/O Priority is an end-to-end synergy feature between the z/OS Workload Manager, Brocade SAN Fabric, and the storage system to manage QoS on a single I/O level.

> **Note:** For more information about the z Systems synergy features, see *IBM DS8870 and IBM z Systems Synergy*, REDP-5186.

► The DS8000 series is certified as meeting the requirements of the IPv6 Ready Logo program. This certification indicates its implementation of IPv6 mandatory core protocols and the ability to interoperate with other IPv6 implementations. The IBM DS8000 can be configured in native IPv6 environments. The logo program provides conformance and interoperability test specifications that are based on open standards to support IPv6 deployment globally. Furthermore, the NIST tested IPv6 with the DS8000, granting it support for the USGv6 profile and testing program.

► The IBM DS8880 supports VMware vStorage application programming interfaces (APIs) for Array Integration (VAAI). VAAI enables certain storage tasks to be offloaded from the server hardware to the storage array. Support is included for the Atomic Test and Set (ATS) primitive, the Cloning Blocks primitive, and the Zeroing Blocks primitive.

► The IBM DS8880 supports IBM Storage Management Console for VMware vCenter. IBM Storage Management Console for VMware vCenter is a software plug-in that integrates into the VMware vCenter server platform. It enables VMware administrators to independently and centrally manage their storage resources on IBM storage systems. In addition to being a virtualization-related solution, the IBM Storage Management Console is a powerful management solution for VMware administrators who want to control storage resources primarily from the VMware vSphere Client GUI.

1.2 DS8880 controller options and frames

The IBM DS8880 family consists of a series of distinct models:

► DS8888 Model 988 (Base/three-phase power) and Model 88E (Expansion)
  This is an all-flash model: DS8888F.

► DS8886 Model 986 (Base/three-phase power) and Model 86E (Expansion)
  This model is also available in an all-flash version: DS8886F.

► DS8886 Model 985 (Base/single-phase power) and Model 85E (Expansion)
  This model is also available in an all-flash version: DS8886F.

► DS8884 Model 984 (Base/single-phase power) and Model 84E (Expansion)
  This model is also available in an all-flash version: DS8884F.

These systems are part of the 283*x* (hybrid) and 533*x* (all-flash) machine type families.

The DS8880 includes the following features:

► IBM POWER8 processor technology

The DS8880 currently features the IBM POWER8 processor-based server technology for high performance. Compared to the IBM POWER7+™ processors that were used in models of the DS8870, IBM POWER8 can deliver at least 2x improvement in IOPS in transaction-processing workload environments.

► Nondisruptive upgrade path

A nondisruptive upgrade path for the DS8880 allows processor, cache, HAs, storage upgrades, and both model 98E and 98B expansion frame upgrades to be installed concurrently without disrupting applications.

► The air-flow system allows optimal horizontal cool down of the storage system. The DS8880 is designed for hot and cold aisle data center design, drawing air for cooling from the front of the system and exhausting hot air at the rear. For more information, see 2.6, "Power and cooling" on page 56.

► The DS8880 offers three system class options:

– DS8888 based on POWER8 E850 architecture

Up to 384 Flash Cards, up to 48 cores, 3.02 GHz, and up to 2 TB of system memory.

– DS8886 based on POWER8 S824 architecture

Up to 1,536 drives (1,440 for model 986), single 8-core or 16/24 dual-core processor configuration at 3.52 or 3.89 GHz, and up to 2 TB of system memory.

– DS8884 based on POWER8 S822 architecture

Up to 768 drives, single socket 6-core processor configuration at 3.89 GHz, and up to 256 GB of system memory.

For more information, see 2.3.1, "IBM POWER8-based servers" on page 35.

> **Note:** The DS8884 Class configuration cannot be converted into a DS8886 or DS8888 Class configuration.

► High-Performance Flash Enclosure

The HPFE Gen2 is a RAID enclosure that can support flash cards with capacities between 400 GB and 3.2 TB (2.5-inch form factor), in a 2U rack space. The enclosures need to be installed in pairs. You can install up to 8 HPFE Gen2 pairs in the DS8888 all-flash system, up to four HPFE pairs for each DS8886 in a hybrid system, and up to two HPFE Gen2 pairs in a DS8884 hybrid system.

► Standard drive enclosures

The DS8880 configurations provide standard drive enclosure support for 24 small form factor (SFF) 2.5-inch (64 mm) drives in 2U of rack space. This option helps improve the storage density for disk drive modules (DDMs) as compared to previous enclosures.

The DS8886 configuration can support 1,536 drives, plus 192 flash cards, in a small, high-density footprint (base frame and up to four expansion frames) that helps to preserve valuable data center floor space with the 19-inch rack 40U extended to 46U.

The DS8884 configuration can support 768 drives, plus 96 flash cards in a small high-density footprint (base frame and up to two expansion frames) that helps to preserve valuable data center floor space with the 19-inch rack 40U.

As an alternative, the DS8886 and the DS8884 also support large form factor (LFF) drive enclosures with twelve 3.5-inch (89 mm) disk drives in 2U of rack space. LFF drive maximums are half of the maximums of the SFF enclosures. SFF and LFF enclosures can be intermixed within the same frame and within a DA pair, although they cannot be intermixed within a drive enclosure pair.

1.3 DS8880 architecture and functions overview

The DS8880 offers continuity with the fundamental architecture of its predecessors, the DS8800 and DS8870 models. This architecture ensures that the DS8880 can use a stable and well-proven operating environment that offers optimal availability. The hardware is optimized to provide higher performance, connectivity, and reliability.

1.3.1 Overall architecture and components

For more information about the available configurations for the DS8880, see Chapter 2, "IBM DS8880 hardware components and architecture" on page 23.

IBM POWER8 processor technology

The POWER8 processor is manufactured by using the IBM 22 nm Silicon-On-Insulator (SOI) technology. Each chip is 649 mm^2 and contains 4.2 billion transistors. The DS8880 now uses the PCIe Gen 3 I/O controllers and an interconnection system that connects all components within the chip. The POWER8 processors that are used in the DS8880 can run at 3.891 GHz or 3.525 GHz, depending on your configuration.

POWER8 processor advancements in multi-core and multithreading are remarkable. A significant performance opportunity comes from parallelizing workloads to enable the full potential of the microprocessor, and the large memory bandwidth. SMT allows a single physical processor core to simultaneously dispatch instructions from more than one hardware thread context. Although x86-based processors are limited to just two threads per core, IBM POWER® technology also enables a processor to run in SMT4 or SMT8 modes. These multithreading capabilities improve the I/O throughput of the DS8880 storage server.

The DS8880 family offers several configurations of CPCs:

► The DS8888 offers two processor configurations:

– 24-core with 1 TB memory
– 48-core with 2 TB memory

► The DS8886 offers three processor configurations:

– 8-core with 128 and 256 GB memory
– 16-core with 256 and 512 GB memory
– 24-core with 1024 and 2048 GB memory

► The DS8884 offers a single processor configuration of a 6-core with 64, 128, and 256 GB memory.

A processor complex is also referred to as a *storage server* or CPC. For more information, see Chapter 3, "Reliability, availability, and serviceability on the IBM DS8880" on page 65.

Internal PCIe-based fabric

The new I/O enclosure pairs are PCIe Gen3-capable. The new I/O enclosure supports six PCIe adapter slots and two additional direct-attached, high-performance flash enclosure ports. The I/O enclosures are attached to the POWER8 servers by two x8 PCIe Gen3 cables. The internal transfer rate to each I/O bay is four times faster compared to the earlier model.

For more information, see Chapter 2, "IBM DS8880 hardware components and architecture" on page 23.

High-performance flash enclosure

The HPFEs connect to the I/O enclosures over a PCIe fabric, which increases bandwidth and transaction-processing capability. The original HPFE that was used in the earlier DS8880 models (980, 981, 982) was a 1U enclosure that contained integrated dual Flash RAID adapters, with up to 30 flash cards 1.8-inch (400/800 GB). The HPFE Gen2, which is based on 2.5-inch flash cards (up to 3.2 TB) and installed in the models 984/985/986/988, comes in pairs of 2U enclosures, with each such pair holding up to 48 flash cards. HPFEs and flash cards provide up to a 4x throughput improvement compared to ordinary flash drives.

When comparing both HPFE generations, the Gen2 integrated flash RAID controller delivers significantly increased performance and capacity figures. Each flash adapter pair, and thus each HPFE pair, delivers up to 500K IOPS, and up to 14 GBps and 9.3 GBps for read and write bandwidth respectively.

For more information, see Chapter 2, "IBM DS8880 hardware components and architecture" on page 23.

Device adapters

The DS8880 Standard Drive Enclosures connect to the processor complexes through 8 Gbps four-port Fibre Channel device adapters (DAs). They are optimized for flash drive (SSD) technology and designed for long-term scalability growth. For more information, see Chapter 2, "IBM DS8880 hardware components and architecture" on page 23.

Switched Fibre Channel Arbitrated Loop

When connecting to the standard drive enclosures, the DS8880 uses an 8 Gb switched FC-AL architecture to connect the DAs to these enclosures. The DAs connect to the controller cards in the drive enclosures by using FC-AL with optical shortwave multi-mode interconnections. The Fibre Channel interface cards (FCICs) provide point-to-point SAS connections to each drive and DA, providing four paths from the DS8880 processor complexes to each disk drive. For more information, see Chapter 2, "IBM DS8880 hardware components and architecture" on page 23.

Drive options

In addition to flash cards, the DS8880 offers the following disk drives to meet the requirements of various workloads and configurations. For more information, see Chapter 7, "IBM DS8880 physical planning and installation" on page 179.

► 400 GB, 800 GB, and 1.6 TB flash drives (SSDs) for higher performance requirements

► 300 and 600 GB 15K revolutions per minute (RPM) enterprise disk drives for high-performance requirements

► 600 GB, 1.2 TB, and 1.8 TB 10K RPM disk drives for standard performance requirements

► 4 TB and 6 TB 7,200 RPM nearline LFF disk drives for large-capacity requirements

Flash drives provide up to 100 times the throughput and 10 times lower response time than 15K RPM spinning disks. They also use less power than traditional spinning drives. For more information, see Chapter 7, "IBM DS8880 physical planning and installation" on page 179.

All drives and flash cards in the DS8880 are encryption-capable. Enabling encryption is optional. It requires at least two key servers with the IBM Security Key Lifecycle Manager software, or the Gemalto Safenet, OASIS Key Management Interoperability Protocol (KMIP) compatible software.

IBM Easy Tier

Easy Tier enables the DS8880 to automatically balance I/O access to disk drives to avoid hot spots on disk arrays. Easy Tier can place data in the storage tier that best suits the access frequency of the data. Highly accessed data can be moved nondisruptively to a higher tier, for example, to flash drives. Cold data or data that is primarily accessed sequentially is moved to a lower tier (for example, to nearline disks). Easy Tier includes additional components:

► Easy Tier Application is an application-aware storage utility to help deploy storage more efficiently by enabling applications and middleware to direct more optimal placement of the data by communicating important information about current workload activity and application performance requirements. It is possible for DB2 applications in z/OS environments to give hints of data placement to Easy Tier at the data set level.

► Easy Tier Heat Map Transfer (HMT) can take the data placement algorithm on the Metro Mirror/Global Copy/Global Mirror (MM/GC/GM) primary site and reapply it to the MM/GC/GM secondary site when failover occurs by using the Easy Tier Heat Map Transfer Utility (HMTU). With this capability, the DS8000 systems can maintain application-level performance. The Easy Tier HMT functions support Metro/Global Mirror (MGM) to transfer a heat map automatically to a tertiary site.

> **Note:** Easy Tier Server has been removed from marketing support. It is replaced with the Flash Cache option of AIX 7.2.

The following newer Easy Tier controls are available:

► For Easy Tier to function effectively, it needs some free extents in each extent pool to be able to move around extents. A storage administrator can keep an eye on this and make sure that not all capacity is allocated. However, there is a new control to let the system reserve some space for Easy Tier extent movements. The control is available, for example by use of the `chsi` command specifying `-etsrmode enable` to enable space reservation for Easy Tier.

► Another new control is the allocation policy for new volumes. In previous implementations, new storage was allocated from Tier 1 if available and then from Tier 2, while Tier 0 was not preferred. You can change the allocation policy now according to your needs with the `chsi` command and the `-ettierorder highutil` or `highperf`. The data allocation order can be changed between ENT → NL → SSD(highutil) and SSD → ENT → NL(highperf).

For more information about the Easy Tier features, see the following resources:

► *IBM DS8000 EasyTier*, REDP-4667
► *DS8870 Easy Tier Application*, REDP-5014
► *IBM DS8870 Easy Tier Heat Map Transfer*, REDP-5015
► *IBM DS8870 and IBM z Systems Synergy*, REDP-5186

Host adapters

The DS8880 offers 16 Gbps and 8 Gbps Fibre Channel host adapters:

- ► The 16 Gbps FC adapter has four 16 Gbps-capable ports. Each port independently auto-negotiates to 4, 8, or 16 Gbps link speed, and it can be configured to FCP or FICON.

- ► The 8 Gbps FC adapter has four or eight 8 Gbps-capable ports. Each port independently auto-negotiates to 2, 4, or 8 Gbps link speed, and it can be configured to FC-AL, FCP, or FICON.

For more information, see Chapter 2, "IBM DS8880 hardware components and architecture" on page 23.

Storage Hardware Management Console for the DS8880

The HMC is the focal point for maintenance activities and contains a version of CSM. The HMC is a dedicated workstation that is physically inside the DS8880. The HMC can proactively monitor the state of your system and notify you and IBM when service is required. It can also be connected to your network to enable centralized management of your system by using the IBM data storage command-line interface (DS CLI). The HMC supports the IPv4 and IPv6 standards. For more information, see Chapter 8, "IBM DS8880 Management Console planning and setup" on page 203.

A second HMC is available as an optional feature. The console can be used as a redundant management console for environments with high availability requirements. It is also in the bottom primary frame, and contains a copy of the CSM that also includes an LDAP proxy server.

Isolated key server

The IBM Security Key Lifecycle Manager software performs key management tasks for IBM encryption-enabled hardware, such as the IBM DS8880. IBM Secure Lifecycle Manager provides, protects, stores, and maintains encryption keys that are used to encrypt information that is written to, and decrypt information that is read from encryption-enabled disks.

The DS8880 ships with FDE drives. To configure a DS8880 to use encryption, at least two IBM key servers are required. An isolated key server with dedicated hardware and non-encrypted storage resources is required. It can be ordered from IBM. For more information, see 7.3.6, "Key manager servers for encryption" on page 198.

Alternatively, starting with DS8000 Release 8.1, customers who need to comply with the KMIP have the possibility to use the Gemalto SafeNet KeySecure as the external key server. For details, see the *IBM DS8880 Data-at-rest Encryption*, REDP-4500.

1.3.2 Storage capacity

The physical storage capacity for the DS8880 is installed in fixed increments that are called *drive sets* and *flash card sets*. A drive set contains 16 DDMs, all of which have the same capacity and the same rotational speed. Nearline drives are available in half sets (8) or full sets (16) of disk drives or DDMs. High-performance flash cards are available in sets of 16.

The available drive options provide industry-class capacity and performance to address a wide range of business requirements. The DS8880 storage arrays can be configured as RAID 6, RAID 10, or RAID 5, depending on the drive type. Note that RAID 6 is now the default and preferred setting for the DS8880. RAID5 can be configured for drives of less than 1 TB, but this is not recommended and requires a risk acceptance. RAID10 continues to be an option for all drive types.

For more information, see 2.2, "DS8880 configurations and models" on page 26.

1.3.3 Supported environments

The DS8000 offers connectivity support across a broad range of server environments, including IBM Power Systems™, z Systems, and Lenovo System x servers, servers from Oracle and HPE, non-IBM Intel, and AMD-based servers.

The DS8000 supports over 90 platforms. For the current list of supported platforms, see the DS8000 System Storage Interoperation Center (SSIC) at this website:

https://www.ibm.com/systems/support/storage/ssic

For more information about the supported environments, see *IBM System Storage DS8000: Host Attachment and Interoperability*, SG24-8887.

This rich support of heterogeneous environments and attachments, with the flexibility to easily partition the DS8000 storage capacity among the attached environments, can help support storage consolidation requirements and dynamic environments.

1.3.4 Configuration flexibility

The DS8000 series uses virtualization techniques to separate the logical view of hosts onto LUNs from the underlying physical layer, providing high configuration flexibility. For more information about virtualization, see Chapter 4, "Virtualization concepts" on page 97.

Dynamic volume expansion

Dynamic volume expansion increases the capacity of open systems and z Systems volumes, while the volume remains connected to a host system. This capability simplifies data growth by providing volume expansion without taking volumes offline. Certain operating systems do not support a change in volume size. Therefore, a host action is required to detect the change after the volume capacity is increased.

Large LUN and large count key data volume support

You can configure LUNs and volumes to span arrays, allowing for larger LUN sizes of up to 16 TB in open systems. Copy Services are not supported for LUN sizes that are greater than 2 TB.

The maximum CKD volume size is 1,182,006 cylinders (1 TB), which can greatly reduce the number of volumes that are managed. This large CKD volume type is called a 3390 Model A. It is referred to as an extended address volume (EAV).

T10 data integrity field support

The DS8880 supports the T10 DIF standard for FB volumes that are accessed by the FCP channel of Linux on z Systems and IBM AIX on Power Systems. You can define LUNs with an option to instruct the DS8880 to use the CRC-16 T10 DIF algorithm to store the data. You can also create T10 DIF-capable LUNs. The support for IBM i variable LUNs now adds flexibility for volume sizes and can increase capacity utilization for IBM i environments.

VMware VAAI support

The VAAI feature offloads specific storage operations to disk arrays for greatly improved performance and efficiency. With VAAI, VMware vSphere can perform key operations faster and use less CPU, memory, and storage bandwidth. The DS8880 supports the VAAI primitives ATS, which is also known as *Compare and Write* for hardware-assisted locking, and Clone Blocks (Extended Copy or XCOPY for hardware-assisted move or cloning). VAAI also supports Write Same, Site Recovery Manager (SRM), vCenter plug-in, and variable LUN size.

OpenStack

The DS8000 supports the OpenStack cloud management software for business-critical private, hybrid, and public cloud deployments. The DS8880 supports features in the OpenStack environment, such as volume replication and volume retype. The current supported release is the OpenStack Newton, but also older versions can be used with the respective IBM Storage Driver software level.

For more information about the DS8870 and OpenStack, see *Using IBM DS8870 in an OpenStack Environment*, REDP-5220.

RESTful API

With the DS8880 support of RESTful API services, DS8880 clients or cloud administrators can design and implement the DS8880 management applications by using the Representational State Transfer (REST) software architecture.

For more information about the RESTful API, see *Exploring the DS8870 RESTful API Implementation*, REDP-5187.

Flexible LUN-to-LSS association

With no predefined association of arrays to logical subsystems (LSSs) on the DS8000 series, users can put LUNs or CKD volumes into LSSs and make the best use of the 256 address range, particularly for z Systems.

Simplified LUN masking

In the new GUI, LUNs are directly mapped to the host, and the user cannot define volume groups. Volume groups still exist on the DS8880 systems, but they are not visible from the GUI because they are created in the background during the assignment of a volume to the host.

Thin-provisioning features

Volumes in the DS8880 can be provisioned as full or thin. When clients plan capacity, they need to consider the number of volumes in the extent pool (or overall storage system) and the degree of over-allocation that is planned for.

These volumes feature enabled over-provisioning capabilities that provide more efficient usage of the storage capacity and reduced storage management requirements. For more information, see Chapter 4, "Virtualization concepts" on page 97. For details, refer to*IBM DS8880 Thin Provisioning*, REDP-5343.

Maximum values of logical definitions

The DS8000 features the following maximum values for the major logical definitions:

► Up to 256 LSSs

► Up to 65,280 logical devices

- ► Up to 16 TB LUNs
- ► Up to 1,182,006 cylinder (1 TB) CKD volumes
- ► Up to 130,560 FICON logical paths (512 logical paths for each control unit image) on the DS8000
- ► Up to 1,280 logical paths for each FC port
- ► Up to 8,192 process logins (509 for each SCSI-FCP port)

1.3.5 Copy Services functions

For IT environments that cannot afford to stop their systems for backups, the DS8880 provides IBM FlashCopy®. This fast replication technique can provide a point-in-time copy of the data in a few seconds or even less.

> **Note:** FlashCopy Space Efficient (based on Track Space Efficient) is not supported by the DS8880. FlashCopy targets based on Extent Space Efficient volumes with small extents are now the alternative for both Open and z/OS. For more information, see *IBM DS8880 Thin Provisioning*, REDP-5343.

For data protection and availability needs, the DS8880 provides Metro Mirror, Global Mirror, Global Copy, Metro/Global Mirror, and z/OS Global Mirror, which are Remote Mirror and Copy functions. These functions are also available and are fully interoperable with previous models of the DS8000 family. These functions provide storage mirroring and copying over large distances for disaster recovery or availability purposes.

Copy Services scope limiting is the ability to specify policy-based limitations on Copy Services requests. With the combination of policy-based limitations and other inherent volume-addressing limitations, you can control these objects:

- ► Volumes that can be in a Copy Services relationship
- ► Which network users or host logical partitions (LPARs) can issue Copy Services requests on which resources
- ► Other Copy Services operations

For more information about Copy Services, see Chapter 5, "DS8000 Copy Services overview" on page 129. You can also see *DS8000 Copy Services*, SG24-8367.

1.3.6 Service and setup

The installation of the DS8000 is performed by IBM in accordance with the installation procedure for this system. The client's responsibility is the installation planning, retrieval, and installation of feature activation codes, logical configuration, and execution.

For maintenance and service operations, the storage HMC is the focus. The management console is a dedicated workstation that is physically inside the DS8880 where it can automatically monitor the state of your system. It notifies you and IBM when service is required. Generally, use a dual-HMC configuration, particularly when Full Disk Encryption is used.

The HMC is also the interface for remote services (call home and remote support), which can be configured to meet client requirements. It is possible to allow one or more of the following configurations:

► Call home on error (machine-detected)
► Connection for a few days (client-initiated)
► Remote error investigation (service-initiated)

The remote connection between the management console and the IBM Service organization is performed by using the Assist On-site (AOS) feature. AOS offers more options, such as Secure Sockets Layer (SSL) security and enhanced audit logging. For more information, see *IBM Assist On-site for Storage Overview*, REDP-4889.

The DS8000 storage system can be ordered with an outstanding four-year warranty (an industry first) on hardware and software.

1.3.7 IBM Certified Secure Data Overwrite

IBM Certified Secure Data Overwrite (SDO) is a process that provides a secure overwrite of all data storage in a DS8880 storage system. Before you perform a secure data overwrite, you must remove all logical configuration. Encryption groups, if configured, must also be disbanded. The process is then initiated by the IBM service support representative (SSR). The process continues unattended until it completes. This process takes a full day to complete. Two DDM overwrite options exist.

DDM overwrite options
Two DDM overwrite options exist for SDO: Cryptoerase and three-pass overwrite.

Cryptoerase
This option performs a cryptoerase of the drives, which invalidates the internal encryption key on the DDMs, rendering the previous key information unreadable. It then performs a single-pass overwrite on all drives.

Three-pass overwrite
This option also performs a cryptoerase of the drives, then performs a three-pass overwrite on all drives. This overwrite pattern allows compliance with the NIST Special Publication 800-88r1 standard.

CPC and HMC
A three-pass overwrite is performed on both the CPC and HMC disk drives. If a secondary HMC is associated with the storage system, SDO is run against the secondary HMC after the run completes on the primary HMC.

SDO process overview
The SDO process can be summarized in these steps:

1. The client removes all logical configuration and encryption groups.
2. The IBM SSR starts SDO from the HMC.
3. SDO performs a dual-cluster restart of the CPCs.
4. SDO cryptoerases all drives and flash cards in the storage system.
5. SDO starts an overwrite method.
6. SDO starts a three-pass overwrite on the CPC and HMC hard disks.
7. When the SDO process completes, SDO generates a certificate.

Certificate

The certificate provides written verification, by drive or Flash Card serial number, of the full result of the overwrite operations. You can retrieve the certificate by using DSCLI, or the IBM SSR can offload the certificate to removable media, and provide the media to you.

1.4 Performance features

The DS8880 offers optimally balanced performance. The DS8880 incorporates many performance enhancements, such as a dual multi-core IBM POWER8 processor complex implementation in the DS8886, fast 16 Gbps, and 8 Gbps Fibre Channel/FICON host adapters, HPFE dedicated flash architecture in second generation, classical flash drives, and high-bandwidth, fault-tolerant point-to-point PCI Express Gen-3 internal connections.

With all of these components, the DS8880 is positioned at the top of the high-performance category.

1.4.1 16 Gbps host adapters

The DS8880 supports 16 Gbps host adapters. This connectivity reduces latency and provides faster single stream (12% - 14% improvements for zHPF) and per port throughput. These cards can work also with 8 or 4 Gbps compatibility, with no FC-AL connections. The *Lights on Fastload* feature avoids path disturbance during code loads on 16 Gbps host adapters.

1.4.2 Sophisticated caching algorithms

IBM Research conducts extensive investigations into improved algorithms for cache management and overall system performance improvements. To implement sophisticated caching algorithms, it is essential to include powerful processors for the cache management. With a 4 KB cache segment size and up to 2 TB overall cache sizes, the tables to maintain the cache segments become large.

Sequential Prefetching in Adaptive Replacement Cache

One of the performance features of the DS8000 is its self-learning cache algorithm, which optimizes cache efficiency and enhances cache hit ratios. This algorithm, which is used in the DS8000 series, is called Sequential Prefetching in Adaptive Replacement Cache (SARC).

SARC provides the following abilities:

► Sophisticated algorithms to determine the data to store in cache that is based on recent access and the frequency needs of the hosts

► Prefetching, which anticipates data before a host request and loads it into cache

► Self-learning algorithms to adaptively and dynamically learn the data to store in cache that is based on the frequency needs of the hosts

Adaptive Multi-stream Prefetching

AMP is a breakthrough caching technology that improves performance for common sequential and batch processing workloads on the DS8000. AMP optimizes cache efficiency by incorporating an autonomic, workload-responsive, and self-optimizing prefetching technology.

Intelligent Write Caching

Intelligent Write Caching (IWC) improves performance through better write-cache management and destaging the order of writes. It minimizes disk actuator movements on writes so that the disks can perform more I/O in total. IWC can also double the throughput for random write workloads. Specifically, database workloads benefit from this new IWC cache algorithm.

SARC, AMP, and IWC play complementary roles. Although SARC is carefully dividing the cache between the RANDOM and the SEQ lists to maximize the overall hit ratio, AMP is managing the contents of the SEQ list to maximize the throughput that is obtained for the sequential workloads. IWC manages the write cache and decides the order and rate to destage to disk.

1.4.3 Flash storage

Today's installations of new storage systems are mostly either hybrid (multitier) installations, mixing spinning drives and flash together in one system. It is sometimes also referred to as a Hybrid-flash array (HFA) then. For clients considering an all-flash strategy, an All-flash array (AFA) can be deployed.

The IBM DS8880 offers the following options:

► The DS8880 can act as a Tier 1 machine, or combine flash and HDDs into a multitier hybrid-flash storage system, with automated algorithms optimizing the tiering and placing hot areas onto flash.

► The DS8880 can hold a large quantity of flash drives (SSDs) to allow an all-flash deployment in larger capacities, thereby already benefitting from the low SSD latency values for each I/O, and having an All-flash array.

► You can also configure the DS8880 with High-Performance Flash Enclosures exclusively, for an all-flash solution or in a hybrid (mixed) configuration.

To improve data transfer rate (IOps) and response time, the DS8880 supports flash drives and high-performance flash cards, which are based on NAND technology. With the flash cards and the specific architecture being used in the High-Performance Flash Enclosures, much higher IOps densities (IOps/GB) are possible, than with ordinary SSDs.

The flash drives feature improves I/O transaction-based performance over traditional spinning drives in standard drive enclosures. The DS8880 is available with 400 GB, 800 GB, and 1.6 TB encryption-capable flash drives.

High-performance flash cards offer even higher throughputs by using the integrated Flash RAID adapters in the HPFE, and PCI Express connections to the processor complexes. The DS8880 is available with encryption-capable flash cards in capacities from 400 GB to 3.2 TB. Flash drives and flash cards are high-IOPS class enterprise storage devices that are targeted at Tier 0, I/O-intensive workload applications that can use a high level of fast-access storage. Flash drives offer many potential benefits over rotational drives, including better IOPS, lower power consumption, less heat generation, and lower acoustical noise. For more information, see Chapter 7, "IBM DS8880 physical planning and installation" on page 179.

1.4.4 Performance for z Systems

The DS8000 series supports the following IBM performance enhancements for z Systems environments:

► PAVs enable a single z Systems server to simultaneously process multiple I/O operations to the same logical volume, which can significantly reduce device queue delays. This reduction is achieved by defining multiple addresses for each volume. With Dynamic PAV, the assignment of addresses to volumes can be automatically managed to help the workload meet its performance objectives and reduce overall queuing.

► HyperPAV is designed to enable applications to achieve equal or better performance than with PAV alone, while also using fewer unit control blocks (UCBs) and eliminating the latency in targeting an alias to a base. With HyperPAV, the system can react immediately to changing I/O workloads.

► New since DS8880 release 8.1, SuperPAV allows z/OS to use an *alias address* from another LCU.

► Multiple Allegiance expands the simultaneous logical volume access capability across multiple z Systems. This function, with PAV, enables the DS8000 to process more I/Os in parallel, which improves performance and enables greater use of large volumes.

► I/O priority queuing allows the DS8000 series to use I/O priority information that is provided by the z/OS Workload Manager to manage the processing sequence of I/O operations at the adapter level.

► I/O Priority Manager provides application-level QoS for workloads that share a storage pool. It provides a way to manage QoS for I/O operations of critical workloads and gives them priority over other I/O operations that are associated with non-critical workloads. For z/OS, the I/O Priority Manager allows increased interaction with the host side (zWLM integration).

► zHPF reduces the effect that is associated with supported commands on current adapter hardware. This configuration improves FICON throughput on the DS8000 I/O ports. The DS8000s also supports the new zHPF I/O commands for multi-track I/O operations, DB2 list-prefetch, sequential access methods, and DB2 castout acceleration.

► zHyperwrite is another enhancement for DB2 clients. In a z/OS Metro Mirror environment, it enables DB2 log updates to be written to the primary and secondary volumes in parallel. This configuration reduces the latency for log writes, and so improving transactional response times and log throughput. The Metro Mirror primary volume needs to be enabled with IBM HyperSwap® by either IBM Geographically Dispersed Parallel Sysplex™ (IBM GDPS®) or CSM.

For more information about performance on z Systems, see *IBM DS8870 and IBM z Systems Synergy*, REDP-5186.

1.4.5 Performance enhancements for IBM Power Systems

Many IBM Power Systems users can benefit from the following DS8000 features:

► End-to-End I/O priorities
► Cooperative caching
► Long busy wait host tolerance
► Automatic port queues

For more information about performance enhancements, see Chapter 6, "Designed for performance" on page 145.

2

IBM DS8880 hardware components and architecture

This chapter describes the hardware components of the IBM DS8880. It provides insights into the architecture and individual components.

This chapter covers the following topics:

► Terminology of the DS8880
► DS8880 configurations and models
► DS8880 architecture overview
► I/O enclosures and adapters
► Storage enclosures and drives
► Power and cooling
► Management Console and network
► Hardware feature summary for DS8880 configurations

2.1 Terminology of the DS8880

It is important to understand the naming conventions that are used to describe the DS8880 components and features. Although most terms are introduced in other chapters of this book, they are repeated and summarized here.

2.1.1 Storage system

The term *storage system* describes a single DS8880 (base frame plus optional expansion frames).

Base frame

The DS8880 has three available base frame models. The model numbers will depend on the hardware release for each DS8888, DS8886, or DS8884 system configuration. The base frame models for each are listed in Figure 2-1.

Each is a complete storage system that can be contained within a single base frame. To increase the storage capacity and connectivity, expansion frames can be added. Each base frame is equipped with a Hardware Management Console (HMC). The base frame can also contain an optional second HMC.

For more information about the base frame configuration, see 2.2.4, "DS8880 base frames" on page 31.

Hardware Release	Base Frame Model	Expansion Frame Model	Max Expansion Frames
DS8888 Configuration			
8.1	982	98F	1
8.2	988	88E	1
DS8886 Configuration			
8.0	981	98E	4
8.2	985 (Single-phase power)	85E	4
8.2	986 (Three-phase power)	86E	4
DS8884 Configuration			
8.0	980	98B	2
8.2	984	84E	2

Figure 2-1   Frame models by system configuration and hardware release

Expansion frame

For each of the DS8888, DS8886, and DS8884 system configurations, one or more optional expansion frames can be installed. The model numbers for the expansion frames likewise will depend on the hardware release and system configuration of the base frame. The expansion frame models for each are also listed in Figure 2-1.

For each system configuration, the first expansion frame provides additional storage capacity and can also contain additional I/O enclosures. Subsequent expansion frames provide additional storage capacity.

A single expansion frame can be added to the DS8888 configuration. Up to four expansion frames can be added to the DS8886 configuration, and up to two expansion frames can be added to the DS8884 configuration.

To add an expansion frame to a DS8888, the system must first be configured with 2048 GB of total system memory. This configuration includes two additional 12-core processors per server, for a total of 48 cores per server.

To add expansion frames to a DS8886 system, the system must first be configured with a minimum of 256 GB of total system memory and 16-core processors per server.

For a DS8884 system, the minimum total system memory that is required to add expansion frames is 128 GB.

All DS8880 system memory and processor upgrades can be performed concurrently.

Expansion frames of previous generation DS8000 storage systems are not supported and cannot be installed in a DS8880 storage system. DS8880 expansion frames can only be added to systems with the same configuration and hardware release. For more information about the expansion frame configuration, see 2.2.5, "DS8880 expansion frames" on page 33.

2.1.2 Management Console

The Management Console, also known as the HMC, is the focal point for management operations of the DS8880. The HMC provides connectivity to the client network, and communications to the system private networks, power subsystem, and other management domains. All storage configuration, user-controlled tasks, and service actions are managed through the HMC. Although many other IBM products use an HMC, the installed Licensed Internal Code (LIC) makes the DS8880 HMC unique to these systems.

The DS8880 HMC also includes the IBM Copy Services Manager code. For details, see *IBM DS8880 Integrated Copy Services Manager and LDAP Client on the HMC*, REDP-5356.

2.1.3 Central processor complex

The DS8880 has two POWER8 servers, which are referred to as central processor complexes (CPCs). The CPC is also known as the processor complex or the internal server. The CPCs for each model type have these characteristics:

► Each DS8888 CPC can have either 24 or 48 processor cores, and either 512 GB or 1024 GB of processor memory, for 1024 GB or 2048 GB of total system memory.

► Each DS8886 CPC can have 8 - 24 processor cores, and 64 - 1,024 GB of processor memory, for 128 GB - 2048 GB of total system memory.

► Each DS8884 CPC has six processor cores and can have 32 - 128 GB of processor memory, for 64 GB - 256 GB of total system memory.

Both CPCs in a DS8880 system share the system workload. The CPCs are redundant, and either CPC can fail over to the other CPC if a failure occurs, or for scheduled maintenance or upgrade tasks. The CPCs are identified as CPC 0 and CPC 1. A logical partition in each CPC runs the AIX V7.*x* operating system and storage-specific LIC. This logical partition is called the *storage node*. The storage servers are identified as *Node 0* and *Node 1*.

Similar to earlier generations of the IBM DS8000 series, the DS8880 consists of one base frame, and up to four optional expansion frames. The CPCs reside in the base frame, and can be upgraded with additional processor cores and system memory to accommodate growing performance, or when additional storage capacity or host connectivity is required.

Upgrades from the smallest to the largest configuration in terms of system memory, processors, storage capacity, and host attachment can be performed concurrently.

Note: The DS8884, DS8886, and DS8888 hardware platforms are specific to each configuration. Upgrades are not supported from one hardware platform to another.

2.2 DS8880 configurations and models

This section presents the current DS8880 configurations and models. The DS8880 storage systems are associated with machine type 283*x*, or 533x for All-Flash models. The last digit of the machine type corresponds to the length of warranty option chosen, where *x* equals the number of years. DS8880 offers a one-year, two-year, three-year, or four-year warranty period. Expansion frames must have the same machine type as the base frame.

The DS8880 is designed for modular expansion. From a high-level view, IBM offers three configurations of the DS8880:

► DS8888 configuration
► DS8886 configuration
► DS8884 configuration

However, the physical frames themselves are almost identical. The main variations are the combinations of CPCs, I/O enclosures, storage enclosures, disks, and direct current uninterruptable power sources (DC-UPSs) that the frames contain.

2.2.1 DS8888 All-Flash configuration

The DS8888 is available with dual 24-core, or dual 48-core processor complexes, with up to 16 Fibre Channel/Fibre Channel connection (FICON) host adapters (HAs) in the base frame, and up to 16 additional Fibre Channel/FICON HAs in the optional expansion frame, for a total of up to 128 host ports in the system as shown in Figure 2-2.
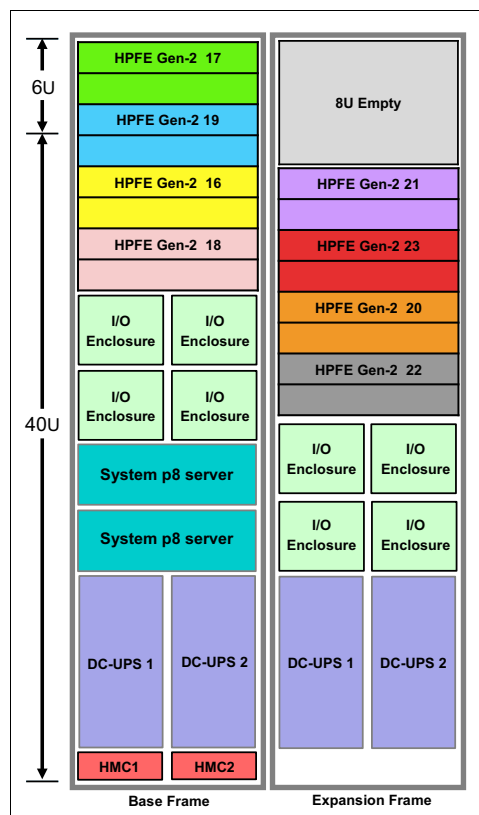
Figure 2-2   Front view of fully configured DS8888 All-Flash configuration

The DS8888 configuration can be configured with either 1 TB or 2 TB of system memory, and up to four High-Performance Flash Enclosure (HPFE) Gen2 pairs in the base frame. An additional four HPFE Gen2 pairs can be added in the expansion frame. Each HPFE Gen2 pair can contain either 16, 32, or 48 Flash Cards, of 400 GB, 800 GB, 1.6 TB, or 3.2 TB capacity. The DS8888 All-Flash configuration offers up to 1228.8 TB of raw flash storage capacity.

Figure 2-2 on page 26 shows a newer fully configured DS8888 All-Flash configuration, for hardware available with release 8.2.1. The first frame contains the processor complexes, I/O enclosures, and up to four HPFE Gen2 pairs. The second frame contains four additional I/O enclosures, and up to four additional HPFE Gen2 pairs.

Each frame contains two standard 3-phase DC-UPSs that supply redundant power for all installed components. All DC-UPSs in a system contain one battery service module (BSM) set, whether the extended power line disturbance (ePLD) feature is installed or not.

Note: The DS8888 configuration supports only three-phase input power.

2.2.2 DS8886 configuration

The DS8886 is available with dual 8-core, dual 16-core, or dual 24-core processor complexes, with up to 16 Fibre Channel/FICON host adapters HAs in the base rack, and up to 16 Fibre Channel/FICON HAs in the first expansion rack, for a total of up to 128 host ports in the system.

The DS8886 configuration is optimized for performance. It is highly scalable, offering a wide range of options for long-term growth. The DS8886 can be configured with 128 GB - 2 TB of system memory and supports up to four expansion frames.

Figure 2-3 shows a fully configured, five-frame DS8886 configuration with single-phase DC-UPSs. The leftmost frame is a base frame that contains the processor complexes, I/O enclosures, standard drive enclosures, and HPFE Gen2s. The second frame is the first expansion frame that contains I/O enclosures, HPFEs, and standard drive enclosures. The third and fourth frames are also expansion frames that contain only standard drive enclosures.
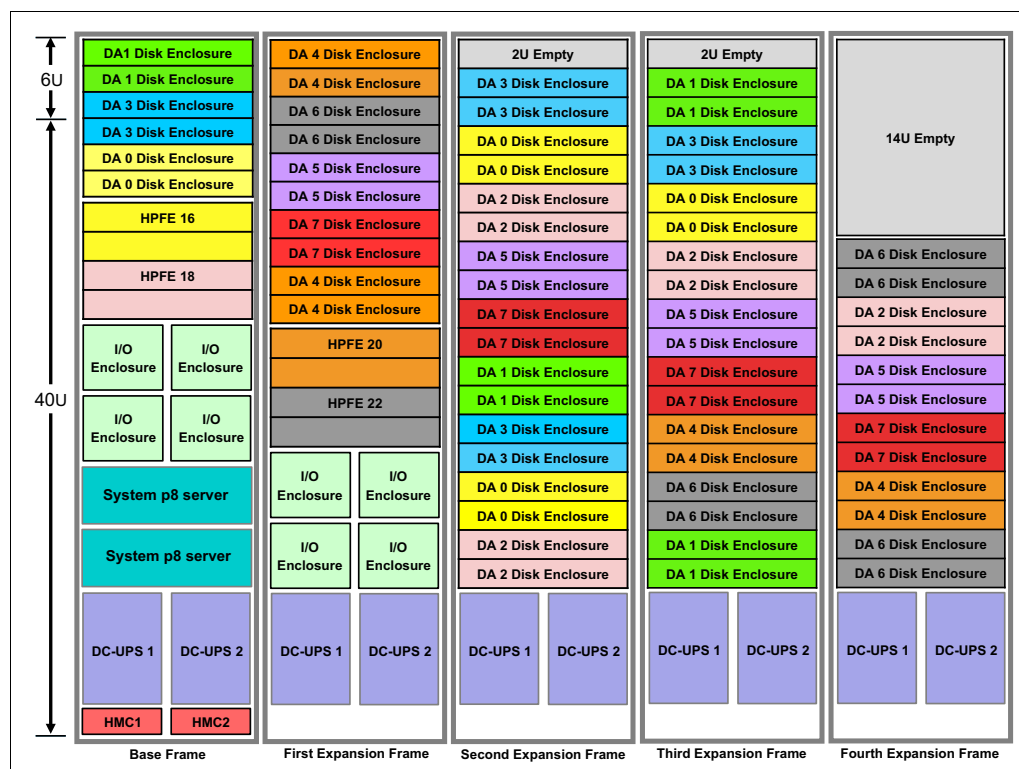
Figure 2-3 Front view of fully configured DS8886 configuration with single-phase DC-UPSs

Each frame contains two standard single-phase DC-UPSs or two optional 3-phase DC-UPSs that supply redundant power for all installed components. All DC-UPSs in a system contain one BSM set, whether the ePLD feature is installed or not.

Note: All DC-UPSs in a system must be either single-phase or 3-phase. DS8886 systems do not support field conversion from single-phase to 3-phase DC-UPSs, or from 3-phase to single-phase DC-UPSs.

Figure 2-4 shows the same configuration with 3-phase DC-UPSs and HPFEs Gen2.
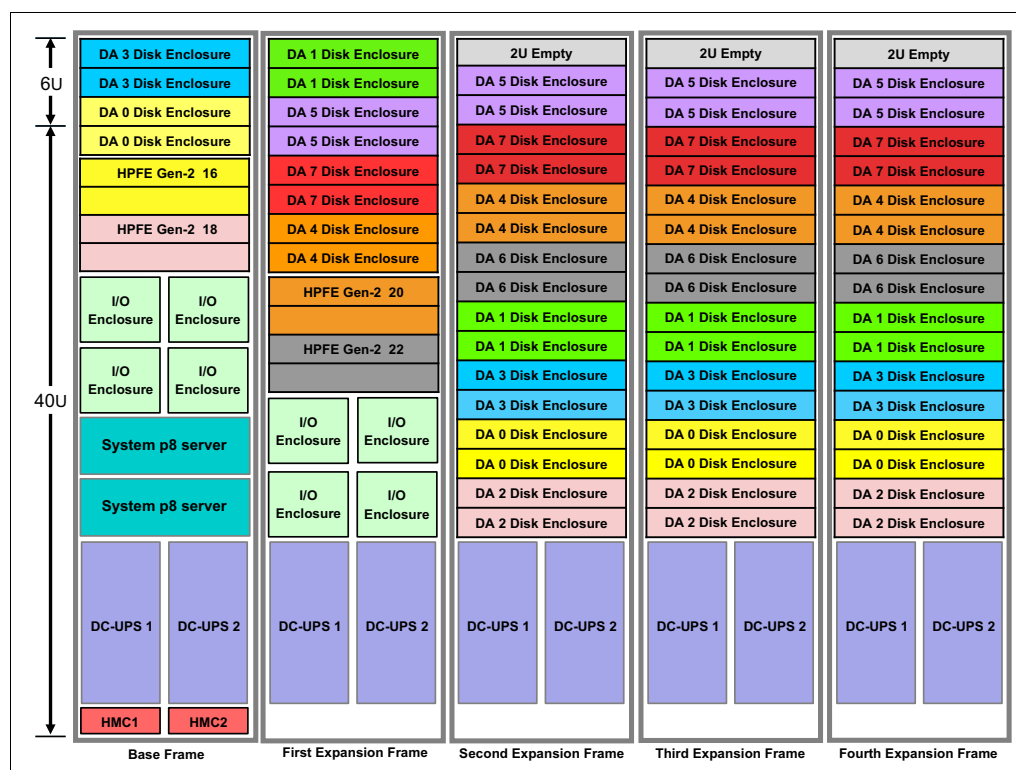
Figure 2-4   Front view of fully configured DS8886 configuration with 3-phase DC-UPSs

2.2.3 DS8884 configuration

The DS8884 configuration is available with a 6-core processor complex, with up to eight Fibre Channel/FICON HAs in the base frame, and up to eight Fibre Channel/FICON HAs in the first expansion frame, for a total of up to 128 host ports in the system.

The DS8884 configuration employs a different standard drive enclosure cabling scheme to reduce initial configuration costs compared to the DS8886 configuration. The DS8884 configuration increases device adapter (DA) utilization, prioritizing cost-effective storage capacity growth. The DS8884 configuration can be configured with 64 GB - 256 GB of system memory. The DS8884 configuration supports up to two expansion frames.

Figure 2-5 shows a fully configured three-frame DS8884 configuration. The leftmost frame is the base frame that contains the processor complexes, I/O enclosures, standard drive enclosures, and HPFEs Gen2. The second frame is the first expansion frame that contains I/O enclosures, HPFEs, and standard drive enclosures. The third frame is also an expansion frame that contains only standard drive enclosures.

Each frame contains two standard single-phase DC-UPSs that supply redundant power for all installed components. All DC-UPSs in a system contain one BSM set, whether the ePLD feature is installed or not.

Note: The DS8884 configuration supports only single-phase input power.
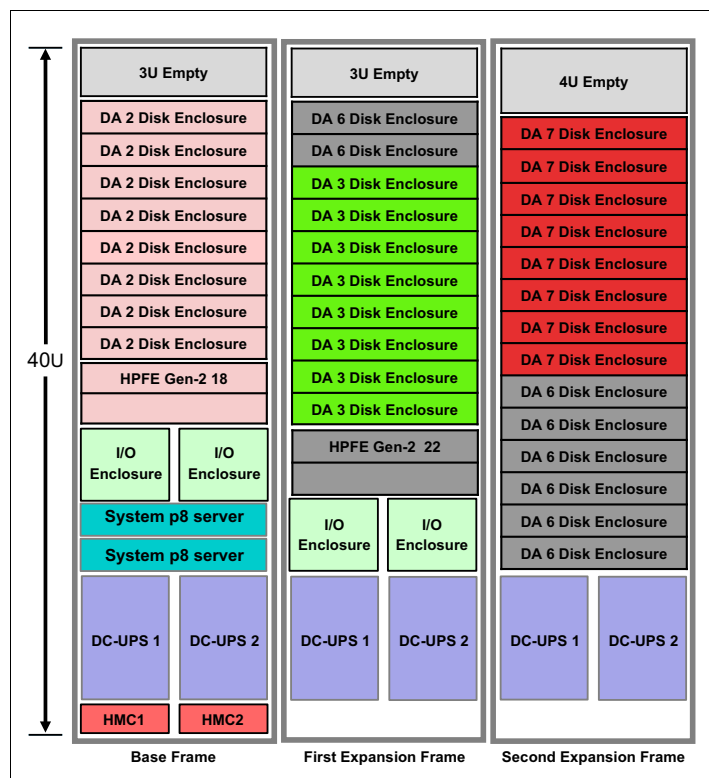
Figure 2-5   Front view of fully configured DS8884 configuration

2.2.4 DS8880 base frames

As mentioned in 2.1.1, "Storage system" on page 24, the DS8880 storage system is available in three different base frame models. Each base frame model number is determined by the configuration and the hardware version. The specific combinations are shown in Figure 2-1 on page 24.

The DS8880 base frames accommodate the following components:

► High-Performance Flash Enclosure Gen2

The following configurations are possible:

- Up to four HPFE Gen2 enclosure pairs in the DS8888 configuration base frame.
- Up to two HPFE Gen2 enclosure pairs in the DS8886 configuration base frame.
- Up to one HPFE Gen2 enclosure pair in the DS8884 configuration base frame.

Each HPFE Gen2 enclosure pair can accommodate sixteen, thirty-two, or forty-eight flash cards. Flash cards are available in 400 GB, 800 GB, 1.6 TB, and 3.2 TB capacities. For more information about the HPFE, see 2.5, "Storage enclosures and drives" on page 48.

► Standard drive enclosures

Standard drive enclosures can accommodate up to twenty-four 2.5-inch small form factor (SFF) serial-attached SCSI (SAS) drives, or twelve 3.5-inch large form factor (LFF) SAS nearline drives. Standard drive enclosures are installed in pairs. For more information about the standard drive enclosures, see 2.5, "Storage enclosures and drives" on page 48. For supported maximums by configuration, see Figure 2-9 on page 38.

All enclosures have redundant power and integrated cooling, which draw air from front to rear. For more information about cooling requirements, see Chapter 7, "IBM DS8880 physical planning and installation" on page 179.

► The power subsystem

Each DS8880 base frame contains two Direct Current Uninterruptible Power Supplies (DC-UPSs), in a fully redundant configuration. Each DC-UPS contains one integrated Battery Service Module (BSM) set. Each DC-UPS has a dedicated input AC power cord. If input AC is not present on one power cord, it is associated DC-UPS continues to operate by using rectified AC from the partner DC-UPS, with no reduction of system power redundancy. If neither AC input is active, the DC-UPSs switch to battery power. If input power is not restored within 4 seconds (40 seconds with ePLD), the DS8880 initiates an orderly system shutdown. For more information about the DC-UPS, see 2.6, "Power and cooling" on page 56.

The DS8888 configuration uses 3-phase DC-UPSs exclusively. The DS8886 configuration can be ordered with single-phase or 3-phase DC-UPSs, whereas the DS8884 configuration uses single-phase DC-UPSs exclusively.

The power subsystem in the DS8880 complies with the latest directives for the Restriction of Hazardous Substances (RoHS), and is engineered to comply with US Energy Star guidelines.

► Hardware Management Consoles (HMCs)

Each base frame comes with two HMCs that are located below the DC-UPSs, and two Ethernet switches (only seen from the rear). For more information about the Management Console, see 2.7, "Management Console and network" on page 58.

► POWER8 CPCs.

Each base frame accommodates two POWER8 CPCs. The POWER8 processor-based servers contain the processors and memory that drive all functions within the DS8880. System memory and processor cores in the DS8888 and DS8886 configurations can be upgraded concurrently. The DS8884 offers concurrent system memory upgrades. For more information about the processor complexes, see 2.3.1, "IBM POWER8-based servers" on page 35.

► I/O enclosures

The DS8888 and DS8886 base frames accommodate four I/O enclosures. The DS8884 base frame accommodates two I/O enclosures. I/O enclosures are installed in pairs. The I/O enclosures provide PCIe Gen3 connectivity from the I/O adapters and the processor complexes.

The I/O enclosures house the following PCIe I/O adapters:

– 8 Gbps Fibre Channel Host Adapters (HAs) with up to four HAs for each I/O enclosure:

 • Either 4-port or 8-port.
 • Either shortwave (SW) or longwave (LW).

– 16 Gbps Fibre Channel HAs with up to four HAs for each I/O enclosure:

 • 4-port only.
 • Either shortwave (SW) or longwave (LW).

> **Note:** An intermix of 8 Gbps Fibre Channel HAs and 16 Gbps Fibre Channel HAs is supported. This intermixture influences the maximum number of possible host ports. For example, a configuration with 16 Gbps Fibre Channel HAs only, each with four ports, supports a maximum of 128 host ports for a DS8886 and 64 host ports for a DS8884.

The HAs can be configured in the following manner:

 • Fibre Channel Arbitrated Loop (FC-AL) for open systems host attachment (only supported by 8 Gbps Fibre Channel HAs).

 • Switched Fibre Channel Protocol (FCP), which is also for open systems host attachment, and for Metro Mirror and Global Copy.

 • FICON for z Systems host connectivity and also for z/OS Global Mirror.

– 8 Gbps Fibre Channel Device Adapters (DAs) in DS8886 and DS8884:

 • Up to two DA pairs for each I/O enclosure pair.
 • Four ports for each DA.
 • Connectivity that is provided to the standard drive enclosures.

For more information about I/O enclosures and I/O adapters, see 2.4, "I/O enclosures and adapters" on page 41.

> **Note:** All DS8880 frames are 19 inches wide with a 40U capacity. The DS8886 frames and DS8888 optionally can be extended from 40U to 46U to allow greater storage capacity in a smaller footprint.

2.2.5 DS8880 expansion frames

All D8880 system configurations *except DS8884F* can also have one or more expansion frames. The total number of frames depends on the configuration and model. The supported expansion frame options are shown in Figure 2-1 on page 24. Systems must meet certain minimum processor and system memory features, before adding expansion frames.

The DS8880 base frames accommodate the following components:

► High-Performance Flash Enclosure Gen2

The following configurations are possible:

- Up to four HPFE Gen2 enclosure pairs in the DS8888F expansion frame.
- Up to two HPFE Gen2 enclosure pairs in the first DS8886 expansion frame.

Each HPFE Gen2 enclosure pair can accommodate 16, 32, or 48 flash cards. Flash cards are available in 400 GB, 800 GB, 1.6 TB, and 3.2 TB capacities. For more information about the HPFE, see 2.5, "Storage enclosures and drives" on page 48.

► Standard drive enclosures

Standard drive enclosures can accommodate up to twenty-four 2.5-inch small form factor (SFF) serial-attached SCSI (SAS) drives, or twelve 3.5-inch large form factor (LFF) SAS nearline drives. Standard drive enclosures are installed in pairs. For more information about the standard drive enclosures, see 2.5, "Storage enclosures and drives" on page 48. For supported maximums by configuration, see Figure 2-9 on page 38.

All enclosures have redundant power and integrated cooling, which draw air from front to rear. For more information about cooling requirements, see Chapter 7, "IBM DS8880 physical planning and installation" on page 179.

► The power subsystem

Each DS8880 base frame contains two DC-UPSs in a fully redundant configuration. Each DC-UPS contains one integrated Battery Service Module (BSM) set. Each DC-UPS has a dedicated input AC power cord. If input AC is not present on one power cord, it is associated DC-UPS continues to operate by using rectified AC from the partner DC-UPS, with no reduction of system power redundancy. If neither AC input is active, the DC-UPSs switch to battery power. If input power is not restored within 4 seconds (40 seconds with ePLD), the DS8880 initiates an orderly system shutdown. For more information about the DC-UPS, see 2.6, "Power and cooling" on page 56.

The DS8888 configuration uses 3-phase DC-UPSs exclusively. The DS8886 configuration can be ordered with single-phase or 3-phase DC-UPSs, whereas the DS8884 configuration uses single-phase DC-UPSs exclusively.

The power subsystem in the DS8880 complies with the latest directives for the Restriction of Hazardous Substances (RoHS), and is engineered to comply with US Energy Star guidelines.

► I/O enclosures

The DS8888 and DS8886 expansion frames accommodate four I/O enclosures. The DS8884 expansion frame accommodates two I/O enclosures. I/O enclosures are installed in pairs. The I/O enclosures provide PCIe Gen3 connectivity from the I/O adapters and the processor complexes.

The I/O enclosures house the following PCIe I/O adapters:

– 8 Gbps Fibre Channel Host Adapters (HAs) with up to four HAs for each I/O enclosure:

  • Either 4-port or 8-port.
  • Either shortwave (SW) or longwave (LW).

– 16 Gbps Fibre Channel HAs with up to four HAs for each I/O enclosure:

  • 4-port only.
  • Either shortwave (SW) or longwave (LW).

> **Note:** An intermix of 8 Gbps Fibre Channel HAs and 16 Gbps Fibre Channel HAs is supported. This intermixture influences the maximum number of possible host ports. For example, a configuration with 16 Gbps Fibre Channel HAs only, each with four ports, supports a maximum of 128 host ports for a DS8886 and 64 host ports for a DS8884.

The HAs can be configured in the following manner:

  • Fibre Channel Arbitrated Loop (FC-AL) for open systems host attachment (only supported by 8 Gbps Fibre Channel HAs).

  • Switched Fibre Channel Protocol (FCP), which is also for open systems host attachment, and for Metro Mirror and Global Copy.

  • FICON for z Systems host connectivity and also for z/OS Global Mirror.

– 8 Gbps Fibre Channel Device Adapters (DAs) in DS8886 and DS8884:

  • Up to two DA pairs for each I/O enclosure pair.
  • Four ports for each DA.
  • Connectivity that is provided to the standard drive enclosures.

For more information about I/O enclosures and I/O adapters, see 2.4, "I/O enclosures and adapters" on page 41.

> **Note:** All DS8880 frames are 19 inches wide with a 40U capacity. The DS8886 frames and DS8888 optionally can be extended from 40U to 46U to allow greater storage capacity in a smaller footprint.

2.2.6 Scalable upgrades

The hardware features that are supported in the DS8880 depend on the total system memory that is installed. This design ensures that the performance and capacity scale correctly. For more information about processor and system memory requirements for hardware upgrades, see Figure 2-9 on page 38.

Each of the DS8880 configurations can be nondisruptively upgraded from the smallest system memory feature to the largest memory feature that is supported by that configuration.

> **Note:** The DS8884 and DS8886 configured systems cannot be upgraded or modified to a DS8886 or DS8888 configuration. All upgrades must be within the same configuration.

2.2.7 Licensed functions

Several of the IBM DS8880 functions require a license key. For more information about licensed functions, see Chapter 9, "IBM DS8880 features and licensed functions" on page 219.

2.3 DS8880 architecture overview

This section provides an architectural overview of the major components of the DS8880:

- ► IBM POWER8 central processor complexes
- ► I/O enclosures and adapters
- ► PCIe connectivity and communication
- ► Storage subsystem
- ► Hardware management

2.3.1 IBM POWER8-based servers

All DS8880 systems include two CPCs. In the DS8888 configuration, these are Power E850 servers, which have four processor sockets. They are populated with either two or four 12-core 3.02 GHz processors, for a combined total of 24 or 48 cores per CPC. The DS8888 CPCs support a maximum of 32 DDR3 custom dual inline memory modules (CDIMMs), with a maximum memory size of 1 TB per CPC.

Figure 2-6 shows the processor complex as configured in the DS8888.
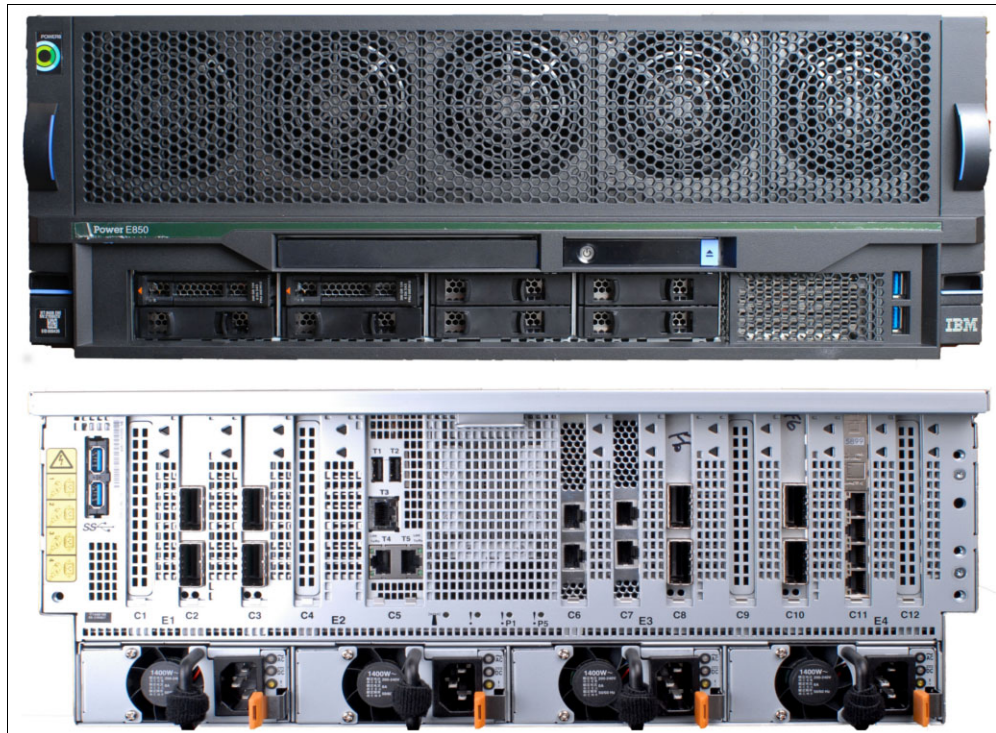
Figure 2-6   DS8888 processor complex front and rear views

The DS8888 CPC is a 4U-high drawer, and features the following configuration:

► Sixteen Double Data Rate-3 (DDR3) CDIMM slots
► One storage cage with two hard disk drives (HDDs)
► Four or eight PCIe x16 Gen3 slots
► Three PCIe x8 Gen3 slots
► Four power supplies

For more information about the server hardware that is used in the DS8888, see *IBM Power System E850 Technical Overview and Introduction*, REDP-5222.

The DS8886 configuration uses Power S824 servers that have two processor sockets. These servers are populated with single or dual 8-core 4.15 GHz processors, or dual 12-core 3.52 GHz processors, for a combined total of 8, 16, or 24 cores per CPC. The DS8886 CPCs support a maximum of 16 DDR3 CDIMMs, with a maximum memory size of 1 TB per CPC.

The DS8886 CPC is a 4U-high drawer, and features the following configuration:

► Sixteen DDR3 CDIMM slots
► One storage cage with two HDDs
► Four PCIe x16 Gen3 slots
► Six PCIe x8 Gen3 slots
► Four power supplies

Figure 2-7 shows the processor complex as configured in the DS8886.

Figure 2-7   DS8886 processor complex front and rear views

For more information about the server hardware that is used in the DS8886, see *IBM Power Systems S814 and S824 Technical Overview and Introduction*, REDP-5097.

The DS8884 configuration uses Power S822 servers, each with a single 6-core 3.89 GHz processor. The DS8884 CPCs support a maximum of 8 DDR3 CDIMMs, with a maximum memory size of 128 GB per CPC. Figure 2-8 shows the processor complex for the DS8884.

The DS8884 CPC is a 2U-high drawer, and features the following configuration:

► Eight DDR3 CDIMM slots
► One storage cage with two hard disk drives
► Two PCIe x16 Gen3 slots
► Four PCIe x8 Gen3 slots
► Two power supplies
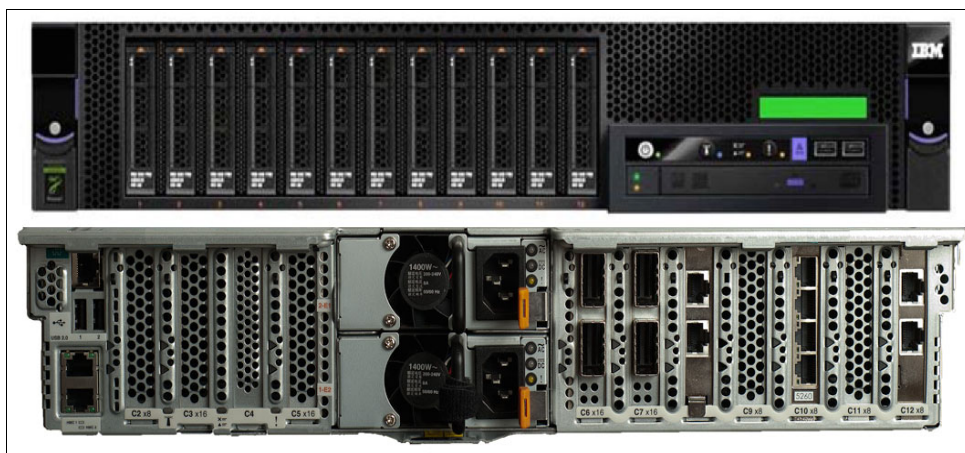
Figure 2-8   DS8884 processor complex front and rear views

For more information about the server hardware that is used in the DS8884, see *IBM Power System S822 Technical Overview and Introduction*, REDP-5102.

The DS8880 base frame contains two processor complexes. The based servers based on POWER8 feature up to four processor single chip modules (SCMs) in the DS8888. For the DS8886, one or two SCMs are used. The DS8884 features a single SCM.

In the DS8888, the SCMs are configured with 12 active cores. In the DS8886, the SCMs are configured with 8 or 12 cores. In the DS8884, the SCM is configured with six cores. Processor cores and system memory are configured according to the hardware that is installed in the storage system. Processors and memory can be upgraded concurrently as required to support storage system hardware upgrades. The supported maximum system hardware components depend on the total processor and system memory configuration.

Figure 2-9 shows the supported components for the DS8880 processor and memory options. NVS values are typically 1/16 of installed system memory.

Processor Cores	System Memory /NVS[1] (GB)[2]	Max I/O Enclosure Pairs	Max Host Adapters (HAs)	Max Device Adapter (DA) Pairs	Max Standard Drive Enclosure Pairs	Max 2.5 in SFF Drives	HPFE Gen2 Pairs	Max Flash Cards	Expansion Frames
DS8888 AF Configuration									
24-core	1024/64	2	16	0	0	0	4	192	0
48-core	2048/64[1]	4	32	0	0	0	8	384	1
DS8886 AF Configuration									
8-core	128/8	2	16	0	0	0	2	96	0
8-core	256/16	2	16	0	0	0	2	96	0
16-core	256/16	2	16	0	0	0	2	96	0
16-core	512/32	2	16	0	0	0	2	96	0
24-core	1024/64	2	16	0	0	0	2	96	0
24-core	2048/64[1]	2	16	0	0	0	2	96	0
DS8884 AF Configuration									
6-core	64/4	1	8	0	0	0	1	48	0
6-core	128/8	1	8	0	0	0	1	48	0
6-core	256/16	1	8	0	0	0	1	48	0
DS8886 Configuration - Three Phase									
8-core	128/8	4	16	2	2	96	2	96	0
8-core	256/16	4	16	2	2	96	2	96	0
16-core	256/16	4	32	8	30	1440	4	192	4
16-core	512/32	4	32	8	30	1440	4	192	4
24-core	1024/64	4	32	8	30	1440	4	192	4
24-core	2048/64[1]	4	32	8	30	1440	4	192	4
DS8886 Configuration - Single Phase									
8-core	128/8	4	16	3	3	144	2	96	0
8-core	256/16	4	16	3	3	144	2	96	0
16-core	256/16	4	32	8	32	1536	4	192	4
16-core	512/32	4	32	8	32	1536	4	192	4
24-core	1024/64	4	32	8	32	1536	4	192	4
24-core	2048/64[1]	4	32	8	32	1536	4	192	4
DS8884 Configuration									
6-core	64/4	2	8	2	6	192	1	48	0
6-core	128/8	4	16	4	18	768	2	96	2
6-core	256/16	4	16	4	18	768	2	96	2
Notes:									
1 - NVS for 2 TB total system memory configuration is 1/32 of system memory									
2 - System memory and NVS totals are aggregate of both CPCs									

Figure 2-9   Supported components for the DS8880 processor/memory options

2.3.2 Processor memory

The DS8888 and DS8886 configurations offer up to 2 TB of total system memory. The DS8884 configuration offers up to 256 GB total. Each processor complex contains half of the total system memory. All memory that is installed in each CPC is accessible to all processors in that CPC. The absolute addresses that are assigned to the memory are common across all processors in the CPC. The set of processors is referred to as a symmetric multiprocessor (SMP) system.

The POWER8 processor that is used in the DS8880 operates in simultaneous multithreading (SMT) mode, which executes multiple instruction streams in parallel. The number of simultaneous instruction streams varies according to processor and Licensed Internal Code (LIC) level. SMT mode enables the POWER8 processor to maximize the throughput of the processor cores by offering an increase in core efficiency.

The DS8880 configuration options are based on the total installed memory, which in turn corresponds with a predetermined number of installed and active processor cores.

The following DS8888 configuration upgrades can be performed nondisruptively:

► Processor upgrade from two processors/24 cores, to four processors/48 cores
► Processor memory upgrade, from 512 GB per server, to 1024 GB per server

The following DS8886 configuration upgrades can be performed nondisruptively:

► Processor configuration upgrade, with 8, 16, or 24 cores per server
► Processor memory upgrade, 64 - 1,024 GB per server

The following DS8884 configuration upgrade can be performed nondisruptively:

► Processor memory upgrade, 32 - 128 GB per server

Note: System memory and processor upgrades are tightly coupled. They cannot be ordered or installed independently from each other.

Caching is a fundamental technique for reducing I/O latency. Like other modern caches, the DS8880 contains volatile memory that is used as a read and write cache and nonvolatile memory that is used to maintain and back up a second copy of the write cache. If power is lost, the batteries keep the system running until all data in nonvolatile storage (NVS) is written to the CPC's internal disks.

The NVS scales to the processor memory that is installed, which also helps to optimize performance. NVS is typically 1/16 of installed CPC memory, with a minimum of 4 GB. The exception is that NVS is 1/32 of installed CPC memory on systems with 2 TB of total system memory.

The effectiveness of a read cache depends on the *read hit ratio*, which is the fraction of requests that are served from the cache without necessitating a read from the disk (read miss).

2.3.3 Flexible service processor

Each Power8 processor complex is managed by a service processor that is called a flexible service processor (FSP). The FSP is an embedded controller that is based on an IBM PowerPC® processor.

The FSP controls power and cooling for the CPC. The FSP performs predictive failure analysis for installed processor hardware, and performs recovery actions for processor or memory errors. The FSP monitors the operation of the firmware during the boot process, and can monitor the operating system for loss of control and take corrective actions.

2.3.4 Peripheral Component Interconnect Express adapters

Each DS8880 processor complex contains multiple PCIe adapters. These adapters allow point-to-point connectivity between CPCs, I/O enclosures, I/O adapters, and the HPFEs. Depending on the configuration, up to four PCIe adapters are in the DS8880 processor complex.

A DS8880 processor complex is equipped with the following PCIe adapters. These adapters provide connectivity between the CPCs and I/O enclosures:

► The DS8888 has four multi-port PCIe Gen3 adapters in slots 2, 3, 8, and 10.
► The DS8886 has four multi-port PCIe Gen3 adapters in slots 3, 5, 6, and 7.
► The DS8884 has two multi-port PCIe Gen3 adapters in slot 6 and 7.

Figure 2-10 shows the PCIe adapter locations in the DS8888 CPC.

Multi-port PCIe adapters
slots 2, 3, 8, and 10
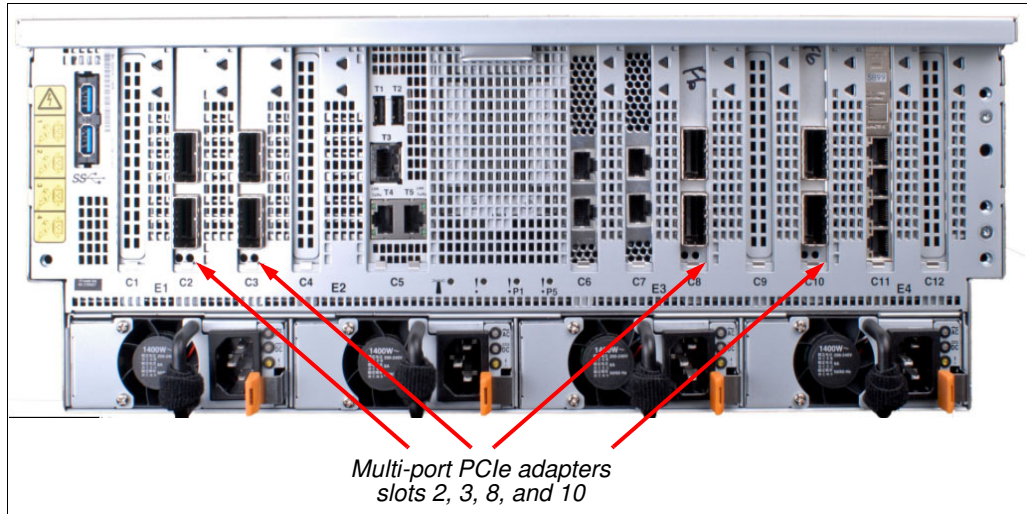
Figure 2-10   PCIe adapter locations in the DS8888 processor complex

Figure 2-11 shows the PCIe adapter locations in the DS8886 CPC.

Multi-port PCIe adapters
slots 3, 5, 6, and 7
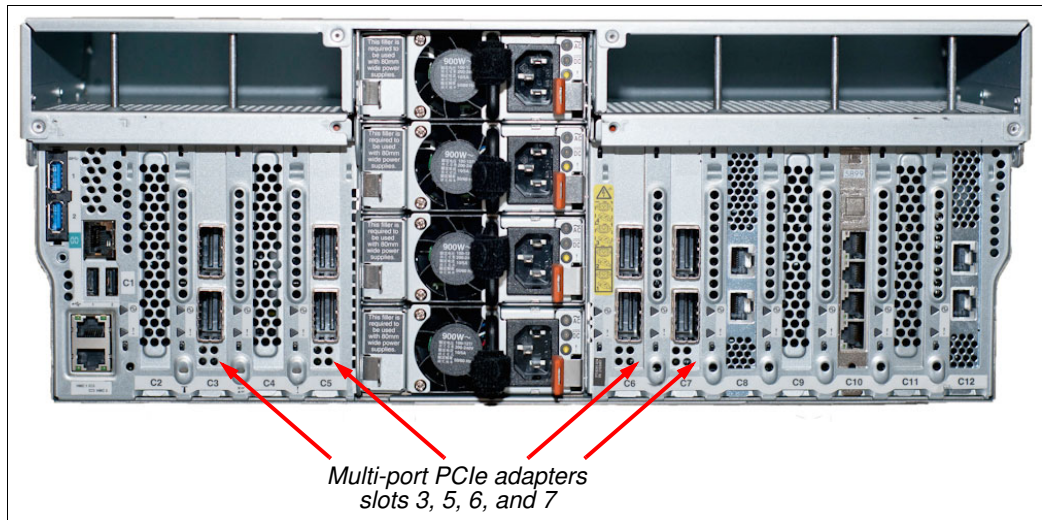
Figure 2-11   PCIe adapter locations in the DS8886 processor complex

Figure 2-12 shows the PCIe adapter locations in the DS8884 CPC.

Multi-port PCIe adapters
slots 6 and 7
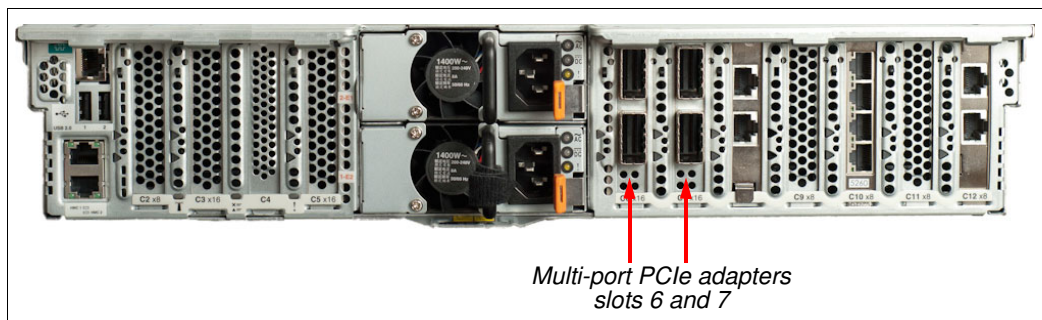
Figure 2-12   PCIe adapter locations in the DS8884 processor complex

2.4 I/O enclosures and adapters

The DS8880 base frame and first expansion frame (if installed) contain I/O enclosures, which are installed in pairs.

The DS8880 systems use an improved I/O enclosure, compared to earlier DS8000 systems. The enclosure is PCIe Gen3-capable, and it is attached to the CPCs with PCIe Gen3 cables. The internal transfer rate to each I/O bay is four times faster than DS8870, which uses PCIe Gen2. The new I/O enclosure now supports six PCIe adapter slots, and it has PCIe attachments that allow direct attachment to the High-Performance Flash Enclosures.

DS8888 and DS8886 CPCs have up to four 2-port PCIe adapters that provide connectivity to the I/O enclosures. The DS8884 CPCs have two 2-port PCIe adapters that provide connectivity.

Figure 2-13 on page 42 to Figure 2-15 on page 43 show the DS8880 CPC to I/O enclosure connectivity. The DS8886 requires a dual 8-core configuration for the first expansion frame, and the DS8884 requires a 6-core processor with 128 GB of total system memory for the first expansion frame.

One or two I/O enclosure pairs can be installed in the base frame of the DS8886. The DS8884 has a single I/O enclosure pair. Two I/O enclosure pairs are installed in the first expansion frame of a DS8886 when it is installed. The DS8884 expansion frame has a single I/O enclosure pair.

Each I/O enclosure can have up to four Host Adapters (HAs) and two Device Adapters (DAs). A maximum of 16 host adapter ports are supported in a single I/O enclosure. The I/O enclosure features integrated PCIe cable connections for direct attachment to the HPFE.

Each I/O enclosure includes the following attributes:

- ▶ Half-width 5U rack-mountable enclosure
- ▶ Six PCIe slots
- ▶ Two PCIe Gen3 connections to the HPFE
- ▶ Redundant power and cooling

Figure 2-13 shows the DS8888 CPC to I/O enclosure connectivity.
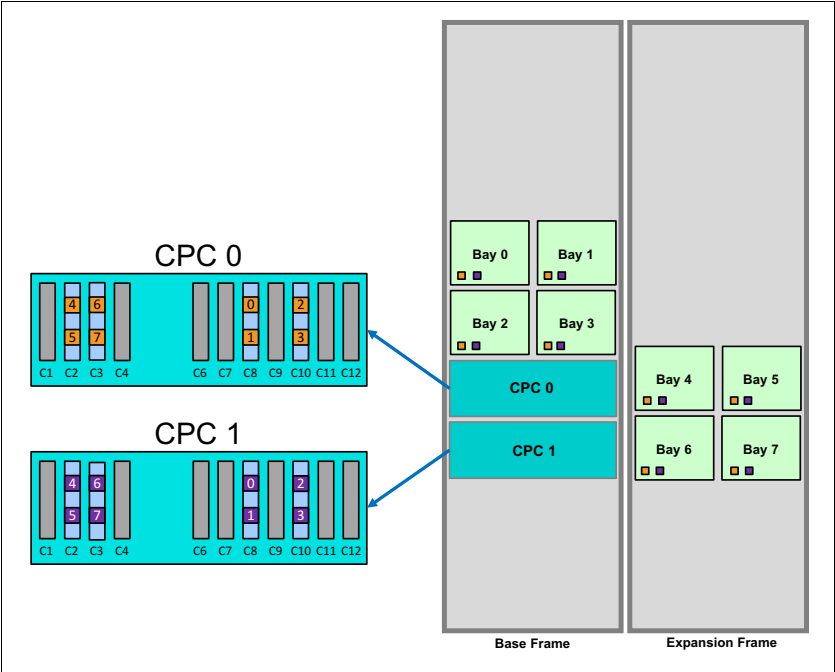
Figure 2-13   DS8888 I/O enclosure connections to the CPC

Figure 2-14 shows the DS8886 CPC to I/O enclosure connectivity.
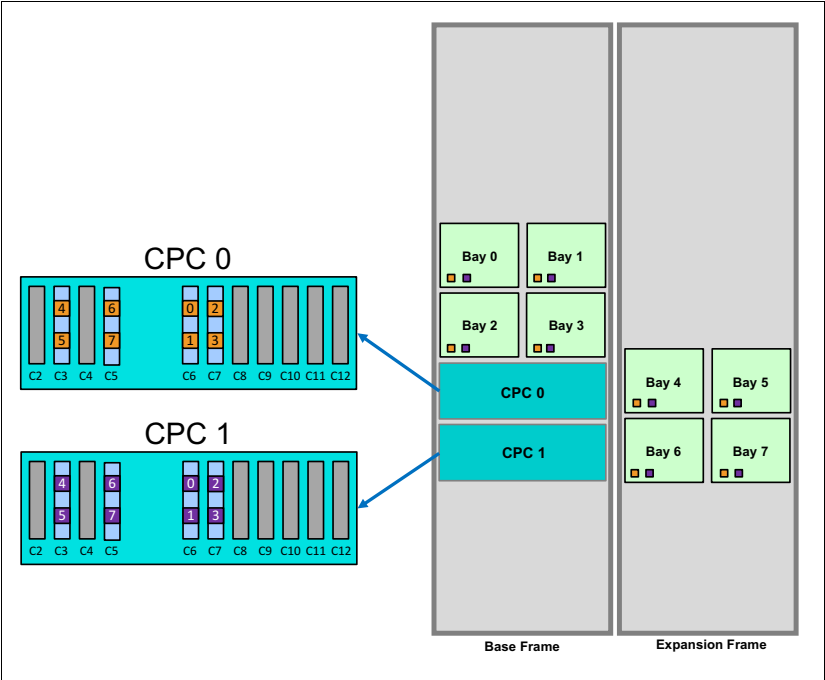
Figure 2-14   DS8886 I/O enclosure connections to the CPC

Figure 2-15 shows the DS8884 CPC to I/O enclosure connectivity.
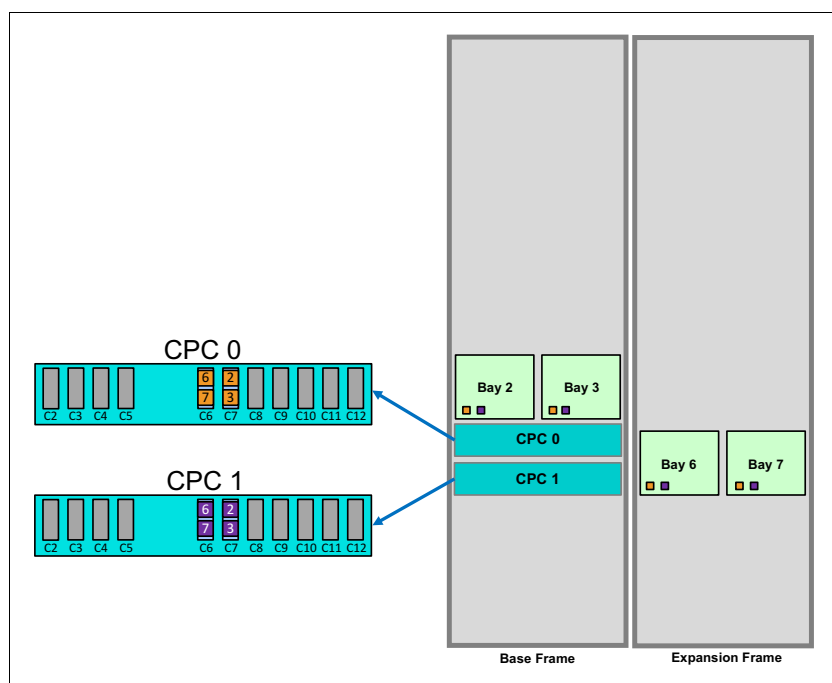
Figure 2-15   DS8884 I/O enclosure connections to the CPC

2.4.1 Cross-cluster communication

Figure 2-16 shows how the DS8880 I/O enclosure hardware is shared between the servers. One CPC is on the left side and one CPC is on the right side. The solid lines denote primary PCIe paths, and the dashed lines denote secondary PCIe paths.
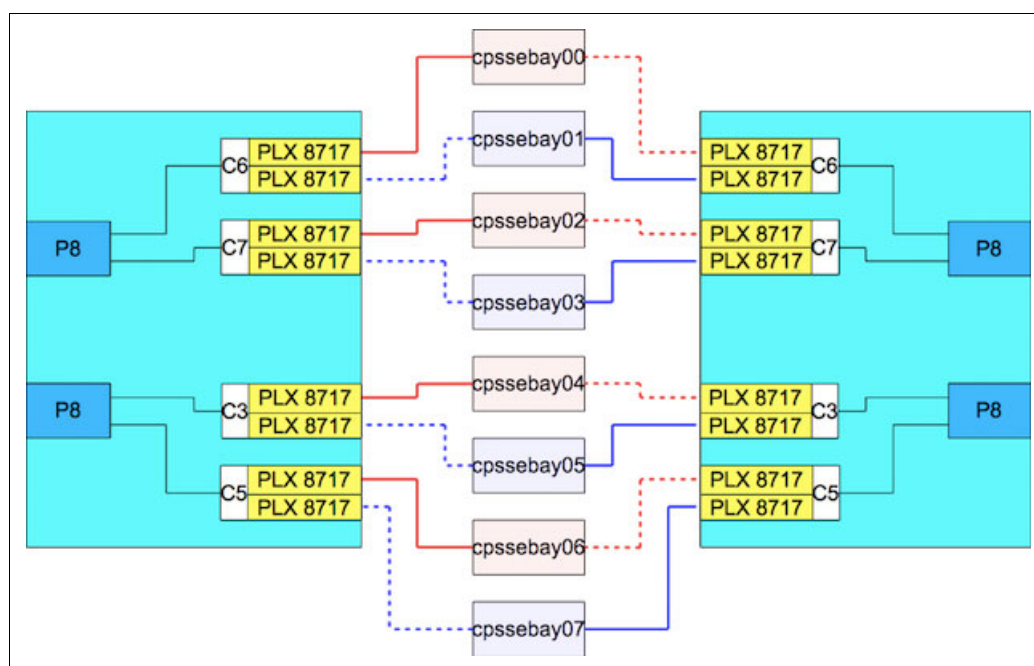
Figure 2-16   DS8880 series PCIe communications paths as configured on the DS8886

The DS8880 uses the PCIe paths through the I/O enclosures to provide high-speed communication paths between the CPCs. Normally, the lowest available even-numbered I/O enclosure is used for communication from server 0 to server 1, whereas the lowest available odd-numbered I/O enclosure is used for communication from server 1 to server 0.

If a failure occurs in one or more I/O enclosures, any of the remaining enclosures can be used to maintain communication between the servers.

2.4.2 I/O enclosure adapters

The DS8880 I/O enclosures provide the connectivity from the host systems to the storage arrays through the processor complexes. Each I/O adapter is optimized for its specific task in the DS8880.

Each I/O enclosure can contain up to four HAs that provide attachment to host systems and up to two DAs to provide attachment for standard drive enclosures. Each I/O enclosure has two PCIe x8 connectors on the I/O enclosure PCIe module that provide attachment for the HPFEs. In DS8888, two additional PCIe x8 redrive cards can be installed in the third and sixth card slots.

Figure 2-17 shows the DS8880 I/O enclosure adapter layout.
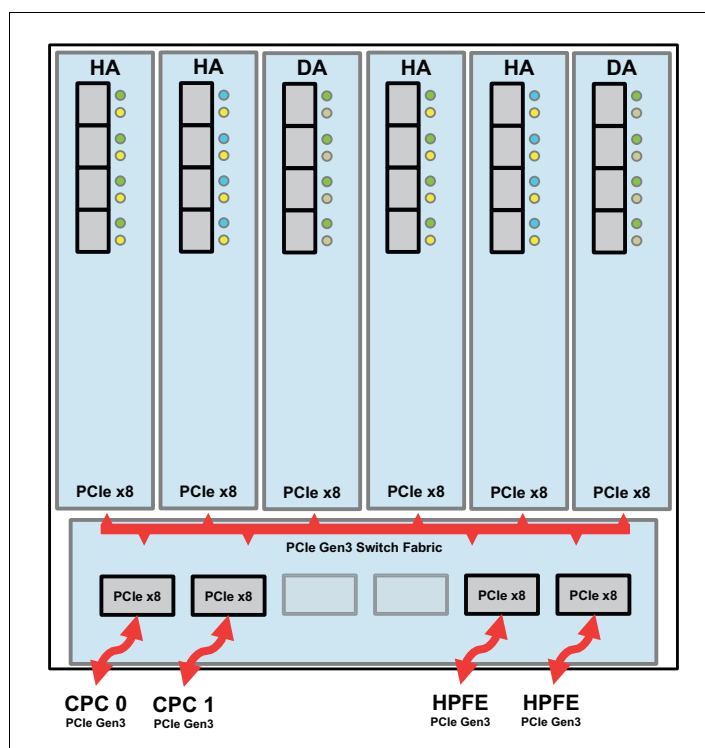
Figure 2-17   DS8880 I/O enclosure adapter layout

DS8880 host adapters

Attached host servers interact with software that is running on the processor complexes to access data that is stored on logical volumes. The servers manage all read and write requests to the logical volumes on the storage arrays.

Two different types of HAs are available (8 Gbps and 16 Gbps). Both can auto-negotiate their data transfer rate. 16 Gbps HAs can negotiate to 16, 8, or 4 Gbps full-duplex data transfer. The 8 Gbps HAs can negotiate to 8, 4, or 2 Gbps full-duplex. The 8 Gbps HAs are available in both 4-port or 8-port cards. The 16 Gbps HAs are 4-port cards only.

Figure 2-18 shows the 16 Gbps host adapter. It provides faster single stream and per port throughput, and reduces latency in comparison to the 8 Gbps adapter. The 16 Gbps HA is equipped with a quad-core PowerPC processor that delivers dramatic (two to three times) full adapter input/output operations per second (IOPS) improvements compared to the 8 Gbps adapter.
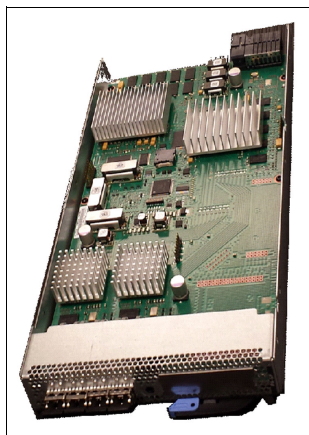
Figure 2-18   Sixteen Gbps host adapter

The 16 Gbps HA contains a new high-performance application-specific integrated circuit (ASIC). To ensure maximum data integrity, it supports metadata creation and checking. Each Fibre Channel port supports a maximum of 509 host login IDs and 1,280 paths. This configuration allows the creation of large storage area networks (SANs).

Note: The maximum number of host adapter ports in any single I/O enclosure is 16, regardless of the HA type. If you install 8-port HAs, they are installed in slots 1 and 4 of the I/O enclosure.

Each HA port can be configured as either FICON or FCP. For 8 Gbps HAs only, it is also possible to configure the port for the FC-AL protocol. For both HAs, the card optics can be either Long Wave or Short Wave.

The DS8888 and DS8886 configurations support a maximum of 16 HAs in the base frame, and an additional 16 HAs in the first expansion frame. The DS8884 configuration supports a maximum of eight in the base frame and an additional eight in the first expansion frame. With sixteen 8-port HAs, the maximum number is 128 HA ports. With sixteen 4-port HAs, the maximum number is 64 HA ports.

HAs are installed in slots 1, 2, 4, and 5 of the I/O enclosure. Figure 2-17 on page 44 shows the locations for four 16 Gbps HA cards in the DS8880 I/O enclosure. DS8880 supports intermix of both adapter types up to the maximum number of ports, as shown on Table 2-1.

Optimum availability: To obtain optimum availability and performance, one HA must be installed in each available I/O enclosure before a second HA is installed in the same enclosure.

Table 2-1   DS8886 port configurations

16 Gbps FC adapters	8 Gbps FC adapters	16 Gbps FC ports	8 Gbps FC ports (4-port/8-port)	Maximum ports
0	16	0	64 - 128	128
1	15	4	60 - 120	124
2	14	8	56 - 112	120
3	13	12	52 - 104	116
4	12	16	48 - 96	112
5	11	20	44 - 88	108
6	10	24	40 - 80	104
7	9	28	36 - 72	100
8	8	32	32 - 64	96
9	7	36	28 - 56	92
10	6	40	24 - 48	88
11	5	44	20 - 40	84
12	4	48	16 - 32	82
13	3	52	12 - 24	78
14	2	56	8 - 16	74
15	1	60	4 - 8	70
16	0	64	0	64

Figure 2-19 shows the preferred HA installation order for the DS8880. The HA locations and installation order for the four I/O enclosures in the base frame are the same for the I/O enclosures in the first expansion frame. If 16-port HAs are installed, they fill the first and second HA pair only, and leave the additional two slots unused.

HA Install Order for Two I/O Enclosures												
Host Bay	1B3						1B4					
Slot Location	1	2	3	4	5	6	1	2	3	4	5	6
HA Pair	2	4	DA	1	3	DA	1	3	DA	2	4	DA

HA Install Order for Four I/O Enclosures												
Host Bay	1B1						1B2					
Slot Location	1	2	3	4	5	6	1	2	3	4	5	6
HA Pair	4	8	DA	2	6	DA	2	6	DA	4	8	DA
Host Bay	1B3						1B4					
Slot Location	1	2	3	4	5	6	1	2	3	4	5	6
HA Pair	3	7	DA	1	5	DA	1	5	DA	3	7	DA

Figure 2-19   DS8880 HA installation order

Fibre Channel

The DS8880 uses the Fibre Channel protocol to transmit Small Computer System Interface (SCSI) traffic inside Fibre Channel frames. It also uses Fibre Channel to transmit FICON traffic, which uses Fibre Channel frames to carry z Systems I/O.

Each of the ports on a DS8880 HA can be independently configured for FCP or FICON. In addition, FC-AL can be configured on 8 Gbps ports only. The type of port can be changed through the DS8880 Storage Management GUI or by using Data Storage Command-Line Interface (DS CLI) commands. A port cannot be FICON and FCP simultaneously, but the protocol can be changed as required.

Fibre Channel-supported servers

The current list of servers that are supported by Fibre Channel attachment is available at this website:

http://www.ibm.com/systems/support/storage/config/ssic/index.jsp

Consult these documents regularly because they contain the current information about server attachment support.

Fibre Channel distances

The following types of HA cards are available:

► LW
► SW

All ports on each card must be LW or SW. The two types cannot be intermixed within an a single adapter card. With LW, you can connect nodes at distances of up to 10 km (6.2 miles) (non-repeated). With SW, you are limited to a distance that depends on the Fibre Channel cable type and the data transfer rate. OM3 Fibre Channel cable is required with 8 Gbps HAs and 16 Gbps HAs. For 16 Gbps, use OM3 or OM4. For the link distance limitations, see Table 2-2.

Table 2-2   SW link distance

Speed	OM3 link distance	OM4 link distance
8 Gbps FC	150 m (492 ft.)	190 m (623.3 ft.)
16 Gbps FC	100 m (328 ft.)	125 m (410.1 ft.)

Device adapters

DAs are Redundant Array of Independent Disks (RAID) controllers that provide access to the installed drives in the standard drive enclosures.

Each of the DS8880 DAs has four 8 Gbps Fibre Channel ports, which are connected to the drive enclosures by using two dual FC-AL loops, by using a switched topology. The DAs are installed in the I/O enclosures, and they are connected to the processor complexes through the PCIe network. The DAs are responsible for managing, monitoring, and rebuilding the RAID arrays. The DAs provide remarkable performance because of a high function and high-performance ASIC. To ensure maximum data integrity, the adapter supports metadata creation and checking.

Each DA connects the processor complexes to two separately switched FC-AL networks. Each network attaches to standard drive enclosures. Every drive in the enclosure is connected to both networks. Whenever the DA connects to a drive, it uses a bridged connection to transfer data. Figure 2-17 on page 44 shows the locations of the DAs in the I/O enclosure.

The DS8886 configuration supports 1 - 4 DA pairs in the base frame, and up to 4 additional DA pairs in the first expansion frame, for a total of up to 8 DA pairs. The DS8884 configuration supports half of the DS8886 configuration, or 1 - 2 DA pairs in the base frame, and up to 2 additional DA pairs in the first expansion frame, for a total of up to 4 DA pairs.

> **Note:** The all-flash models DS8888F, DS8886F and DS8884F do not support the installation of Device Adapter cards.

2.5 Storage enclosures and drives

This section covers the storage enclosures and drives, which consist of the following components:

► High Performance Flash Enclosure (HPFE) Gen2 Pairs, and Flash Cards
  – Flash Enclosure pairs connect to Flash RAID controllers using SAS cabling.
  – Flash RAID controllers connect to I/O enclosures using 8x PCIe Gen3 cabling.

► Standard Drive Enclosures, and rotational drives, commonly referred to as disk drive modules (DDMs)

  – Fibre Channel interface cards (FCICs) in the standard drive enclosures connect to Device Adapter cards using 8 Gbps Fibre. These connections create switched Fibre Channel arbitrated loop networks to the installed drives.

For more information about the storage subsystem, see 3.5, "RAS on the storage subsystem" on page 82.

2.5.1 Drive enclosures

The DS8880 has two types of drive enclosures:

► High-Performance Flash Enclosure Gen2
► Standard Drive Enclosure

High-performance flash enclosure Gen2 pairs

The DS8880 HPFE Gen2 consists of pair of high-speed 2U Flash Enclosures. One Gen2 enclosure is shown in Figure 2-20. HPFE Gen2 enclosures are always installed in pairs. Each enclosure pair supports 16, 32, or 48 flash cards.

Figure 2-20   High-Performance Flash Enclosure Gen2

Each HPFE Gen2 pair is connected to a redundant pair of Flash-optimized RAID controllers, which are installed in a separate pair of enclosures. These enclosures are called microbays.

Figure 2-21 shows an internal view of a microbay, showing the RAID controller on the left, PCIe bridge on the right, and dual cooling fans. Each microbay is powered by the I/O enclosure pair, through a redundant pair of Power Junction Assemblies (PJAs). Each PJA receives power from an I/O enclosure, and distributes power to both microbays in the pair.
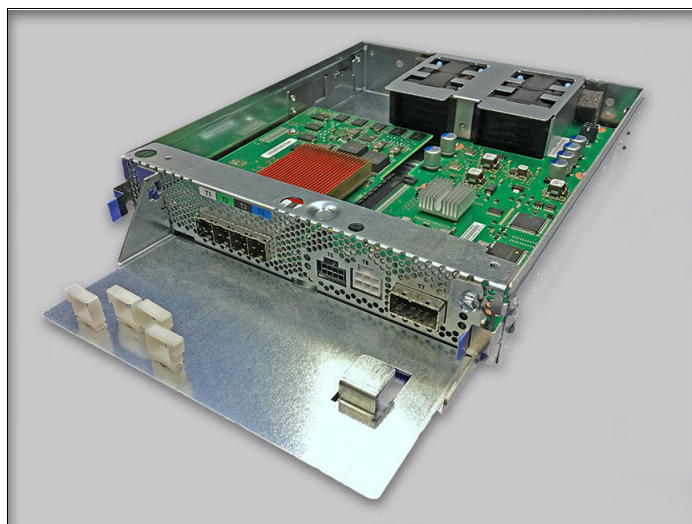
Figure 2-21   Internal view of a microbay

PCIe bus attachment

The HPFE Gen2 Flash RAID adapters are connected to the PCIe module in two different I/O enclosures to provide redundancy and workload balance. The connections are made by using 8-lane PCIe Gen3 cabling.

The x8 PCIe Gen3 connections provide up to 8 GBps, full duplex data transfer rates to the RAID adapters, with none of the protocol burden that is associated with Fibre Channel architecture. Each I/O enclosure pair supports up to two HPFE Gen2 pairs. The left side of Figure 2-22 shows a simplified view of the PCIe cabling topology. The right side shows the SAS cabling from the Flash RAID adapters to the Flash Enclosure pair.
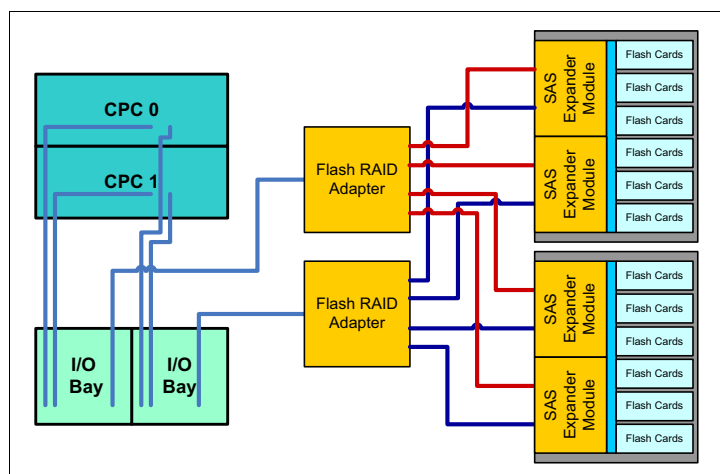
Figure 2-22   High-performance flash enclosure PCIe and SAS cabling

The DS8888 configuration can support up to four HPFE Gen2 pairs in the base frame, and up to four HPFE Gen2 pairs in the optional expansion frame, for a total of eight HPFE Gen2 pairs, with a maximum 384 flash cards.

The DS8886 configuration can support up to two HPFE Gen2 pairs in the base frame, and up to two HPFE Gen2 pairs in the first expansion frame, for a total of four HPFE Gen2 pairs, with a maximum 192 flash cards.

The DS8884 configuration can support one HPFE Gen2 pair in the base frame, one HPFE Gen2 pair in the first expansion frame, for a total of two HPFE Gen2 pairs, with a maximum 96 flash cards.

To learn more about the HPFE Gen2, see *DS8880 High Performance Flash Enclosure Gen2*, REDP-5422.

High-performance flash cards
The DS8880 HPFE Gen2 supports 2.5-inch high-performance flash cards (Table 2-3).

Each HPFE Gen2 pair can contain 16, 32, or 48 flash cards. Flash cards can have 400 GB, 800 GB, 1.6 TB, or 3.2 TB capacity. All flash cards in a flash enclosure must be the same type and same capacity.

All high-performance flash cards are Full Disk Encryption (FDE) capable. For more information about licensed features, see Chapter 9, "IBM DS8880 features and licensed functions" on page 219.

Table 2-3   Supported high-performance flash cards

Feature code	Drive capacity	Drive type	Drive speed in RPM (K=1000)	RAID support (Default RAID-6)
1610	400 GB	2.5-inch flash card	N/A	5, 6, 10[a]
1611	800 GB	2.5-inch flash card	N/A	5, 6, 10[a]
1612	1.6 TB	2.5-inch flash card	N/A	6, 10[b]
1613	3.2 TB	2.5-inch flash card	N/A	6, 10[b]

a. RAID-5 is supported, but not recommended
b. RAID-5 is not supported

Note: To learn more about the DS8880 drive features, see the *IBM System Storage DS8880 Introduction and Planning Guide*, GC27-8525.

Arrays and spares
Each HPFE Gen2 pair can contain up to six array sites. The first set of 16 flash cards creates two 8-flash-card array sites. RAID 6 arrays will be created by default on each array site. RAID 5 is optional for flash cards smaller than 1 TB, but is not advised. RAID 10 is optional for all flash card sizes.

During logical configuration, RAID 6 arrays and the required number of spares are created. Each HPFE Gen2 pair will have two global spares, created from the first increment of 16 Flash Cards. The first two arrays to be created from these array sites are 5+P+Q. Subsequent RAID 6 arrays in the same HPFE Gen2 Pair will be 6+P+Q.

For more information about DS8880 virtualization and configuration, see Chapter 4, "Virtualization concepts" on page 97.

2.5.2 Standard drive enclosures

The DS8880 traditional spinning drives, also known as DDMs, and flash drives, also known as solid-state drives (SSDs), are installed in standard drive enclosures. Standard drive enclosures are installed in pairs.

A standard drive enclosure pair supports either twenty-four 2.5-inch small form factor (SFF) or twelve 3.5-inch large form factor (LFF) drives. Each drive is an industry-standard serial-attached SCSI (SAS) drive.

Flash drives can be installed in standard drive enclosure pairs in drive sets of 8 or 16, depending on the capacity of the drive. The 2.5-inch SFF drives are installed in drive sets of 16. The 3.5-inch LFF drives are in installed in drive sets of eight.

Half of each drive set is installed in each enclosure of the enclosure pair.

> **Note:** If the standard drive enclosure pair is not fully populated, you must install fillers in the empty slots.

Each drive connects into the enclosure midplane. The midplane provides physical connectivity for the drives, FCICs, and power supply units (PSUs).

Each drive enclosure has a redundant pair of FCICs with two inbound and two outbound 8 Gbps Fibre Channel (FC) connections. The FCICs provide the Fibre Channel to SAS conversion and interconnection logic for the drives. In addition, the FCICs include SCSI Enclosure Services (SES) processors that provide all enclosure services.

Figure 2-23 shows the front and rear views of the standard drive enclosure.
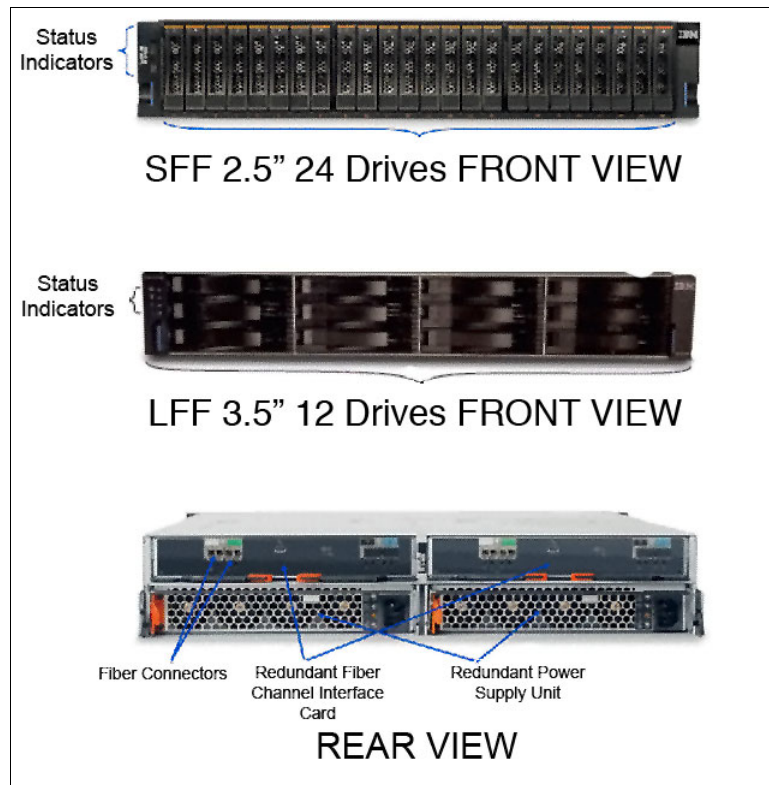
Figure 2-23   DS8880 standard drive enclosures

Switched Fibre Channel Arbitrated Loop

The DS8880 uses switched FC-AL technology to link the Fibre Channel DA pairs to the standard drive enclosures. Switched FC-AL uses the standard FC-AL protocol. However, the physical implementation is different. Switched FC-AL technology includes the following key features:

► Standard FC-AL communication protocol from the DA to drives.

► Direct point-to-point connections are established between the DA and drives.

► Isolation capabilities to provide easy problem determination during drive failures.

► Predictive failure statistics.

► Simplified expansion, where no cable rerouting is required when additional drive enclosure pairs are installed to the Fibre Channel network. The DS8880 architecture uses dual redundant switched FC-AL access to each of the drive enclosures. This configuration features the following key benefits:

– Two independent switched FC-AL networks provide high-performance connections to the drives.

– Four access paths are available to each drive.

– Each DA port operates independently.

– This configuration doubles the bandwidth over traditional FC-AL loop implementations.

Figure 2-24 shows each drive attached to two separate FCICs with connections to the drives. By using two DAs, redundant data paths exist to each drive. Each DA can support two switched Fibre Channel networks.
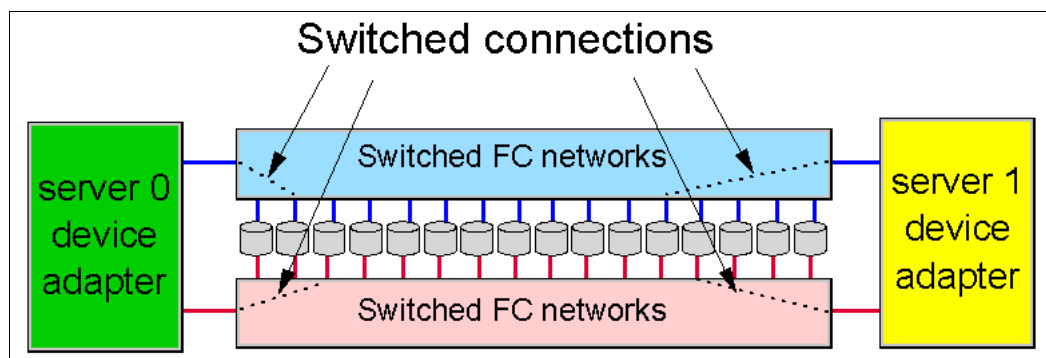
Figure 2-24   DS8880 drive enclosure (only 16 drives are shown for simplicity)

Arrays across loops

For the DS8880, the upper enclosure connects to one dual loop and the lower enclosure connects the other dual loop, in a drive enclosure pair.

Each enclosure places two FC switches onto each dual loop. Drives are installed in groups of 16. Half of the new drives are installed in one drive enclosure, and the other half is placed into the other drive enclosure of the pair.

A standard drive enclosure array site consists of eight drives: Four from each enclosure of the pair. When a RAID array is created from the drives of an array site, half of the array is in each storage enclosure. The performance is increased because the bandwidth is aggregated across the hardware of both enclosures.

One storage enclosure of the pair is on one FC switched loop, and the other storage enclosure of the pair is on a second switched loop. This configuration splits the array across

two loops, which is known as array across loops (AAL), as shown in Figure 2-26. Only 16 drives are shown, with eight in each drive enclosure. When fully populated, 24 drives are in each enclosure.

Figure 2-25 shows the DA pair layout. One DA pair creates two dual switched loops.
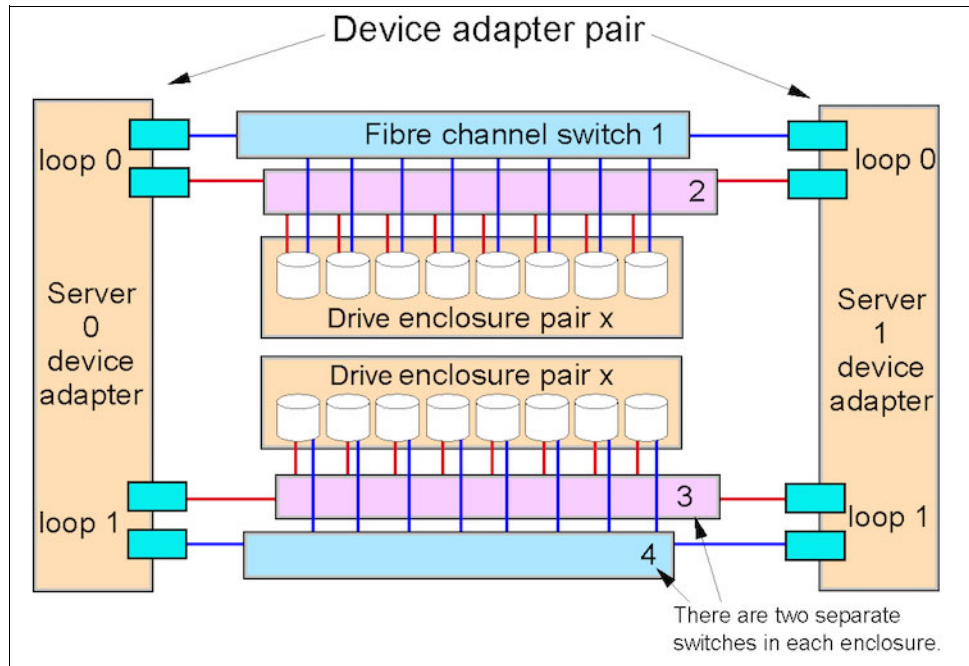
Figure 2-25   DS8880 switched loop layout (only eight drives per enclosure are shown for simplicity)

Figure 2-26 shows the layout of the array sites. Array site 1 in green (the darker drives) consists of the four left drives in each enclosure. Array site 2 in yellow (the lighter disks) consists of the four right drives in each enclosure. When an array is created, the array is accessible across both dual loops.
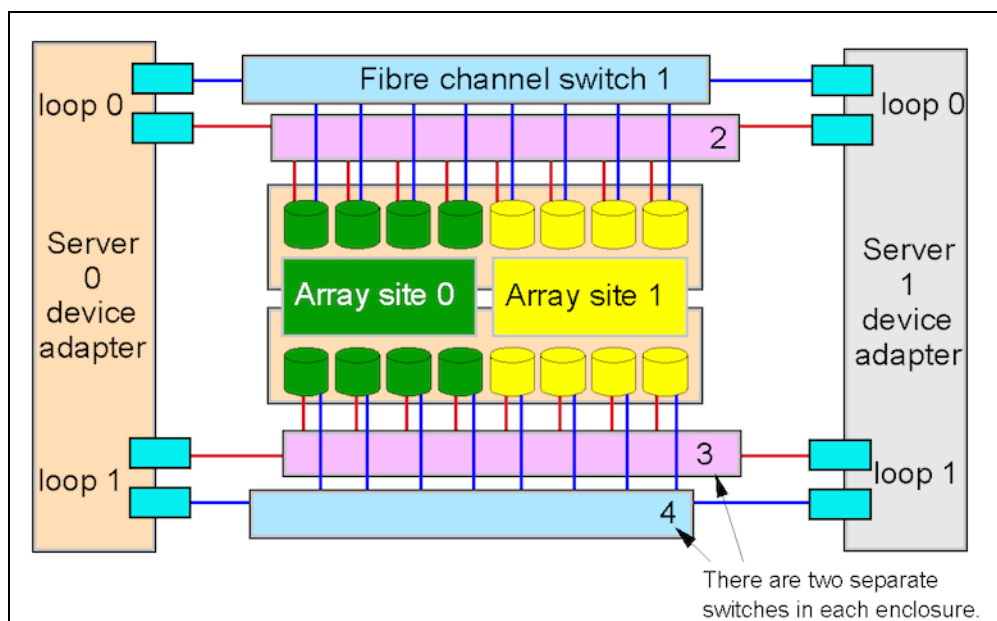
Figure 2-26   Array across loop

Benefits of AAL

AAL is implemented to increase performance. When the DA writes a stripe of data to a RAID array, it sends half of the write to each switched loop. By splitting the workload in this manner, the workload is balanced across the loops. This configuration aggregates the bandwidth of the two loops and improves performance. If RAID 10 is used, two RAID 0 arrays are created. Each loop hosts one RAID 0 array. When the read I/O is serviced, half of the reads can be sent to each loop, which improves performance by balancing the workload across the loops.

Expansion

Device adapters (DAs) are installed in the I/O enclosures in pairs. Each DA of the pair is in a separate I/O enclosure of the I/O enclosure pair. The DS8886 configuration can support up to four DA pairs in the base frame, and four DA pairs in the first expansion frame for a total of eight DA pairs. The DS8884 configuration can support up to two DA pairs in the base frame, and two DA pairs in the first expansion frame, for a total of four DA pairs.

Standard drive enclosure pairs are connected to a DA pair. Each DA pair can support up to four drive enclosure pairs. If more storage capacity is required, an additional DA pair can be installed (up to the maximum supported quantity), and then additional standard drive enclosure pairs can be installed (up to the maximum supported quantity).

Drive sets can also be added to standard drive enclosures that are not fully populated. Half of the drive set is added to each enclosure of the drive enclosure pair. Performance will be superior if drive enclosure pairs are distributed across more DA pairs, aggregating the bandwidth of the installed hardware.

For more information about DA pairs and standard drive enclosure distribution and cabling, see 2.2.1, "DS8888 All-Flash configuration" on page 26 and 2.2.3, "DS8884 configuration" on page 29.

DS8880 standard and flash drives

The DS8880 supports the following drive types (Table 2-4 and Table 2-5 on page 56):

► 2.5-inch High-performance flash cards (see "High-performance flash cards" on page 51)
► 2.5-inch SAS flash drives, which are also known as SSDs
► 2.5-inch SAS enterprise drives (15K or 10K revolutions per minute (RPM))
► 3.5-inch SAS nearline drives (7200 RPM)

Table 2-4 shows the flash drive types supported by DS8880.

Table 2-4   DS8880 supported flash drive types

Feature code	Drive capacity	Drive type	Drive speed in RPM (K=1000)	RAID support (Default RAID-6)	Drives for each set
6158	400 GB	Flash drive (SSD)	N/A	5, 6, 10[a]	16
6258	800 GB	Flash drive (SSD)	N/A	5, 6, 10[a]	16
6358	1.6 TB	Flash drive (SSD)	N/A	6, 10[b]	16

a. RAID 5 is supported, but not recommended
b. RAID 5 is not supported

Table 2-5 shows the enterprise and nearline drive types supported by DS8880.

Table 2-5 DS8880 supported enterprise and nearline drive types

Feature code	Drive capacity	Drive type	Drive speed in RPM (K=1000)	RAID support (Default RAID-6)	Drives for each set
5308	300 GB	2.5-inch SAS Ent	15K	5, 6, 10[a]	16
5618	600 GB	2.5-inch SAS Ent	15K	5, 6, 10[a]	16
5708	600 GB	2.5-inch SAS Ent	10K	5, 6, 10[a]	16
5768	1.2 TB	2.5-inch SAS Ent	10K	6, 10[b]	16
5778	1.8 TB	2.5-inch SAS Ent	10K	6, 10[b]	16
5868	4 TB	3.5-inch SAS NL	7.2K	6, 10[b]	8
5878	6 TB	3.5-inch SAS NL	7.2K	6, 10[b]	8

a. RAID 5 is supported, but not recommended
b. RAID 5 is not supported

Arrays and spares

During logical configuration, arrays and spares are created from groups of eight drives, which are called *array sites*. The required number of spares for each DA pair is four drives of the same capacity and speed. For example, the first four RAID 6 arrays that are created are 5+P+Q+S. Additional RAID 6 arrays (of the same capacity and speed) that are configured on the DA pair are 6+P+Q.

The next group of 16 drives creates two 8-drive array sites. These next two arrays to be created from these array sites are 6+P+Q. For more information about DS8880 virtualization and configuration, see Chapter 4, "Virtualization concepts" on page 97.

2.6 Power and cooling

The DS8880 power and cooling system is highly redundant. The components are described in this section. For more information, see 3.6, "RAS on the power subsystem" on page 90.

2.6.1 Rack Power Control

The DS8880 features a pair of redundant rack power control (RPC) cards, which are used to control the power sequencing of the system. As in earlier DS8000 models, the DS8880 RPCs are connected to the FSP in each processor complex, which allows them to communicate to the HMC and the storage system. The DS8880 RPC cards also add a second communication path to each of the processor complex operating partitions. The DS8880 RPCs also feature dedicated communication paths to each DC-UPS.

2.6.2 Direct current uninterruptible power supply

Each DS8880 frame contains two DC-UPSs. Depending on model configuration, these can be either single-phase 8 kW 8U DC-UPSs, or three-phase, 12 kW 12U DC-UPSs. The DC-UPSs cannot be intermixed in any DS8880 storage system.

The DC-UPS provides rectified AC power distribution and power switching for redundancy. Two redundant DC-UPSs are in each frame of the DS8880. Each DC-UPS features internal fans to supply cooling for that power supply.

The frame features two AC power cords. Each cord feeds a single DC-UPS. The DC-UPS distributes rectified AC. If AC is not present at the input line, the partner DC-UPS can provide rectified AC to the DC-UPS that lost input power, with no reduction in DC-UPS redundancy. If no AC input is present for either DC-UPS in the frame, the DC-UPSs switch to battery power. Depending on whether the ePLD feature is installed on the system, the system runs on battery power for either 4 or 40 seconds before it initiates an orderly shutdown.

2.6.3 Battery service module set

The DC-UPS contains integrated battery sets that are known as BSM sets. The BSM set consists of four BSM modules. The BSM sets help protect data if external power to the frame is lost. If AC input power to the frame is lost, the batteries are used to maintain power to the processor complexes and I/O enclosures for sufficient time to allow the contents of NVS memory (modified data that is not yet destaged from cache) to be written to the hard disk drives that are internal to the processor complexes (not the storage enclosure drives).

The BSMs sets consist of two BSM types:

► Each BSM set contains one primary BSM. The primary BSM is the only BSM with an electrical connection to the DC-UPS.

► Each BSM set also contains three secondary BSMs that are cabled to the primary BSM.

2.6.4 Extended Power Line Disturbance feature

The duration that the DC-UPSs can run on battery before system shutdown is initiated depends on whether the ePLD feature is installed.

The optional ePLD feature allows the system to run for up to 40 seconds without line input power and then gracefully shuts down the system if power is not restored. If the ePLD feature is not installed, the system initiates shutdown after 4 seconds if frame power is not restored. For more information about why this feature might be necessary, see 3.6.3, "Line power fluctuation" on page 93.

2.6.5 Power cord options

The power cord must be ordered specifically for the input voltage to meet certain requirements. The power cord connector requirements vary widely throughout the world. The power cord might not include the suitable connector for the country in which the system is installed. In this case, the connector must be replaced by an electrician after the system is delivered. For more information, see the *IBM System Storage DS8880 Introduction and Planning Guide*, GC27-8525.

2.6.6 Power distribution

In each frame, the two DC-UPSs supply output power to four Power Distribution Units (PDUs).

In the base frame, the PDUs supply power to the processor complexes, the I/O enclosures, the standard drive enclosures, and the HPFEs. In the first expansion frame, the PDUs supply power to the I/O enclosures, standard drive enclosures, and the HPFEs. In the second and third expansion frames, the PDUs supply power only to the standard drive enclosures. No HPFEs, I/O enclosures, or processor complexes are in these frames.

2.6.7 Enclosure power supplies

The CPCs, I/O enclosures, and standard drive enclosures feature dual redundant PSUs for each enclosure. The PSUs are supplied power from the DC-UPS through the PDUs. Each PSU in an enclosure is supplied from different DC-UPSs for redundancy. The PSUs have their own internal cooling fans. Each enclosure also has its own redundant cooling fans. All cooling fans draw cool air from the front of the frame and exhaust hot air to the rear of the frame.

> **Note:** The DS8880 is designed for efficient air flow and to be compliant with *hot and cold aisle* data center configurations.

2.6.8 Power junction assembly

The PJA is a component of the DS8880 power subsystem that combines and distributes power from multiple sources. Dual PJAs provide redundant power to the Management Console (HMC) and the Ethernet switches. Additional redundant PJAs distribute power from the I/O enclosures to the High-Performance Flash Enclosure Gen2 microbays in systems that contain HPFE Gen2 pairs.

2.7 Management Console and network

All configuration base frames ship with one Management Console, which is also known as the HMC, and two private network Ethernet switches. An optional second Management Console is available as a redundant point of management, and is also installed in the base frame. The introduction of the 19-inch rack introduced a new HMC, which is an SFF mini PC HMC.

Figure 2-27 shows a diagram of the new mini PC HMC and its location in the base frame. The mini PC HMC resides in the bottom of the base frame. As an option, a redundant HMC is available. This secondary HMC no longer needs to be added to a customer rack. This secondary HMC sits next to the primary HMC.
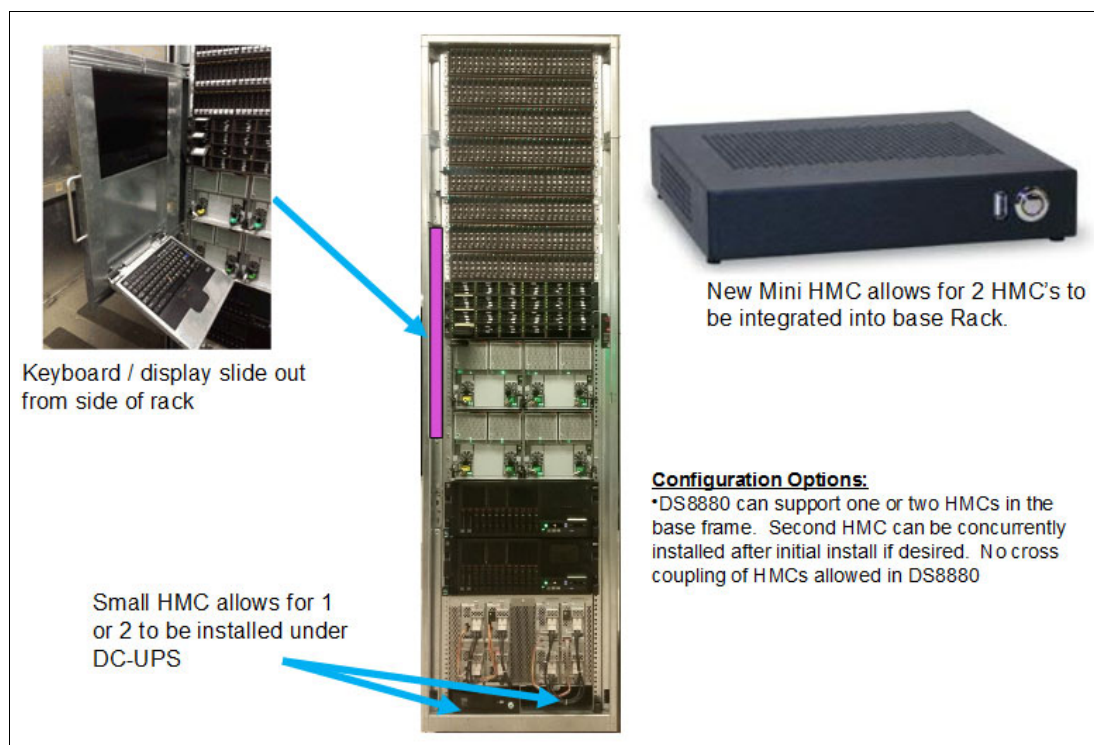
Figure 2-27   Diagram of new mini PC HMC and location in the base frame

The storage administrator runs all DS8880 logical configuration tasks by using the Storage Management GUI or DS CLI. All client communications to the storage system are through the HMC.

Clients that use the DS8880 advanced functions, such Metro Mirror or FlashCopy, for example, communicate to the storage system with Copy Services Manager (CSM). CSM is built into Spectrum Control Standard and Advanced. It replaces the IBM Tivoli Productivity Center for Replication through the Management Console.

The Management Console provides connectivity between the storage system and Encryption Key Manager servers.

The Management Console also provides the functions for remote call-home and remote support connectivity.

For more information about the HMC, see Chapter 8, "IBM DS8880 Management Console planning and setup" on page 203.

2.7.1 Ethernet switches

The DS8880 base frame has two 8-port Ethernet switches. The two switches provide two redundant private management networks. Each CPC includes connections to each switch to allow each server to access both private networks. These networks cannot be accessed externally, and no external connections are allowed. External client network connection to the DS8880 system is through a dedicated connection to the Management Console. The switches receive power from the PJAs and do not require separate power outlets. The ports on these switches are shown in Figure 2-28.
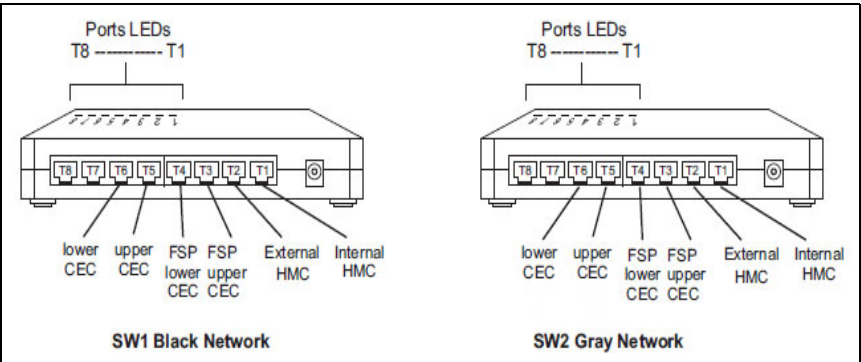
Figure 2-28   Ethernet switch ports

> **Important:** The internal Ethernet switches that are shown in Figure 2-28 are for the DS8880 private network only. No client network connection must ever be made directly to these internal switches.

2.7.2 DS8880 operator panel

The DS8880 status indicators are on the front door. Figure 2-29 shows the operator panel for a DS8880 storage system.

Index	LED Function
1	Line Cord 1
2	Line Cord 2
3	Identify or System Attention
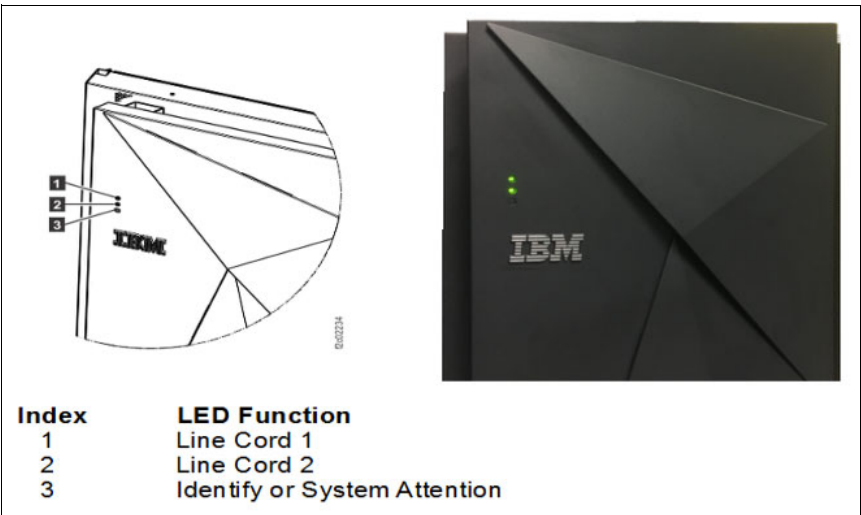
Figure 2-29   DS8880 frame operator indicators

Each panel has two power cord indicators, one for each power cord (frame input power). For normal operation, both of these indicators are illuminated green if each power cord supplies the correct power to the frame. Another LED is below the line indicators with a normal state of off. If this LED is lit solid yellow (in only the base frame), an open event exists in the problem log, and service is required. If this LED flashes (in any frame), the frame is being serviced.

The DS8880 has no physical power on/off switch because power sequencing is managed through the HMC. This configuration ensures that all data in nonvolatile storage, which is known as *modified data,* is destaged correctly to the drives before the DS8880 powers down.

2.8 Hardware feature summary for DS8880 configurations

The following tables summarize the hardware features that are available for the different DS8880 configurations.

The DS8888 is a high-performance All-Flash storage system that uses only High-Performance Flash Enclosures. Table 2-6 shows the hardware features of the DS8888.

Table 2-6   Hardware features of the DS8888

Processor cores	System memory[a]	Maximum I/O enclosure pairs	Host adapters[b]	Device adapter pairs	HPFE Gen2 Pairs	Maximum standard drive enclosure pairs	Expansion frames
24-core	1024 GB	4	2 - 16	0	1 - 4	0	0
48-core	2048 GB	8	2 - 32	0	1 - 8	0	0 - 1

a. System memory total for both CPCs.
b. Host adapters can be four ports or eight ports, with a maximum of 16 ports for each I/O enclosure.

Table 2-7 shows the maximum capacities for the DS8888 configuration

Table 2-7   Maximum capacity for the DS8888 configuration

Processor cores	System memory[a]	Maximum quantity of 2.5-inch drives	Maximum usable capacity of 2.5-inch drives	Maximum quantity of 3.5-inch drives	Maximum usable capacity of 3.5-inch drives	Maximum quantity of 2.5-inch flash cards	Maximum usable capacity of 2.5-inch flash cards[b]	Maximum total number of drives
24-core	1024 GB	0	0	0	0	192	392 TB	n/a
48-core	2048 GB	0	0	0	0	384	784 TB	n/a

a. System memory for each CPC.
b. Usable capacity is calculated by using RAID 6 for 2.5-inch flash cards.

The DS8886 is a high-density, high-performance storage system that includes standard disk enclosures and High-Performance Flash Enclosures. Table 2-8 shows the hardware features for the DS8886 configuration.

Table 2-8   Hardware features of the DS8886

Processor cores	System memory[a]	Maximum I/O enclosure pairs	Host adapters[b]	Device adapter pairs	HPFE Gen2 Pairs	Maximum standard drive enclosure pairs	Expansion frames
8-core	128 GB	4	2 - 16	1 - 3	0 - 2	3	0
8-core	256 GB	4	2 - 16	1 - 3	0 - 2	3	0
16-core	256 GB	8	2 - 32	1 - 8	0 - 4	32	0 - 4
16-core	512 GB	8	2 - 32	1 - 8	0 - 4	32	0 - 4
24-core	1024 GB	8	2 - 32	1 - 8	0 - 4	32	0 - 4
24-core	2048 GB	8	2 - 32	1 - 8	0 - 4	32	0 - 4

a. System memory for each CPC.
b. Host adapters can be four ports or eight ports, with a maximum of 16 ports for each I/O enclosure.

Table 2-9 shows the maximum capacities for the Enterprise Class configuration.

Table 2-9   Maximum capacity for the DS8886 configuration

Processor cores	System memory[a]	Maximum quantity of 2.5-inch drives	Maximum usable capacity of 2.5-inch drives[b]	Maximum quantity of 3.5-inch drives	Maximum usable capacity of 3.5-inch drives[c]	Maximum quantity of 2.5-inch flash cards	Maximum usable capacity of 2.5-inch flash cards[d]	Maximum total number of drives[e]
8-core	128 GB	144	138.3 TB	72	288 TB	96	196 TB	288
8-core	256 GB	144	138.3 TB	72	288 TB	96	196 TB	288
16-core	256 GB	1536	1.61 PB	768	3.04 PB	192	392 TB	1728
16-core	512 GB	1536	1.61 PB	768	3.04 PB	192	392 TB	1728
24-core	1024 GB	1536	1.61 PB	768	3.04 PB	192	392 TB	1728
24-core	2048 GB	1536	1.61 PB	768	3.04 PB	192	392 TB	1728

a. System memory total for both CPCs.
b. Usable capacity is calculated by using RAID 5 for 2.5-inch drives.
c. Usable capacity is calculated by using RAID 6 for 3.5-inch nearline (NL) drives.
d. Usable capacity is calculated by using RAID 6 for 2.5-inch flash cards.
e. Combined total of 2.5-inch disk drives and 2.5-inch flash cards.

The DS8884 is an entry-level, high-performance storage system that includes standard disk enclosures and high-performance flash enclosures.

Table 2-10 shows the hardware features for the DS8884.

Table 2-10   Hardware features for the Business Class configuration

Processor cores	System memory[a]	Maximum I/O enclosure pairs	Host adapters[b]	Device adapter pairs	HPFE Gen2 Pairs	Maximum standard drive enclosure pairs	Expansion frames
6-core	64 GB	1	2 - 8	1	0 - 1	4	0
6-core	128 GB	2	2 - 16	1 - 4	0 - 2	18	0 - 2
6-core	256 GB	2	2 - 16	1 - 4	0 - 2	18	0 - 2

a. System memory for each CPC.
b. Host adapters can be four ports or eight ports, with a maximum of 16 ports for each I/O enclosure.

Table 2-11 shows the maximum capacities for the DS8884 configuration.

Table 2-11   Maximum capacities for the DS8884 configuration

Processor cores	System memory[a]	Maximum quantity of 2.5-inch drives	Maximum usable capacity of 2.5-inch drives[b]	Maximum quantity of 3.5-inch drives	Maximum usable capacity of 3.5-inch drives[c]	Maximum quantity of 2.5-inch flash cards	Maximum usable capacity of 2.5-inch flash cards[d]	Maximum total drives[e]
6-core	64 GB	192	201.8 TB	120	366.4 TB	48	98 TB	252
6-core	128 GB	768	1.12 PB	384	1.52 PB	96	196 TB	864
6-core	256 GB	768	1.12 PB	384	1.52 PB	96	196 TB	864

a. System memory total for both CPCs.
b. Usable capacity is calculated by using RAID 6 for 2.5-inch drives.
c. Usable capacity is calculated by using RAID 6 for 3.5-inch NL drives.
d. Usable capacity is calculated by using RAID 6 for 2.5-inch flash cards.
e. Combined total of 2.5-inch disk drives and 2.5-inch flash cards.

3

Reliability, availability, and serviceability on the IBM DS8880

This chapter describes the reliability, availability, and serviceability (RAS) characteristics of the IBM DS8880.

This chapter covers the following topics:

- ▶ DS8880 processor complex features
- ▶ CPC failover and failback
- ▶ Data flow in the DS8880
- ▶ RAS on the Hardware Management Console
- ▶ RAS on the storage subsystem
- ▶ RAS on the power subsystem
- ▶ Other features

3.1 DS8880 processor complex features

Reliability, availability, and serviceability are important concepts in the design of the IBM DS8880. Hardware features, software features, design considerations, and operational guidelines all contribute to make the DS8880 reliable. At the heart of the DS8880 is a pair of POWER8 processor-based servers. These servers, which are known as central processor complexes (CPCs), share the load of receiving and moving data between the attached hosts and the storage arrays. However, they are also redundant so that if either CPC fails, the system switches to the remaining CPC and continues to run without any host I/O interruption. This section looks at the RAS features of the CPCs, including the hardware, the operating system, and the interconnections.

3.1.1 POWER8 Hypervisor

The POWER8 Hypervisor (PHYP) is a component of system firmware that is always active, regardless of the system configuration, even when it is disconnected from the Management Console. PHYP runs on the flexible service processor (FSP). It requires the FSP processor and memory to support the resource assignments to the logical partition (LPAR) on the server. It operates as a hidden partition with no CPC processor resources that are assigned to it, but it does allocate a small amount of memory from the partition.

The PHYP provides the following capabilities:

► Reserved memory partitions set aside a portion of memory to use as cache and a portion to use as nonvolatile storage (NVS).

► Preserved memory support allows the contents of the NVS and cache areas to be protected if a server restarts.

► I/O enclosure initialization, power control, and slot power control prevent a CPC that is rebooting from initializing an I/O adapter that is in use by another server.

► It provides automatic restart of a hung or stopped partition. The PHYP also monitors the service processor and runs a reset or reload if it detects the loss of the service processor. It notifies the operating system if the problem is not corrected.

The AIX operating system uses PHYP services to manage the Translation Control Entry (TCE) tables. The operating system communicates the wanted I/O bus address to logical mapping, and the PHYP returns the I/O bus address to physical mapping within the specific TCE table. The PHYP needs a dedicated memory region for the TCE tables to translate the I/O address to the partition memory address. The PHYP then can monitor direct memory access (DMA) transfers to the Peripheral Component Interconnect Express (PCIe) adapters.

3.1.2 POWER8 processor

The IBM POWER8 processor implements 64-bit IBM Power Architecture® technology and represents a leap forward in technology achievement and associated computing capability. The multi-core architecture of the POWER8 processor modules is matched with innovation across a wide range of related technologies to deliver leading throughput, efficiency, scalability, and RAS.

Areas of innovation, enhancement, and consolidation

The POWER8 processor represents an important performance increase compared to previous generations. The POWER8 processor features the following areas of innovation, enhancement, and consolidation:

► On-chip L3 cache that is implemented in embedded dynamic random access memory (eDRAM), which improves latency and bandwidth. Lower energy consumption and a smaller physical footprint are benefits.

► Cache hierarchy and component innovation.

► Advances in memory subsystem.

► Advances in off-chip signaling.

► The POWER8 processor features intelligent threads that can vary based on the workload demand. The system automatically selects whether a workload benefits from dedicating as much capability as possible to a single thread of work, or if the workload benefits more from spreading the capability across two or four threads of work. With more threads, the POWER8 processor can deliver more total capacity as more tasks are accomplished in parallel. With fewer threads, those workloads that need fast individual tasks can get the performance that they need for maximum benefit.

The remainder of this section describes the RAS features of the POWER8 processor. These features and abilities apply to the DS8880. You can read more about the POWER8 and processor configuration, from the DS8880 architecture point of view, in 2.3.1, "IBM POWER8-based servers" on page 35.

POWER8 RAS features

The following sections describe the RAS leadership features of IBM POWER8 systems.

POWER8 processor instruction retry

As with previous generations, the POWER8 processor can run processor instruction retry and alternative processor recovery for many core-related faults. This ability significantly reduces exposure to permanent and intermittent errors in the processor core.

With the instruction retry function, when an error is encountered in the core in caches and certain logic functions, the POWER8 processor first automatically retries the instruction. If the source of the error was truly transient, the instruction succeeds and the system can continue normal operation.

POWER8 alternative processor retry

Hard failures are more difficult because permanent errors are replicated each time that the instruction is repeated. Retrying the instruction does not help in this situation because the instruction continues to fail. Similar to the IBM POWER6+™, IBM POWER7®, and POWER7+ processors, the POWER8 processors can extract the failing instruction from the faulty core and retry it elsewhere in the system for many faults. The failing core is then dynamically unconfigured and scheduled for replacement. The entire process is transparent to the partition that owns the failing instruction. Systems with POWER8 processors are designed to avoid a full system outage.

POWER8 cache protection

Processor instruction retry and alternative processor retry protect processor and data caches. L1 cache is divided into sets. The POWER8 processor can deallocate all but one set before a Processor Instruction Retry is run. In addition, faults in the Segment Lookaside Buffer (SLB) array are recoverable by the IBM POWER Hypervisor™. The SLB is used in the core to run address translation calculations.

The L2 and L3 caches in the POWER8 processor are protected with double-bit detect single-bit correct error correction code (ECC). Single-bit errors are corrected before they are forwarded to the processor, and they are then written back to L2 or L3.

In addition, the caches maintain a cache line delete capability. A threshold of correctable errors that is detected on a cache line can result in purging the data in the cache line and removing the cache line from further operation without requiring a restart. An ECC uncorrectable error that is detected in the cache can also trigger a purge and delete of the cache line. This action results in no loss of operation because an unmodified copy of the data can be held in system memory to reload the cache line from main memory. Modified data is handled through special uncorrectable error handling. L2 and L3 deleted cache lines are marked for persistent deconfiguration on subsequent system restarts until they can be replaced.

POWER8 first-failure data capture

First-failure data capture (FFDC) is an error isolation technique. FFDC ensures that when a fault is detected in a system through error checkers or other types of detection methods, the root cause of the fault is captured without the need to re-create the problem or run an extended tracing or diagnostics program.

For most faults, a good FFDC design means that the root cause is detected automatically without intervention by a service representative. Pertinent error data that relates to the fault is captured and saved for analysis. In hardware, FFDC data is collected from the fault isolation registers and the associated logic. In firmware, this data consists of return codes, function calls, and so on.

FFDC *check stations* are carefully positioned within the server logic and data paths to ensure that potential errors can be identified quickly and accurately tracked to a field-replaceable unit (FRU).

This proactive diagnostic strategy is a significant improvement over the classic, less accurate *restart and diagnose* service approach.

Redundant components

High opportunity components (those components that most effect system availability) are protected with redundancy and the ability to be repaired concurrently.

The use of the following redundant components allows the system to remain operational:

► POWER8 cores, which include redundant bits in L1 instruction and data caches, L2 caches, and L2 and L3 directories
► Power System S824, Power System S822, and Power System E850 use CPC main memory and CDIMMs, which use an innovative ECC algorithm from IBM research that improves single-bit error correction and memory failure identification
► Redundant cooling
► Redundant power supplies
► Redundant 16x loops to the I/O subsystem

Self-healing

For a system to be self-healing, it must be able to recover from a failing component by detecting and isolating the failed component. The system is then able to take the component offline, fix, or isolate it, and then reintroduce the fixed or replaced component into service without any application disruption. Self-healing technology includes the following examples:

► Bit steering to redundant memory if a memory module failed to keep the server operational.

► *Chipkill*, which is an enhancement that enables a system to sustain the failure of an entire DRAM chip. An ECC word uses 18 DRAM chips from two DIMM pairs, and a failure on any of the DRAM chips can be fully recovered by the ECC algorithm. The system can continue indefinitely in this state with no performance degradation until the failed DIMM can be replaced.

► Single-bit error correction by using ECC without reaching error thresholds for main, L2, and L3 cache memory.

► L2 and L3 cache line delete capability, which provides more self-healing.

► ECC extended to inter-chip connections on the fabric and processor bus.

► *Hardware scrubbing* is a method that is used to address intermittent errors. IBM POWER8 processor-based systems periodically address all memory locations. Any memory locations with a correctable error are rewritten with the correct data.

► Dynamic processor deallocation.

Memory reliability, fault tolerance, and integrity

POWER8 uses ECC circuitry for system memory to correct single-bit memory failures. In POWER8, an ECC word consists of 72 bytes of data. Of these bytes, 64 are used to hold application data. The remaining 8 bytes are used to hold check bits and more information about the ECC word. This innovative ECC algorithm from IBM research works on DIMM pairs on a rank basis. With this ECC code, the system can dynamically recover from an entire DRAM failure (Chipkill). It can also correct an error even if another symbol (a byte, which is accessed by a 2-bit line pair) experiences a fault. This feature is an improvement from the Double Error Detection/Single Error Correction ECC implementation on the POWER6+ processor-based systems.

The memory DIMMs also use hardware scrubbing and thresholding to determine when memory modules within each bank of memory must be used to replace modules that exceeded their threshold of error count (*dynamic bit-steering*). Hardware scrubbing is the process of reading the contents of the memory during idle time and checking and correcting any single-bit errors that accumulated by passing the data through the ECC logic. This function is a hardware function on the memory controller chip, and does not influence normal system memory performance.

Fault masking

If corrections and retries succeed and do not exceed threshold limits, the system remains operational with full resources, and no external administrative intervention is required.

Mutual surveillance

The service processor monitors the operation of the IBM POWER Hypervisor firmware during the boot process and monitors for loss of control during system operation. It also allows the POWER Hypervisor to monitor service processor activity. The service processor can take the correct action (including calling for service) when it detects that the POWER Hypervisor firmware lost control. The POWER Hypervisor can also request a service processor repair action, if necessary.

3.1.3 AIX operating system

Each CPC is a server that runs the IBM AIX Version 7.1 operating system (OS). This OS is the IBM well-proven, scalable, and open standards-based UNIX-like OS. This version of AIX includes support for Failure Recovery Routines (FRRs).

For more information about the features of the IBM AIX operating system, see this website:

http://www.ibm.com/systems/power/software/aix/index.html

3.1.4 Cross-cluster communication

In the DS8880, the I/O enclosures are connected point-to-point and each CPC uses a PCIe architecture. DS8880 uses the PCIe paths between the I/O enclosures to provide the cross-cluster (XC) communication between CPCs. This configuration means that no separate path is between the XC communications and I/O traffic, which simplifies the topology. During normal operations, the XC communication traffic uses a small portion of the overall available PCIe bandwidth (less than 1.7%), so XC communication traffic has a negligible effect on I/O performance.

Figure 3-1 shows the redundant PCIe fabric design for XC communication in the DS8880. If the I/O enclosure that is used as the XC communication path fails, the system automatically uses an available alternative I/O enclosure for XC communication.
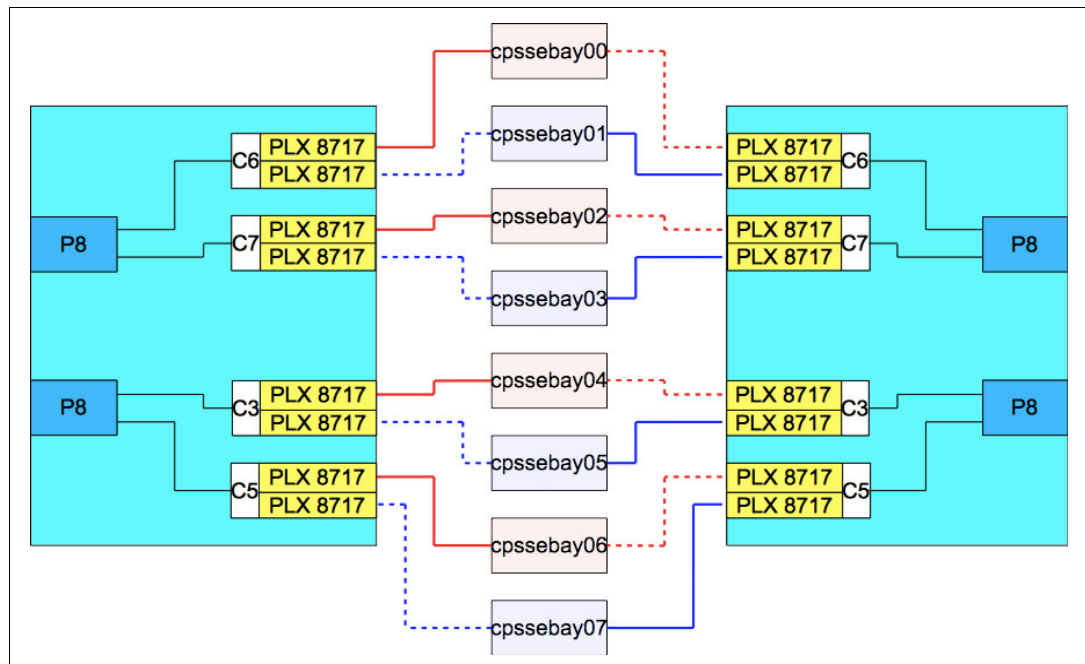
Figure 3-1   DS8886 XC communication through the PCIe fabric and I/O enclosures

3.1.5 Environmental monitoring

The environment (power, fans, and temperature) is monitored by the FSP. Environmental critical and non-critical conditions generate emergency power-off warning (EPOW) events. Critical events (for example, a complete input power loss) trigger the correct signals from the hardware to start an emergency shutdown to prevent data loss without operating system or firmware involvement. Non-critical environmental events are logged and reported by using Event Scan.

The temperature is also monitored. If the ambient temperature rises above a preset operating range, the rotational speed of the cooling fans increases. Temperature monitoring also warns the Licensed Internal Code (LIC) of potential environmental problems. An orderly system shutdown, including a service call to IBM, occurs when the operating temperature exceeds a critical level.

Voltage monitoring provides a warning and an orderly system shutdown when the voltage is out of the operational specification range.

Additional monitoring support can be found with DSCLI `showsu` command by using the Added Energy Report Test Mode, ER Recorded, ER Power Usage, ER Inlet Temp, ER I/O Usage, and ER Data Usage fields.

3.1.6 Resource unconfiguration

If recoverable errors exceed threshold limits, resources can be unconfigured to keep the system operational. This ability allows deferred maintenance at a more convenient time. Dynamic deconfiguration of potentially failing components is nondisruptive, which allows the system to continue to run. Persistent deconfiguration occurs when a failed component is detected. It is then deactivated on a subsequent restart.

Dynamic deconfiguration functions include the following components:
► Processor
► L3 cache lines
► Partial L2 cache deconfiguration
► PCIe bus and slots

Persistent deconfiguration functions include the following components:
► Processor
► Memory
► Unconfigure or bypass failing I/O adapters
► L2 cache

After a hardware error that is flagged by the service processor, the subsequent restart of the processor complex starts extended diagnostic testing. If a processor or memory is marked for persistent deconfiguration, the boot process attempts to proceed to completion with the faulty device automatically unconfigured. Failing I/O adapters are unconfigured or bypassed during the boot process.

3.2 CPC failover and failback

To understand the process of CPC failover and failback, the logical construction of the DS8880 must be reviewed. For more information, see Chapter 4, "Virtualization concepts" on page 97.

3.2.1 Dual cluster operation and data protection

For processing host data, a basic premise of RAS is that the DS8880 always tries to maintain two copies of the data while the data moves through the storage system. Two areas of the primary memory of the nodes are used for holding host data: Cache memory and NVS. NVS for DS8886 is 1/16 of system memory with a minimum of 8 GB total on most configurations.

For a DS8886/DS8888 with 2 TB of total system memory, NVS is 1/32 of system memory (64 GB). For DS8884, all configurations use 1/16 of system memory for a minimum of 4 GB of NVS. NVS contains write data until the data is destaged from cache to the drives. NVS data is written to the CPC hard disk drives (HDDs) in an emergency shutdown due to a complete loss of input AC power.

When a write is sent to a volume and both the nodes are operational, the *write data* is placed into the cache memory of the owning node and into the NVS of the other processor complex. The NVS copy of the write data is accessed only if a write failure occurs and the cache memory is empty or possibly invalid. Otherwise, the NVS copy of the write data is discarded after the destage from cache to the drives is complete.

The location of write data when both CPCs are operational is shown in Figure 3-2.
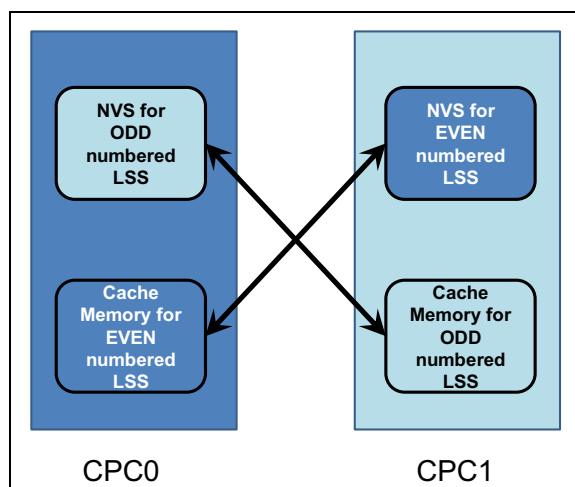
Figure 3-2   Write data when CPCs are both operational

Figure 3-2 shows how the cache memory of node 0 in CPC0 is used for all logical volumes that are members of the even logical subsystems (LSSs). Likewise, the cache memory of node 1 in CPC1 supports all logical volumes that are members of odd LSSs. For every write that is placed into cache, a copy is placed into the NVS memory that is in the alternative node. Therefore, the following normal flow of data for a write when both CPCs are operational is used:

1. Data is written to cache memory in the owning node. At the same time, data is written to NVS memory of the alternative node.

2. The write operation is reported to the attached host as completed.

3. The write data is destaged from the cache memory to a drive array.

4. The write data is discarded from the NVS memory of the alternative node.

Under normal operation, both DS8880 nodes are actively processing I/O requests. The following sections describe the failover and failback procedures that occur between the CPCs when an abnormal condition affects one of them.

3.2.2 Failover

In the example that is shown in Figure 3-3, CPC0 failed. CPC1 needs to take over all of the CPC0 functions. All storage arrays are accessible by both CPCs.
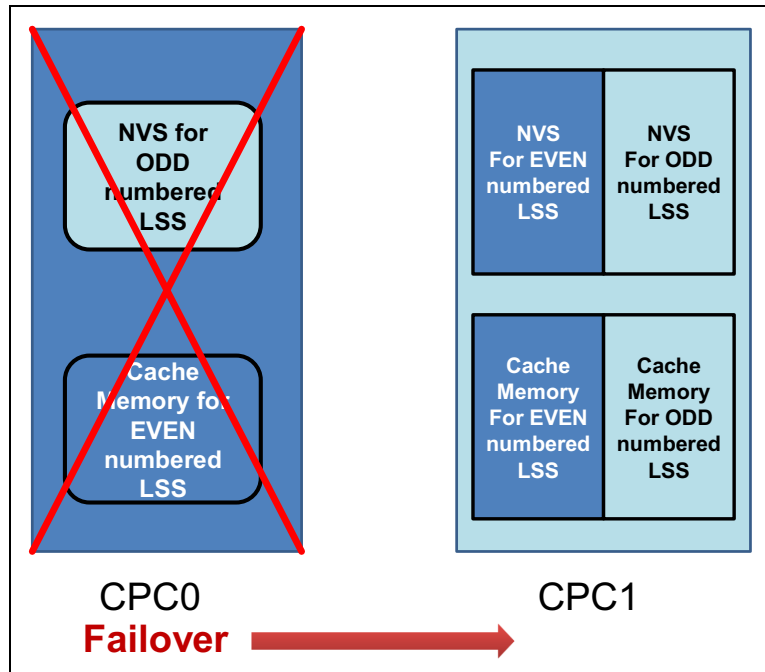
Figure 3-3   CPC0 failover to CPC1

At the moment of failure, node 1 in CPC1 includes a backup copy of the node 0 write data in its own NVS. From a data integrity perspective, the concern is the backup copy of the node 1 write data, which was in the NVS of node 0 in CPC0 when it failed. Because the DS8880 now has only one copy of that data (active in the cache memory of node 1 in CPC1), it performs the following steps:

1. Node 1 destages the contents of its NVS (the node 0 write data) to the drive subsystem. However, before the actual destage and at the beginning of the failover, the following tasks occur:

   a. The surviving node starts by preserving the data in cache that was backed up by the failed CPC NVS. If a restart of the single working CPC occurs before the cache data is destaged, the write data remains available for subsequent destaging.

   b. The existing data in cache (for which still only a single volatile copy exists) is added to the NVS so that it remains available if the attempt to destage fails or a server restart occurs. This function is limited so that it cannot consume more than 85% of NVS space.

2. The NVS and cache of node 1 are divided in two, half for the odd LSSs and half for the even LSSs.

3. Node 1 begins processing the I/O for all the LSSs, taking over for node 0.

This entire process is known as a *failover*. After failover, the DS8880 operates as shown in Figure 3-3. Node 1 now owns all the LSSs, which means all reads and writes are serviced by node 1. The NVS inside node 1 is now used for both odd and even LSSs. The entire failover process is transparent to the attached hosts.

The DS8880 can continue to operate in this state indefinitely. No functionality is lost but the redundancy is lost, and performance is decreased because of the reduced system cache. Any critical failure in the working CPC renders the DS8880 unable to serve I/O for the arrays. Because of this failure, the IBM Support team must begin work immediately to determine the scope of the failure and to build an action plan to restore the failed CPC to an operational state.

3.2.3 Failback

The *failback* process always begins automatically when the DS8880 determines that the failed CPC resumed to an operational state. If the failure was relatively minor and recoverable by the DS8880 operating system, the software starts the resume action. If a service action occurred and hardware components were replaced, the IBM service support representative (SSR) or remote support engineer resumes the failed CPC.

This example in which CPC0 failed assumes that CPC0 was repaired and resumed. The failback begins with server 1 in CPC1 starting to use the NVS in node 0 in CPC0 again, and the ownership of the even LSSs being transferred back to node 0. Normal I/O processing, with both CPCs operational, then resumes. Just like the failover process, the failback process is transparent to the attached hosts.

In general, recovery actions (failover or failback) on the DS8880 do not affect I/O operation latency by more than 8 seconds.

If you require real-time response in this area, contact IBM to determine the latest information about how to manage your storage to meet your requirements.

3.2.4 NVS and power outages

During normal operation, the DS8880 preserves write data by storing a duplicate in the NVS of the alternative CPC. To ensure that this write data is not lost because of a power event, the DS8880 direct current uninterruptible power supplies (DC-UPSs) contain Battery Service Module (BSM) sets. The BSM sets provide continuity of power during AC power loss that is long enough for the CPCs to write modified data to their internal hard disks. The design is to not move the data from NVS to the disk arrays. Instead, each CPC features dual internal disks that are available to store the contents of NVS.

> **Important:** Unless the extended power-line disturbance (ePLD) feature is installed, the BSM sets ensure storage operation for up to 4 seconds in a power outage. After this period, the BSM sets keep the CPCs and I/O enclosures operable long enough to write NVS contents to internal CPC hard disks. The ePLD feature can be ordered so that drive operations can be maintained for 40 seconds after a power disruption.

If any frames lose AC input (which is known as *wall power* or *line power*) to both DC-UPSs, the CPCs are informed that they are running on batteries. In continuous power unavailability for 4 seconds, the CPCs begin a shutdown procedure, which is known as an *on-battery condition*. It is during this emergency shutdown that the entire contents of NVS are written to the CPC HDDs so that the write data will be available for destaging after the CPCs are operational again.

If power is lost to a single DC-UPS, the partner DC-UPS provides power to this UPS, and the output power to other DS8880 components remains redundant.

The following sections describe the steps that are used if dual AC power is lost to the entire frame.

Power loss

When an on-battery condition shutdown begins, the following events occur:

1. All host adapter I/O is blocked.

2. Each CPC begins copying its NVS data to internal disk (not the storage drives). For each CPC, two copies are made of the NVS data. This process is referred to as fire hose dump (FHD).

3. When the copy process is complete, each CPC shuts down.

4. When the shutdown in each CPC is complete, the DS8880 is powered off.

Power restored

When power is restored, the DS8880 needs to be powered on manually, unless the remote power control mode is set to *automatic*.

> **Note:** Be careful before you decide to set the remote power control mode to *automatic*. If the remote power control mode is set to *automatic*, after input power is lost, the DS8880 is powered on automatically as soon as external power becomes available again. For more information about how to set power control on the DS8880, see the IBM System Storage DS8000 IBM Knowledge Center at the following website:
>
> `https://www.ibm.com/support/knowledgecenter/ST8NCA/product_welcome/ds8000_kcwelcome.html`

After the DS8880 is powered on, the following events occur:

1. When the CPCs power on, the PHYP loads and power-on self-test (POST) runs.

2. Each CPC begins the initial microcode load (IML).

3. At an early stage in the IML process, the CPC detects NVS data on its internal disks and restores the data to destage it to the storage drives.

3.3 Data flow in the DS8880

The DS8880 connectivity between the CPC and the I/O enclosures uses the many strengths of the PCIe architecture.

For more information, see 2.3.4, "Peripheral Component Interconnect Express adapters" on page 39.

3.3.1 I/O enclosures

As shown in Figure 3-1 on page 70, each CPC on a DS8886 is connected to all four I/O enclosures (base frame) or all eight I/O enclosures when the first expansion frame is installed, by PCI Express cables. This configuration makes each I/O enclosure an extension of each processor complex. The DS8884 has a base of two I/O enclosures, and a maximum of four.

The DS8880 I/O enclosures use adapters with PCI Express connections. The I/O adapters in the I/O enclosures are concurrently replaceable. Each slot can be independently powered off for concurrent replacement of a failed adapter, installation of a new adapter, or removal of a failed adapter.

In addition, each I/O enclosure has N+1 power and cooling in the form of two power supplies with integrated fans. The power supplies can be concurrently replaced and a single power supply can provide power to the whole I/O enclosure.

3.3.2 Host connections

Each DS8880 Fibre Channel host adapter (HA) provides four or eight ports for connectivity directly to a host or to a Fibre Channel storage area network (SAN). The 16 Gb host adapters have four ports.

Single or multiple path

The HAs are shared between the CPCs. To illustrate this concept, Figure 3-4 shows a potential system configuration. In this example, two I/O enclosures are shown. Each I/O enclosure has up to two Fibre Channel host adapters. If a host server has only a single path to a DS8880, as shown in Figure 3-4, it can access volumes that belong to all LSSs because the HA directs the I/O to the correct CPC. However, if an error occurs on the HA, HA port, I/O enclosure, or in the SAN, all connectivity is lost because this configuration has no redundancy, and it is not advised. The same is true for the host bus adapter (HBA) in the attached host, making it a single point of failure (SPOF) without a redundant HBA.

> **Important:** For host connectivity, hosts that access the DS8880 must have at least two connections to I/O ports on separate host adapters in separate I/O enclosures.

Figure 3-4 shows a single-path host connection.
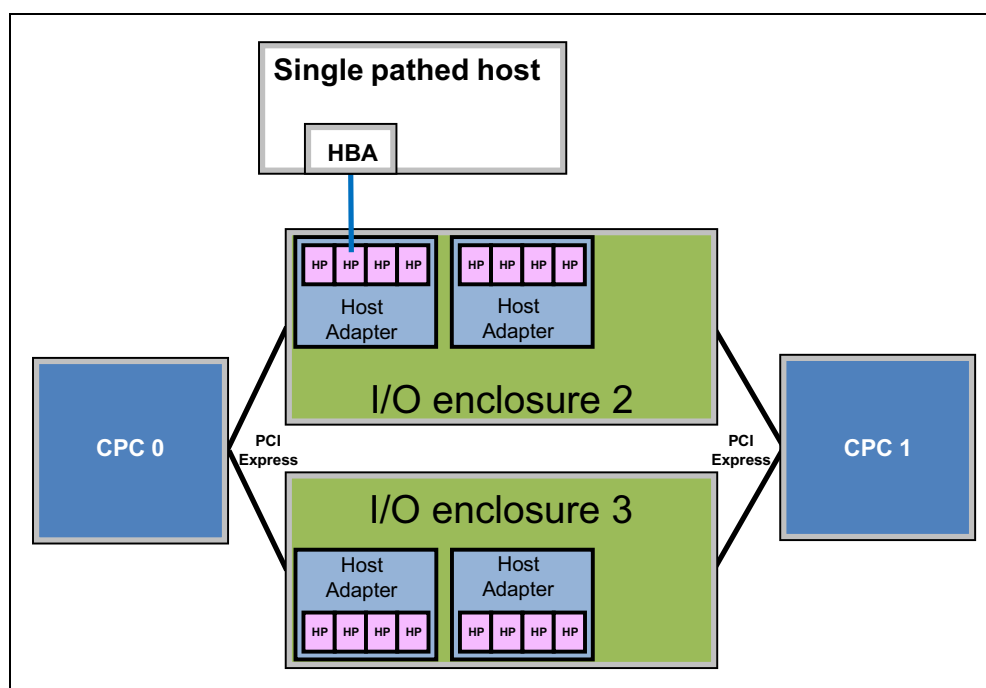
Figure 3-4   A single-path host connection

A more robust design is shown in Figure 3-5, in which the host is attached to separate Fibre Channel host adapters in separate I/O enclosures. This configuration is also important because during a LIC update, a host adapter port might need to be taken offline. This configuration allows host I/O to survive a hardware failure on any component on either path.
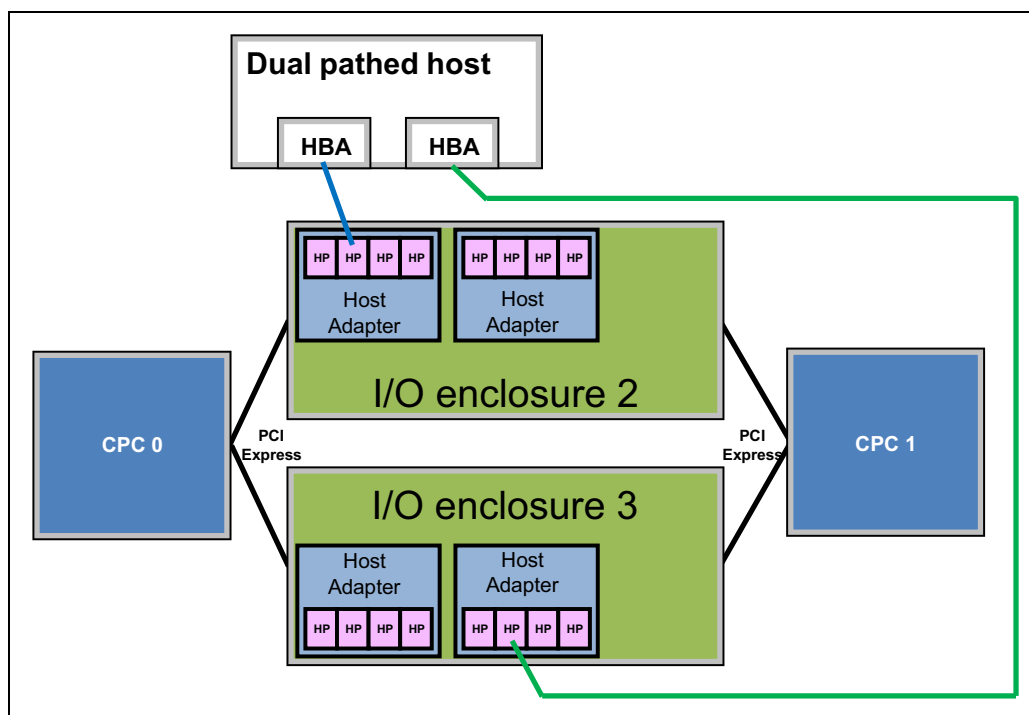
Figure 3-5   A dual-path host connection

SAN/FICON switches

Because many hosts can connect to the DS8880 each by using multiple paths, the number of host adapter ports that are available in the DS8880 might not be sufficient to accommodate all of the connections. The solution to this problem is the use of SAN switches or directors to switch logical connections from multiple hosts. In a z Systems environment, a SAN switch or director that supports Fibre Channel connection (FICON) is required.

A logic or power failure in a switch or director can interrupt communication between hosts and the DS8880. Provide more than one switch or director to ensure continued availability. Configure ports from two separate host adapters in two separate I/O enclosures to go through each of two directors. The complete failure of either director leaves the paths that are configured to the alternative director still available.

Using channel extension technology

For Copy Services scenarios in which single mode fiber distance limits are exceeded, use of channel extension technology is required. The following website contains information about network devices that are marketed by IBM and other companies to extend Fibre Channel communication distances. They can be used with DS8000 Series Metro Mirror, Global Copy, Global Mirror, Metro/Global Mirror (MGM) Support, and z/OS Global Mirror. For more information, see the DS8000 Series Copy Services Fibre Channel Extension Support Matrix:

http://www.ibm.com/support/docview.wss?uid=ssg1S7003277

Support for T10 Data Integrity Field standard

The DS8880 incorporates the ANSI T10 Data Integrity Field (DIF) standard for fixed-block (FB) volumes.

When data is read, the DIF is checked before the data leaves the DS8880 and again when the data is received by the host system. Previously, it was only possible to ensure the data integrity within the storage system with ECC. However, T10 DIF can now check end-to-end data integrity through the SAN. Checking is done by hardware, so no performance impact occurs.

For more information about T10 DIF implementation in the DS8880, see "T10 data integrity field support" on page 109.

Robust and resilient SAN data transfer

Forward Error Correction (FEC) is enabled on 16 Gbps host adapters by implementing redundant data transfer in terms of ECC. FEC is also implemented on z13 (and later) Systems to achieve the end-to-end prevention of I/O errors.

To provide more proactive system diagnosis information about SAN fabric systems, *read diagnostic parameters*, which comply to industry standards, are implemented on the DS8880. This function provides host software with the capability to perform predictive failure analysis on degraded SAN links before they fail.

Multipathing software

Each attached host operating system requires multipathing software to manage multiple paths to the same device, and to provide at least redundant routes for host I/O requests. When a failure occurs on one path to a logical device, the multipathing software on the attached host can identify the failed path and route the I/O requests for the logical device to alternative paths. Furthermore, it can likely detect when the path is restored. The multipathing software that is used varies by attached host operating system and environment, as described in the following sections.

Open systems

In most open systems environments, multipathing is available at the operating system level. The Subsystem Device Driver (SDD), which was provided and maintained by IBM for several operating systems, is the old approach for a multipathing solution.

For the AIX operating system, the DS8880 is supported through the AIX multipath I/O (MPIO) framework, which is included in the base AIX operating system. Choose either to use the base AIX MPIO support or to install the Subsystem Device Driver Path Control Module (SDDPCM).

For multipathing under Microsoft Windows, Subsystem Device Driver Device Specific Module (SDDDSM) is available.

For more information about the multipathing software that might be required for various operating systems, see the IBM System Storage Interoperation Center (SSIC) at this website:

https://www.ibm.com/systems/support/storage/ssic/

Multipathing is covered in more detail in the *IBM DS8000 Host Systems Attachment Guide*, SC27-8527.

z Systems

In the z Systems environment, the preferred practice is to provide multiple paths from each host to a storage system. Typically, four or eight paths are installed. The channels in each host that can access each logical control unit (LCU) in the DS8880 are defined in the hardware configuration definition (HCD) or input/output configuration data set (IOCDS) for that host. Dynamic Path Selection (DPS) allows the channel subsystem to select any available (non-busy) path to start an operation to the disk subsystem. Dynamic Path Reconnect (DPR) allows the DS8880 to select any available path to a host to reconnect and resume a disconnected operation, for example, to transfer data after disconnection because of a cache miss.

These functions are part of the z Systems architecture and they are managed by the channel subsystem on the host and the DS8880.

A physical FICON path is established when the DS8880 port sees light on the fiber, for example, a cable is plugged in to a DS8880 host adapter, a processor or the DS8880 is powered on, or a path is configured online by z/OS. Now, logical paths are established through the port between the host, and part or all of the LCUs in the DS8880 are controlled by the HCD definition for that host. This configuration happens for each physical path between a z Systems host and the DS8880. Multiple system images can be in a CPU. Logical paths are established for each system image. The DS8880 then knows the paths that can be used to communicate between each LCU and each host.

Control-unit initiated reconfiguration (CUIR) varies off a path or paths to all z Systems hosts to allow service to an I/O enclosure or host adapter, then varies on the paths to all host systems when the host adapter ports are available. This function automates channel path management in z Systems environments in support of selected DS8880 service actions.

CUIR is available for the DS8880 when it operates in the z/OS and IBM z/VM® environments. CUIR provides automatic channel path vary on and vary off actions to minimize manual operator intervention during selected DS8880 service actions.

CUIR also allows the DS8880 to request that all attached system images set all paths that are required for a particular service action to the offline state. System images with the correct level of software support respond to such requests by varying off the affected paths, and either notifying the DS8880 subsystem that the paths are offline, or that it cannot take the paths offline. CUIR reduces manual operator intervention and the possibility of human error during maintenance actions, and reduces the time that is required for the maintenance. This function is useful in environments in which many z/OS or z/VM systems are attached to a DS8880.

3.3.3 Metadata checks

When application data enters the DS8880, special codes or *metadata*, which is also known as *redundancy checks*, are appended to that data. This metadata remains associated with the application data while it is transferred throughout the DS8880. The metadata is checked by various internal components to validate the integrity of the data as the data moves throughout the disk system. It is also checked by the DS8880 before the data is sent to the host in response to a read I/O request. The metadata also contains information that is used as an extra level of verification to confirm that the data that is returned to the host is coming from the location that you want on the disk.

The metadata check is independent of the DS8880 T10 DIF support for fixed-block volumes (Linux on z Systems and AIX on Power Systems). For more information about T10 DIF implementation in the DS8880, see "T10 data integrity field support" on page 109.

3.4 RAS on the Hardware Management Console

The Hardware Management Console (HMC) is used to configure, manage, and maintain the DS8880. One HMC (the *primary*) is included in every DS8880 base frame. A second HMC (the *secondary*) can be ordered, and is also in the base frame. The DS8880 HMCs work with IPv4, IPv6, or a combination of both IP standards. For more information about the HMC and network connections, see 8.1.1, "Management Console hardware" on page 205 and 7.3, "Network connectivity planning" on page 194.

The HMC is the DS8880 management focal point. If no HMC is operational, it is impossible to run maintenance, modifications to the logical configuration, or Copy Services tasks, such as the establishment of FlashCopies, by using the data storage command-line interface (DS CLI) or Storage Management GUI. Ordering and implementing a secondary HMC provides a redundant management focal point.

3.4.1 Licensed Internal Code updates

The DS8880 contains many discrete redundant components. The DS8880 architecture allows concurrent code updates. This ability is achieved by using the redundant design of the DS8880. The following components have firmware that can be updated concurrently:

► FSP
► DC-UPS
► Rack power control (RPC) cards
► Host adapters
► Fibre Channel interface cards (FCICs)
► Device adapters (DAs)
► Drives, flash drives, and flash cards

The DS8880 CPCs have an operating system (AIX) and licensed machine code (LMC) that can be updated. As IBM continues to develop and improve the DS8880, new releases of firmware and LMC become available that offer improvements in function and reliability. For more information about LIC updates, see Chapter 13, "Licensed machine code" on page 409.

3.4.2 Call home and remote support

This section describes the call home feature and remote support capability.

Call home
Call home is the capability of the DS8880 to notify the client and IBM Support to report a problem. Call home is configured in the HMC at installation time. Call home to IBM Support is done over the customer network through a secure protocol. Customer notification can also be configured as email or Simple Network Management Protocol (SNMP) alerts. An example of an email notification output is shown in Example 3-1.

Example 3-1   Typical email notification output

```
REPORTING SF MTMS:      2107-981*75DMC10
FAILING SF MTMS:        2107-981*75DMC10
REPORTING SF LPAR:      SF75DMC10ESS11
PROBLEM NUMBER:         168
PROBLEM TIMESTAMP:      Oct 30, 2016 2:52:22 AM MST
REFERENCE CODE:         BE83CB93

************************  START OF NOTE LOG  **************************
```

```
BASE RACK ORDERED MTMS 2834-981*75DMC10
LOCAL HMC MTMS 8100LI7*P210927 HMC ROLE Primary
LOCAL HMC INBOUND MODE Attended  MODEM PHONE Unavailable
LOCAL HMC INBOUND CONFIG Continuous
LOCAL HMC OUTBOUND CONFIG None     FTP: disabled

REMOTE HMC Single HMC
AOS STATUS Not Running AOS VERSION 4.0
AOS ACL=(DS8k, Storage) AOS TRACE Enable
HMC CE default  HMC REMOTE default
HMC PE default  HMC DEVELOPER default
2107 BUNDLE 88.0.135.0
HMC BUILD 20151021.1
LMC LEVEL v25.80.0 build level 20151025.2
FIRMWARE LEVEL SRV0 01SV810135  SRV1 01SV810135

PARTITION NAME SF75DMC10ESS11
PARTITION HOST NAME SF75DMC10ESS11
PARTITION STATUS SFI 2107-981*75DMC11 SVR 8286-42A*21C05AV LPAR SF75DMC10ESS11 STATE =
AVAILABLE

FIRST REPORTED TIME Oct 30, 2016 2:52:22 AM MST
LAST REPORTED TIME Oct 30, 2016 2:52:22 AM MST
CALL HOME RETRY #0 of 12 on Oct 30, 2016 2:52:22 AM MST.
REFCODE BE83CB93..... <=== System Reference Code (SRC)
SERVICEABLE EVENT TEXT
Device adapter reset reached threshold, adapter fenced. ... <=== Description of Problem

FRU group MEDIUM  FRU class FRU
FRU Part Number 31P1452  FRU CCIN DAD4
FRU Serial Number YP10BG37K005
FRU Location Code U1500.1B2.RJBAY02-P1-C6
FRU Previously Replaced No
FRU Previous PMH N/A

*************************  END OF NOTE LOG  **************************
```

For more information about planning the connections that are needed for HMC installations, see Chapter 8, "IBM DS8880 Management Console planning and setup" on page 203.

For more information about setting up SNMP notification, see Chapter 14, "Monitoring with Simple Network Management Protocol" on page 419.

Remote support

Remote Support is the ability of an IBM support representative to remotely access the DS8880. This capability can be configured at the HMC, and access is by Assist On-site (AOS).

For more information about remote support operations, see Chapter 15, "Remote support" on page 433.

For more information about AOS, see *IBM Assist On-site for Storage Overview*, REDP-4889.

3.5 RAS on the storage subsystem

The DS8880 was designed to safely store and retrieve large amounts of data. Redundant Array of Independent Disks (RAID) is an industry-wide method to store data on multiple physical disks to enhance data redundancy. Many variants of RAID are in use today. The DS8880 supports RAID 6, RAID 10, and RAID 5. The DS8880 does not support a non-RAID configuration of disks, which is also known as just a bunch of disks (JBOD).

3.5.1 RAID configurations

The following RAID configurations are supported on DS8880:

► 5+P+Q+S RAID 6 configuration: The array consists of five data drives and two parity drives. The remaining drive on the array site is used as a spare.

► 6+P+Q RAID 6 configuration: The array consists of six data drives and two parity drives.

► 3+3+2S RAID 10 configuration: The array consists of three data drives that are mirrored to three copy drives. Two drives on the array site are used as spares.

► 4+4 RAID 10 configuration: The array consists of four data drives that are mirrored to four copy drives.

► 6+P+S RAID 5 configuration (with drive sizes less than 1 TB): The array consists of six data drives and one parity drive. The remaining drive of the array site is used as a spare.

► 7+P RAID 5 configuration (with drive sizes less than 1 TB): The array consists of seven data drives and one parity drive.

> **Note:** The following characteristics refer to RAID:
>
> ► Spare drives are globally available to the DA pair.
>
> ► The +P/+Q indicators do not mean that a single drive is dedicated to holding the parity bits for the RAID. The DS8880 uses floating parity technology so that no single drive is always involved in every write operation. The data and parity stripes float among the member drives of the array to provide optimum write performance.

Capacity Magic is an easy-to-use tool to help with capacity planning for physical and usable capacities based on installation drive capacities and quantities in intended RAID configurations.

If RAID reconfiguration is required, an available capacity of free space is required, and it must be allowed for. For example, if the storage system is 98% provisioned with RAID-6 arrays, it might not be possible to complete an online reconfiguration to reconfigure a RAID 6 array into a RAID 10 array.

RAID 6 is the default when creating new arrays by using the DS Storage Manager GUI.

> **Important:** The following restrictions apply:
>
> ► Do not use RAID 5 with drive sizes larger than 1 TB.
> ► Within one HPFE Gen2 pair of six array sites, a RAID intermix is not possible.
>
> Consult with your IBM SSR for the latest information about supported RAID configurations.
>
> The Request for Price Quotation (RPQ)/Storage Customer Opportunity REquest (SCORE) System process can be used to submit requests to reenable RAID configurations which otherwise are considered as not recommended. For example, RAID5 can be configured for drives of less than 1 TB, but this is not recommended and will require a risk acceptance

3.5.2 Drive path redundancy

Each drive in the DS8880 is attached to two Fibre Channel switches. These switches are built into the standard drive enclosure FCICs. Figure 3-6 shows the redundancy features of the DS8880 switched Fibre Channel Arbitrated Loop (FC-AL) drive architecture.

Each drive has two separate connections to the enclosure backplane. This configuration allows a drive to be simultaneously attached to both FC switches. If either drive enclosure FCIC is removed from the enclosure, the switch that is included in that FCIC is also removed. However, the FC switch in the remaining FCIC retains the ability to communicate with all the drives and both DAs in a pair. Similarly, each DA has a path to each switch, so it can also tolerate the loss of a single path. If both paths from one DA fail, it cannot access the switches. However, the partner DA retains connection.
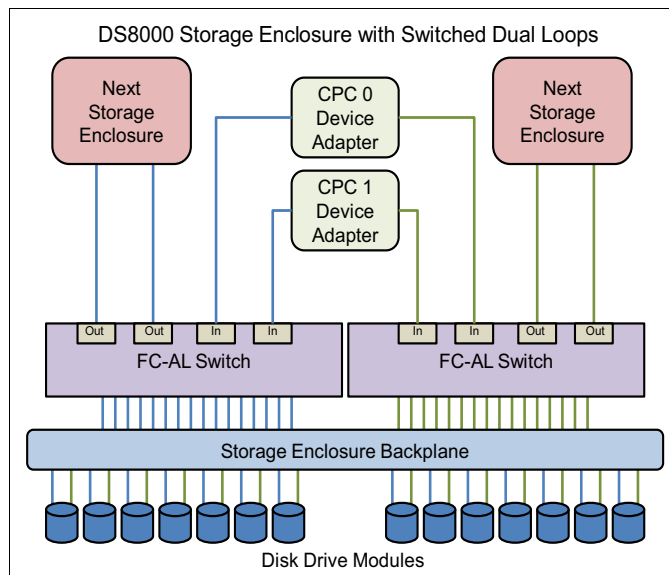
Figure 3-6   Standard drive enclosure: Switched FC-AL paths

Figure 3-6 also shows the connection paths to the neighboring drive enclosures. Because expansion is performed in this linear fashion, adding enclosures is nondisruptive.

For more information about the drive subsystem of the DS8880, see 2.5, "Storage enclosures and drives" on page 48.

3.5.3 Predictive Failure Analysis

The storage drives that are used in the DS8880 incorporate Predictive Failure Analysis (PFA) and can anticipate certain forms of failures by keeping internal statistics of read and write errors. If the error rates exceed predetermined threshold values, the drive is nominated for replacement. Because the drive did not yet fail, data can be copied directly to a spare drive by using the technique that is described in 3.5.5, "Smart Rebuild" on page 84. This copy ability avoids the use of RAID recovery to reconstruct all of the data onto the spare drive.

3.5.4 Disk scrubbing

The DS8880 periodically reads all sectors on a disk. This reading is designed to occur without any interference with application performance. If ECC detects correctable bad bits, the bits are corrected immediately. This ability reduces the possibility of multiple bad bits accumulating in a sector beyond the ability of ECC to correct them. If a sector contains data that is beyond ECC's ability to correct, RAID is used to regenerate the data and write a new copy onto a spare sector of the drive. This scrubbing process applies to drives that are array members and spares.

3.5.5 Smart Rebuild

Smart Rebuild is a function that is designed to help reduce the possibility of secondary failures and data loss of RAID arrays. It can be used to rebuild a RAID 5 array when certain drive errors occur and a normal determination is made that it is time to use a spare to proactively replace a failing drive. If the suspect drive is still available for I/O, it is kept in the array rather than being rejected as under a standard RAID rebuild. A spare is brought into the array at the same time.

The suspect drive and the new spare are set up in a temporary RAID 1 association, allowing the troubled drive to be duplicated onto the spare rather than running a full RAID reconstruction from data and parity. The new spare is then made a regular member of the array and the suspect drive is rejected from the RAID array. The array never goes through an n-1 stage in which it might suffer a complete failure if another drive in this array encounters errors. The result saves substantial time and provides a new level of availability that is not in other RAID 5 products.

Smart Rebuild is not applicable in all situations, so it is not always used. If two drives with errors are in a RAID 6 configuration, or if the drive mechanism failed to the point that it cannot accept any I/O, the standard RAID rebuild procedure is used for the RAID array. If communications across a drive fabric are compromised, such as an FC-AL loop error that causes the drive to be bypassed, standard RAID rebuild procedures are used because the suspect drive is not available for a one-to-one copy with a spare. If Smart Rebuild is not possible or does not provide the designed benefits, a standard RAID rebuild occurs.

Smart Rebuild drive error patterns are continuously analyzed as part of one of the normal tasks that are run by the DS8880 LIC. Drive firmware is optimized to report predictive errors to the DA. At any time, when certain drive errors (following specific criteria) reach a determined threshold, the RAS LIC component starts Smart Rebuild within the hour. This enhanced technique, when it is combined with a more frequent schedule, leads to considerably faster identification if drives show signs of imminent failure.

A fast response in fixing drive errors is vital to avoid a second drive failure in the same RAID 5 array, and to avoid potential data loss. The possibility of having an array that has no redundancy, such as when a RAID rebuild occurs, is reduced by shortening the time when a specific error threshold is reached until Smart Rebuild is triggered, as described in the following scenarios:

► Smart Rebuild might avoid the circumstance in which a suspected drive is rejected because Smart Rebuild process is started before rejection. Therefore, Smart Rebuild prevents the array from going to a standard RAID rebuild, during which the array has no redundancy and is susceptible to experiencing a second drive failure.

► Crossing the drive error threshold is detected by DS8880 LIC immediately because DS8880 LIC continuously analyzes drive errors.

► RAS LIC component starts Smart Rebuild after Smart Rebuild threshold criteria are met. The Smart Rebuild process runs every hour and does not wait for 24 hours as in the past.

IBM remote support representatives can manually start a Smart Rebuild if needed, such as when two drives in the same array logged temporary media errors. In this case, it is considered appropriate to manually start the rebuild.

3.5.6 RAID 5 overview

The DS8880 supports RAID 5 arrays when used with small drive sizes (up to 1 TB). RAID 5 is a method of spreading volume data plus parity data across multiple drives. RAID 5 provides faster performance by striping data across a defined set of drives. Data protection is provided by the generation of parity information for every stripe of data. If an array member fails, its contents can be regenerated by using the parity data.

The DS8880 uses the idea of *striped parity*, which means that no single drive in an array is dedicated to holding parity data, which makes such a drive active in every I/O operation. Instead, the drives in an array rotate between holding data stripes and holding parity stripes, balancing out the activity level of all drives in the array.

RAID 5 implementation in DS8880

In DS8880 standard drive enclosures, an array that is built on one array site contains eight disks. The first four array sites on a DA pair have a spare assigned, and the rest of the array sites have no spare assigned, if all disks are the same capacity and speed. An array site with a spare creates a RAID 5 array that is 6+P+S (where the P stands for parity and S stands for spare). The other array sites on the DA pair are 7+P arrays.

A High-Performance Flash Enclosure pair (Gen2) contains six array sites (like a standard enclosure pair), but the HPFE Gen-2 pair then has two spare flash cards for each such enclosure pair.

Drive failure with RAID 5

When a drive fails in a RAID 5 array, the device adapter rejects the failing drive and takes one of the hot spare drives. Then the DA starts the rebuild, which is an operation to reconstruct the data that was on the failed drive onto one of the spare drives. The spare that is used is chosen based on a smart algorithm that looks at the location of the spares and the size and location of the failed drive. The RAID rebuild is run by reading the corresponding data and parity in each stripe from the remaining drives in the array, running an exclusive-OR operation to re-create the data, and then writing this data to the spare drive.

While this data reconstruction occurs, the device adapter can still service read and write requests to the array from the hosts. Performance might degrade while the rebuild operation is in progress because DA and switched network resources are used to complete the reconstruction. Because of the switch-based architecture, this effect is minimal. Also, any read requests for data on the failed drive require data to be read from the other drives in the array, and then the DA reconstructs the data.

Performance of the RAID 5 array returns to normal when the data reconstruction onto the spare drive completes. The time that is required for rebuild varies depending on the capacity of the failed drive and the workload on the array, the switched network, and the DA. The use of array across loops (AAL) speeds up rebuild time and decreases the impact of a rebuild.

> **Note:** RAID 5 is not recommended anymore, and is disabled for drive sizes larger than 1 TB.

Even with additional optimizations like Smart Rebuild, given the relative probabilities of a dual-drive failure in a RAID 5 in combination with potentially long rebuild times of large-capacity drives justify another RAID type to achieve the highest availability levels possible with such a storage system. Consult your local IBM representative if using RAID 5 is still needed (RPQ/SCORE with large drives).

3.5.7 RAID 6 overview

The DS8880 supports RAID 6 protection. RAID 6 presents an efficient method of data protection in double drive errors, such as two drive failures, two coincident medium errors, or a drive failure and a medium error. RAID 6 protection provides more fault tolerance than RAID 5 in drive failures and uses less raw drive capacity than RAID 10.

> **Note:** RAID 6 is in most cases the default and preferred setting in the DS8880.

RAID 6 provides around a 1,000 times improvement over RAID 5 for impact risk. RAID 6 allows more fault tolerance by using a second independent distributed parity scheme (dual parity). Data is striped on a block level across a set of drives, similar to RAID 5 configurations. The second set of parity is calculated and written across all of the drives and reduces the usable space compared to RAID 5. The striping is shown in Figure 3-7 on page 87.

RAID 6 is best used with large-capacity drives because they have a longer rebuild time. One risk is that longer rebuild times increase the possibility that a second drive error occurs during the rebuild window. Comparing RAID 6 to RAID 5 performance gives about the same results on reads. For random writes, the throughput of a RAID 6 array is only two-thirds of a RAID 5, due to the extra parity handling. Workload planning is important before RAID 6 for write-intensive applications is implemented, including Copy Services targets. In case of very high random-write ratios, RAID 10 can be the better choice.

When RAID 6 is sized correctly for the I/O demand, it is a considerable reliability enhancement, as shown in Figure 3-7.

One Stride with 5 data drives (5+P+Q):

Drives	0	1	2	3	4	5	6
	0	1	2	3	4	P00	P01
	5	6	7	8	9	P10	P11
	10	11	12	13	14	P20	P21
	15	16	17	18	19	P30	P31
							P41

P00 = 0+1+2+3+4; P10 = 5+6+7+8+9; etc.

P01 = 9+13+17+0; P11 = 14+18+1+5; etc.

P41 = 4+8+12+16

NOTE: For illustrative purposes only – implementation details vary, parity is striped across all drives
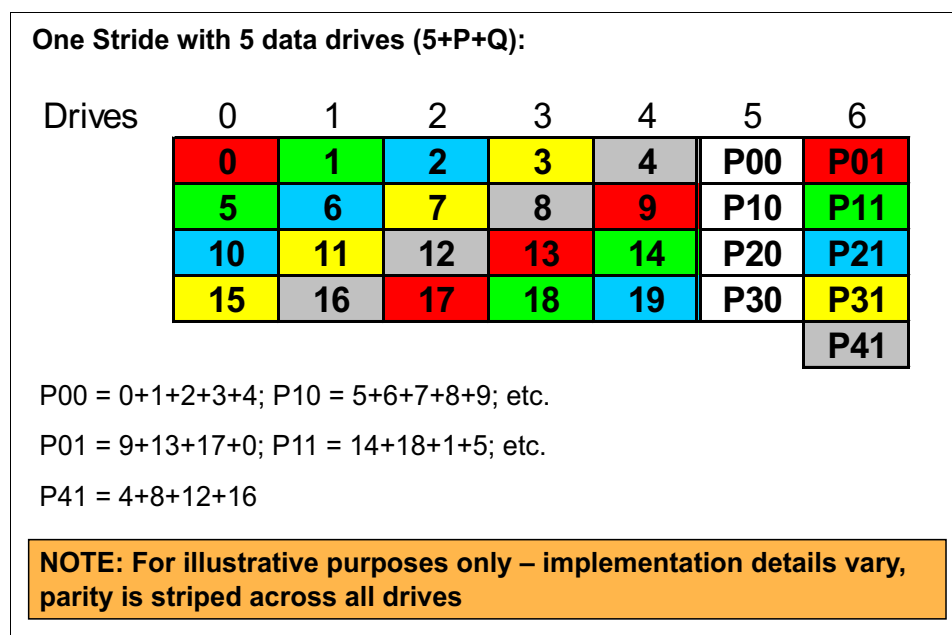
Figure 3-7   Illustration of one RAID 6 stripe on a 5+P+Q+S array

RAID 6 implementation in the DS8880

A RAID 6 array in one array site of a DS8880 can be built in one of the following configurations:

► In a seven-drive array, two drives are always used for parity, and the eighth drive of the array site is needed as a spare. This type of RAID 6 array is referred to as a 5+P+Q+S array, where P and Q stand for parity and S stands for spare.

► A RAID 6 array, which consists of eight drives, is built when all necessary spare drives are configured for the DA pair. An eight-drive RAID 6 array also always uses two drives for parity, so it is referred to as a 6+P+Q array.

Drive failure with RAID 6

When a drive fails in a RAID 6 array, the DA starts to reconstruct the data of the failing drive onto one of the available spare drives. A smart algorithm determines the location of the spare drive to use, depending on the capacity and the location of the failed drive. After the spare drive replaces a failed drive in a redundant array, the recalculation of the entire contents of the new drive is run by reading the corresponding data and parity in each stripe from the remaining drives in the array. This data is then written to the spare drive.

During the rebuild of the data on the new drive, the DA can still handle I/O requests of the connected hosts to the affected array. Performance degradation might occur during the reconstruction because DAs and switched network resources are used to do the rebuild. Because of the switch-based architecture of the DS8880, this effect is minimal. Additionally, any read requests for data on the failed drive require data to be read from the other drives in the array, and then the DA reconstructs the data. Any subsequent failure during the reconstruction within the same array (second drive failure, second coincident medium errors, or a drive failure and a medium error) can be recovered without data loss.

Performance of the RAID 6 array returns to normal when the data reconstruction on the spare drive is complete. The rebuild time varies, depending on the capacity of the failed drive and the workload on the array and the DA. The completion time is comparable to a RAID 5 rebuild, but slower than rebuilding a RAID 10 array in a single drive failure.

3.5.8 RAID 10 overview

RAID 10 provides high availability by combining features of RAID 0 and RAID 1. RAID 0 optimizes performance by striping volume data across multiple drives. RAID 1 provides drive mirroring, which duplicates data between two drives. By combining the features of RAID 0 and RAID 1, RAID 10 provides a second optimization for fault tolerance. Data is striped across half of the drives in the RAID 1 array. The same data is also striped across the other half of the array, which creates a mirror. Access to data is preserved if one drive in each mirrored pair remains available. RAID 10 offers faster data reads and writes than RAID 5 or RAID 6 because it does not need to manage parity. However, with half of the drives in the group used for data and the other half mirroring that data, RAID 10 arrays have less usable capacity than RAID 6 or RAID 5 arrays.

RAID 10 is commonly used for workloads that require the highest performance from the drive subsystem. Although with RAID 6, each front-end random write I/O might theoretically lead to 6 back-end I/Os including the parity updates (RAID penalty), this number is 4 for a RAID 5 and only 2 for a RAID 10 (not counting cache optimizations). Typical areas of operation for RAID 10 are workloads with a high random-write ratio. Either member in the mirrored pair can respond to the read requests.

Tip: Consider RAID 10 for random-write ratios over 35%.

RAID 10 implementation in DS8880

In the DS8880, the RAID 10 implementation is achieved by using six or eight drives. If spares need to be allocated from the array site, six drives are used to make a three-drive RAID 0 array, which is then mirrored to a three-drive array (3x3). If spares do not need to be allocated, eight drives are used to make a four-drive RAID 0 array, which is then mirrored to a four-drive array (4x4).

Drive failure with RAID 10

When a drive fails in a RAID 10 array, the DA rejects the failing drive and takes a hot spare into the array synthesists. Data is then copied from the good drive to the hot spare drive. The spare that is used is chosen based on a smart algorithm that looks at the location of the spares and the size and location of the failed drive. Remember, a RAID 10 array is effectively a RAID 0 array that is mirrored. Therefore, when a drive fails in one of the RAID 0 arrays, you can rebuild the failed drive by reading the data from the equivalent drive in the other RAID 0 array.

While this data copy is occurring, the DA can still service read and write requests to the array from the hosts. Performance might degrade while the copy operation is in progress because DA and switched network resources are used to rebuild the RAID. Because a good drive is available, this effect is minimal. Read requests for data on the failed drive likely are not affected because they are all directed to the good copy on the mirrored drive. Write operations are not affected.

Performance of the RAID 10 array returns to normal when the data copy onto the spare drive completes. The time that is taken for rebuild can vary, depending on the capacity of the failed drive and the workload on the array and the DA.

In relation to a RAID 5, RAID 10 rebuild completion time is faster because rebuilding a RAID 5 6+P configuration requires six reads plus one parity operation for each write. However, a RAID 10 configuration requires one read and one write (essentially, a direct copy).

Array across loops and RAID 10

The DS8880, as with previous DS8000 generations, implements the concept of AAL. With AAL, an array site is split into two halves. Half of the site is on the first drive loop of a DA pair and the other half is on the second drive loop of that DA pair. AAL is implemented primarily to maximize performance, and is used for all the RAID types in the DS8880. However, with RAID 10, you can take advantage of AAL to provide a higher level of redundancy. The DS8880 RAS code deliberately ensures that one RAID 0 array is maintained on each of the two loops that are created by a DA pair. This configuration means that in the unlikely event of a complete loop outage, the DS8880 does not lose access to the RAID 10 array. This access is not lost because when one RAID 0 array is offline, the other remains available to service drive I/O. Figure 2-26 on page 54 shows a diagram of this strategy.

3.5.9 Spare creation

This section describes methods of spare creation.

Standard drive enclosures

When the arrays are created on a DS8880 standard drive enclosure, the LIC determines the array sites that contain spares. The first array sites on each DA pair that are assigned to arrays contribute one or two spares (depending on the RAID option) until the DA pair has access to at least four spares, with two spares placed on each loop.

A minimum of one spare is created for each array site that is assigned to an array until the following conditions are met:

► A minimum of four spares per DA pair exist.

► A minimum of four spares for the largest capacity array site on the DA pair exist.

► A minimum of two spares of capacity and revolutions per minute (RPM) greater than or equal to the fastest array site of any capacity on the DA pair exist.

Spare rebalancing

The DS8880 implements a spare rebalancing technique for spare drives. When a drive fails and a hot spare is taken, it becomes a member of that array. When the failed drive is repaired, DS8880 LIC might choose to allow the hot spare to remain where it was moved. However, it can instead choose to migrate the spare to a more optimum position. This migration is performed to better balance the spares across the FC-AL loops to provide the optimum spare location based on drive capacity and spare availability.

It might be preferable that the drive that is in use as an array member is converted to a spare. In this case, the data on that disk drive module (DDM) is migrated in the background onto an existing spare by using the Smart Rebuild technique. For more information, see 3.5.5, "Smart Rebuild" on page 84. This process does not fail the disk that is being migrated. However, the process reduces the number of available spares in the DS8880 until the migration process is complete.

In a drive intermix on a DA pair, it is possible to rebuild the contents of a 300 GB drive onto a 600 GB spare drive. However, approximately one-half of the 600 GB drive is wasted because that space cannot be used. When the failed 300 GB drive is replaced with a new 300 GB drive, the DS8880 LIC migrates the data back onto the recently replaced 300 GB drive.

When this process completes, the 300 GB DDM rejoins the array and the 600 GB drive becomes the spare again. This same algorithm applies when the hot spare that is taken at the time of the initial drive failure has a speed mismatch.

HPFE does not need to rebalance the spare.

Hot pluggable drives

Replacement of a failed drive does not affect the operation of the DS8880 because the drives are fully hot pluggable. Each drive plugs into a switch, so no loop break is associated with the removal or replacement of a drive. In addition, no potentially disruptive loop initialization process occurs.

Enhanced sparing

The drive sparing policies support having spares for all capacity and speed drives on the DA pair. When any DA pair has only a single spare for any drive type, a call home to IBM is generated. Because of spare over-allocation, several drives can be in a Failed/Deferred Service state. All failed drives are included in the call home when any drive type has one spare. For example, in a configuration with 16 flash drives on a DA pair, two spares are created. If one flash drive fails, all failed drives in the storage system of any type are reported to IBM.

The following DS CLI command can be used to know whether actions can be deferred:

```
lsddm -state not_normal
```

An example of where repair can be deferred is shown in Example 3-2.

Example 3-2   DS CLI lsddm command shows DDM state

```
dscli> lsddm -state not_normal IBM.2107-75DMC71
Date/Time: 14 November 2016 18:18:29 CET IBM DSCLI Version: 7.8.20.233 DS: IBM.2107-75DMC71
ID                            DA Pair dkcap (10^9B) dkuse        arsite State
============================================================================================
IBM.2107-D02-06MZ7/R1-P1-D24 0                600.0 unconfigured S1     Failed/Deferred Service
```

If immediate repair for drives in the Failed/Deferred Service state is needed, an RPQ/SCORE process can be used to submit a request to disable enhanced sparing service. For more information, contact your marketing representative for the details of this RPQ.

3.6 RAS on the power subsystem

All power and cooling components that constitute the DS8880 power subsystem are fully redundant. Key elements that allow this high level of redundancy are two DC-UPSs for each frame for a 2N redundancy. By using this configuration, DC-UPSs are duplicated in each frame so that only one DC-UPS by itself provides enough power for all components inside a frame, if the other DC-UPS becomes unavailable.

As described in "Battery service module sets" on page 92, each DC-UPS has its own battery backup function. Therefore, the battery system in DS8880 also has 2N redundancy. The battery of a single DC-UPS allows the completion of the FHD if a dual AC loss occurs (as described in 3.2.4, "NVS and power outages" on page 74).

The CPCs, I/O enclosures, drive enclosures, and primary HMC components in the frame all feature duplicated power supplies.

Smart internal power distribution connectivity makes it possible to maintain redundant power distribution on a single power cord. If one DC-UPS power cord is pulled (equivalent to a failure in one of the client circuit breakers), the partner DC-UPS can provide power to this UPS and feed each internal redundant power supply inside the frame. For example, if a DC-UPS power cord is pulled, the two redundant power supplies of any CPC continue to be powered on. This ability gives an extra level of reliability in the unusual case of failure in multiple power elements.

In addition, internal Ethernet switches and tray fans, which are used to provide extra cooling to the internal HMC, receive redundant power.

3.6.1 Power components

This section describes the power subsystem components of the DS8880 from an RAS standpoint.

Direct current uninterruptible power supply

Two DC-UPS units exist for each frame for a 2N redundancy. The DC-UPS is a built-in power converter that can monitor power and integrated battery functions. It distributes full wave rectified AC to Power Distribution Units (PDUs), which then provide that power to all areas of the system.

If AC is not present at the input line, the output is switched to rectified AC from the partner DC-UPS. If neither AC input is active, the DC-UPS switches to battery power for up to 4 seconds, or 40 seconds if the ePLD feature is installed. Each DC-UPS has internal fans to supply cooling for that power supply. If AC input power is not restored before the *ride through* time expires, an emergency shutdown results, and FHD copies the data in NVS to the CPC hard disk drives to prevent data loss.

DC-UPS supports high or low voltage single-phase and three-phase as input power. The correct power cables and power select jumper must be used. For information about power cord feature codes, see the *IBM DS8880 Introduction and Planning Guide*, GC27-8525.

All elements of the DC-UPS can be replaced concurrently with client operations. Furthermore, BSM set replacement and DC-UPS fan assembly are performed while the corresponding direct current supply unit remains operational.

The following important enhancements also are available:

► DC-UPS data collection is improved.
► During DC-UPS firmware update, the current power state is maintained so that the DC-UPS remains operational during this service operation. Because of its dual firmware image design, dual power redundancy is maintained in all internal power supplies of all frames during the DC-UPS firmware update.

Each DC-UPS unit consists of one direct current supply unit (DSU) and one BSM set. Figure 2-28 on page 60 shows the DSU (rear view) and BSMs (front view).

> **Important:** If a DS8880 is installed so that both DC-UPSs are attached to the same circuit breaker or the same power distribution unit, the DS8880 is not well-protected from external power failures. This configuration can cause an unplanned outage.

Direct current supply unit

Each DC-UPS has a direct current supply unit (DSU), which contains the control logic of the DC-UPS and it is where images of the power firmware are located. It is designed to protect the DSU from failures during a power firmware update, avoiding physical intervention or hardware replacement, except in a permanent hardware failure.

A DSU contains the necessary battery chargers that are dedicated to monitor and charge the BSM set that is installed in the DC-UPS.

Battery service module sets

The BSM set provides backup power to the system when both AC inputs to a frame are lost. Each DC-UPS supports one BSM set. If the ePLD feature is ordered, no additional BSM set is installed. For more information, see 3.6.3, "Line power fluctuation" on page 93.

A BSM set consists of four battery enclosures. Each of these single-battery enclosures is known as a BSM. A group of four BSMs (battery enclosures) makes up a BSM set. Two types of BSMs are available: The primary BSM and the secondary BSM. The primary BSM is the only BSM with an electrical connector to the DSU and it can be installed only in the top location. This primary BSM is the only BSM that has status LEDs. Three secondary BSMs exist.

The DS8880 BSMs have a fixed working life of five years. The fixed working life is four years if the ePLD feature is installed.

Power Distribution Unit

The PDUs are used to distribute power from the DC-UPS through the power distribution units to the power supplies in drive enclosures, CPCs, I/O enclosures, Ethernet switches, and HMC fans.

Four PDUs are in all frames. A PDU module can be replaced concurrently.

Drive enclosure power supplies

The drive enclosure power supply units provide power for the drives, and they house the cooling fans for the drive enclosure. The fans draw air from the front of the frame, through the drives, and then move it out through the back of the frame. The entire frame cools from front to back, complying with the data center hot aisle/cold aisle cooling strategy. Redundant fans are in each power supply unit, and redundant power supply units are in each drive enclosure. The drive enclosure power supply can be replaced concurrently. Figure 2-23 on page 52 shows a front and rear view of a disk enclosure.

The PDUs for drive enclosures can supply power for 5 - 7 drive enclosures. Each drive enclosure power supply plugs into two separate PDUs, which are supplied from separate DC-UPSs.

CPC power supplies and I/O enclosure power supplies

Each CPC and I/O enclosure has dual redundant power supplies to convert power that is provided by PDUs into the required voltages for that enclosure or complex. Each I/O enclosure and each CPC has its own cooling fans.

Power junction assembly

The power junction assembly (PJA) provides redundant power to HMCs and Ethernet switches.

Rack power control card

RPC cards manage the DS8880 power subsystem and provide control, monitoring, and reporting functions. RPC cards are responsible for receiving DC-UPS status and controlling DC-UPS functions. Two RPC cards are included for redundancy. When one RPC card is unavailable, the remaining RPC card performs all RPC functions.

The following RPC enhancements are available in DS8880 (compared to previous generations of DS8000):

► The DS8880 RPC card contains a faster processor and more parity-protected memory.

► Two different buses are for communication between each RPC card and each CPC. These buses provide redundant paths to have an error recovery capability in a failure of one of the communication paths.

► Each RPC card has two firmware images. If an RPC firmware update fails, the RPC card can still boot from the other firmware image. This design also leads to a reduced period during which one of the RPC cards is unavailable because of an RPC firmware update. Because of the dual firmware image, an RPC card is only unavailable for the time that is required (only a few seconds) to boot from the new firmware image after it is downloaded. Because of this configuration, full RPC redundancy is available during most of the time that is required for an RPC firmware update.

► The RPC cards control power to the attached I/O enclosures. They also monitor environmental components such as power, fans, and temperature for the I/O enclosures. Environmental critical and noncritical conditions can generate EPOW events. Critical events trigger the correct signals from the hardware to the affected components to prevent any data loss without operating system or firmware involvement. Non-critical environmental events are also logged and reported.

3.6.2 Line power loss

The DS8880 uses an area of server memory as NVS. This area of memory is used to hold modified data that is not yet written to the storage drives. If line power fails, meaning that both DC-UPSs in a frame report a loss of AC input power, the DS8880 must protect that data. For a full explanation of the NVS and cache operation, see 3.2, "CPC failover and failback" on page 71.

3.6.3 Line power fluctuation

The DS8880 frames contain BSM sets that protect modified data if dual AC power is lost to the entire frame. If a power fluctuation occurs that causes a momentary interruption to power (often called a *brownout*), the DS8880 tolerates this condition for approximately 4 seconds. If the ePLD feature is not installed on the DS8880 system, the drives are powered off and the servers begin copying the contents of NVS to the internal disks in the CPCs. For many clients who use uninterruptible power supply (UPS) technology, brownouts are not an issue. UPS-regulated power is reliable, so more redundancy in the attached devices is often unnecessary.

Extended power line disturbance

If power at the installation is not always reliable, consider adding the ePLD feature.

Without the ePLD feature, a standard DS8880 offers protection of about 4 seconds from power line disturbances. Installing this feature increases the protection to 40 seconds (running on battery power for 40 seconds) before an FHD begins. For a full explanation of this process, see 3.2.4, "NVS and power outages" on page 74.

3.6.4 Power control

Power control is modified through the HMC, which communicates sequencing information to the service processor in each CPC and RPC. Power control of the DS8880 can be performed by using the Service Maintenance Console Web User Interface (WUI) or by using the DS8880 Storage Management GUI or DS CLI commands.

Figure 3-8 shows the power control window of the Storage Management GUI.

In addition, the following switches in the base frame of a DS8880 are accessible when the rear cover is open:

► Local/remote switch. This switch has two positions: Local and remote.

► Local power on/local force power off switch. When the local/remote switch is in local mode, the local power on/local force power off switch can manually power on or force power off to a complete system. When the local/remote switch is in remote mode, the HMC is in control of power on/power off.

> **Important:** The Local/remote switch must not be used by DS8880 users. The switch can be used only under certain circumstances and as part of an action plan that is performed by an IBM SSR.
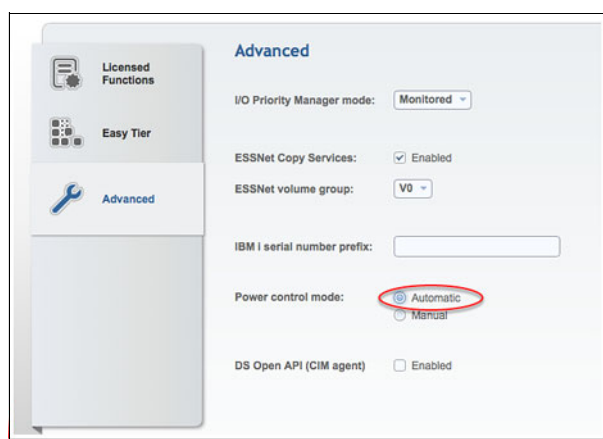
Figure 3-8   DS8880 Modify power control from the Storage Management GUI

3.7 Other features

Many more features of the DS8880 enhance reliability, availability, and serviceability. Several of these features are described in this section.

3.7.1 Internal network

Each DS8880 base frame contains 2-Gigabit Ethernet switches to allow the creation of a fully redundant private management network. Each CPC in the DS8880 has a connection to each switch. The primary HMC (and the secondary HMC, if installed) has a connection to each switch. This configuration means that if a single Ethernet switch fails, all communication can complete from the HMCs to other components in the storage system that are using the alternative private network.

> **Note:** The Ethernet switches are for use that is internal to DS8880 private networks. No external connection to the private networks is allowed. Client connectivity to the DS8880 is allowed only through the provided external Ethernet connector at the rear of the base frame.

3.7.2 Earthquake resistance

The Earthquake Resistance Kit is an optional seismic kit for stabilizing the storage system frame so that the frame complies with IBM earthquake resistance standards. It helps to prevent personal injury and increases the probability that the system will be available following an earthquake by limiting potential damage to critical system components.

Storage system frames with this optional seismic kit include hardware at the bottom of the frame that secures it to the floor. Depending on the flooring in your environment (specifically, non-raised floors), installation of the required floor mounting hardware might be disruptive. This kit must be special-ordered for the DS8880. For more information, contact your IBM sales representative.

3.7.3 Secure Data Overwrite

Secure Data Overwrite (SDO) is a process that provides a secure overwrite of all data drives in a DS8880 series storage system. Removal of all logical configuration and encryption groups are a required client activity before SDO can be run. The SDO process is started by the IBM SSR. The process continues unattended until it is complete. This process takes a full day to complete. Two DDM overwrite options are available.

Drive overwrite options
Two options are available for SDO. This section describes both options. Both options comply with National Institute of Standards and Technology (NIST) Special Publication 800-88r1.

Three-pass overwrite
This option runs a cryptoerase on the drives, then runs a three-pass overwrite on all drives.

Cryptoerase

This option runs a cryptoerase on the drives, which obfuscates the internal encryption key on the drives, rendering the previous information unreadable. It then runs a single-pass overwrite on all drives. Compared to the three-pass overwrite SDO, this option reduces the overall duration of the SDO process.

CPC and HMC

With either option, a three-pass overwrite is run on areas of both the CPC and HMC disk drives that contain any client-related information. If a secondary HMC is associated with the storage system, SDO runs against the secondary HMC after completion on the primary HMC. This SDO detects the previous SDO and runs an overwrite only on the secondary HMC hard disks.

SDO process overview

The basic steps in the SDO process are listed:

1. Client removal of all logical configuration and encryption groups.
2. IBM SSR starts SDO on the HMC.
3. SDO runs a dual cluster restart of the CPCs.
4. SDO cryptoerases all drives in the storage system.
5. SDO starts an overwrite method.
6. SDO starts a three-pass overwrite on the CPC and HMC hard disks.
7. When complete, SDO generates a certificate.

Certificate

The certificate can be offloaded by using DS CLI. If this offload is not possible, the IBM SSR can offload the certificate to removable media.

4

Virtualization concepts

This chapter describes virtualization concepts as they apply to the DS8880.

This chapter covers the following topics:

► Virtualization definition
► Benefits of virtualization
► Abstraction layers for drive virtualization
► Extent pools

4.1 Virtualization definition

For the purposes of this chapter, *virtualization* is the abstraction process from the physical drives to one or more logical volumes that are presented to hosts and systems in a way that they appear as though each were a physical drive.

4.2 Benefits of virtualization

The DS8880 physical and logical architecture provides highly configurable enterprise storage virtualization. DS8880 virtualization includes the following benefits:

► Flexible logical volume configuration:
  – Multiple RAID types (RAID 6, RAID 5, and RAID 10)
  – Storage types (count key data (CKD) and fixed-block architecture (FB)) aggregated into separate extent pools
  – Volumes are allocated from extents of the extent pool
  – Storage pool striping
  – Dynamically add and remove volumes
  – Logical volume configuration states
  – Dynamic Volume Expansion (DVE)
  – Extent space-efficient (ESE) volumes for thin provisioning of FB and CKD volumes
  – Support for small and large extents
  – Extended address volumes (EAVs) (CKD)
  – Parallel Access Volumes across LCUs (Super PAV for CKD)
  – Dynamic extent pool merging for IBM Easy Tier
  – Dynamic volume relocation for Easy Tier
  – Easy Tier Heat Map Transfer
► Flexible logical volume sizes:
  – CKD: up to 1 TB (1,182,006 cylinders) by using EAVs
  – FB: up to 16 TB (limit of 2 TB when used with Copy Services, limit of 1 TB with small extents)
► Flexible number of logical volumes:
  – Up to 65280 (CKD)
  – Up to 65280 (FB)
  – 65280 total for mixed CKD + FB
► No strict relationship between RAID ranks and logical subsystems (LSSs)
► LSS definition allows flexible configuration of the number and size of devices for each LSS:
  – With larger devices, fewer LSSs can be used
  – Volumes for a particular application can be kept in a single LSS
  – Smaller LSSs can be defined, if required (for applications that require less storage)
  – Test systems can have their own LSSs with fewer volumes than production systems

4.3 Abstraction layers for drive virtualization

Virtualization in the DS8880 refers to the process of preparing physical drives for storing data that belongs to a volume that is used by a host. This process allows the host to think that it is using a storage device that belongs to it, but it is really being implemented in the storage system. In open systems, this process is known as creating logical unit numbers (LUNs). In IBM z Systems, it refers to the creation of 3390 volumes.

The basis for virtualization begins with the *physical drives,* which are mounted in storage enclosures and connected to the internal storage servers. To learn more about the drive options and their connectivity to the internal *storage servers*, see 2.5, "Storage enclosures and drives" on page 48.

Virtualization builds upon the physical drives in a series of layers:

► Array sites
► Arrays
► Ranks
► Extent pools
► Logical volumes
► Logical subsystems

4.3.1 Array sites

An *array site* is formed from a group of eight identical drives with the same capacity, speed, and drive class. The specific drives that are assigned to an array site are automatically chosen by the system at installation time to balance the array site capacity across both drive enclosures in a pair, and across the connections to both storage servers. The array site also determines the drives that are required to be reserved as spares. No predetermined storage server affinity exists for array sites.

Figure 4-1 shows an example of the physical representation of array site. The disk drives are dual connected to SAS-FC bridges. An array has drives from two chains.
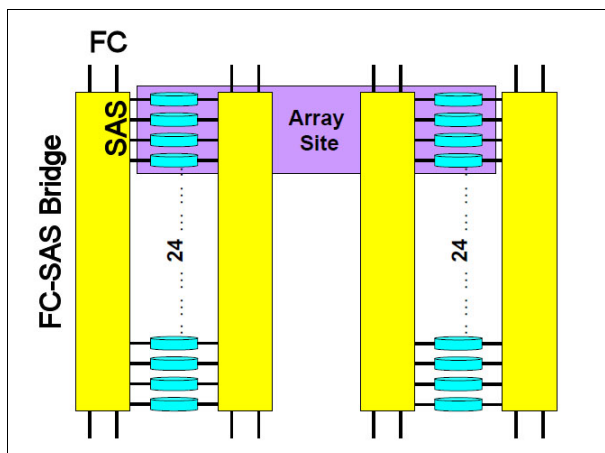
Figure 4-1   Array site across two chains

For the layout logic described here and in the following subchapters, one HPF (Gen2) Enclosure pair can be considered as the equivalent of one standard enclosure pair.

Array sites are the building blocks that are used to define *arrays*.

4.3.2 Arrays

An *array* is created from one *array site.* When an array is created, its Redundant Array of Independent Disks (RAID) level, array type, and array configuration are defined. This process is also called *defining* an array. In all DS8000 series implementations, one array is always defined as using one array site.

The following RAID levels are supported in DS8880:

- ► RAID 6 (preferred default)
- ► RAID 10
- ► RAID 5 (possible for drive sizes below 1 TB)

For more information, see "RAID 5 implementation in DS8880" on page 85, "RAID 6 implementation in the DS8880" on page 87, and "RAID 10 implementation in DS8880" on page 88.

> **Important:** For drive sizes larger than 1 TB, using RAID 6 is highly preferred, and is the default in the DS GUI. RAID 5 is not possible for such large drive capacities. As with large drives, the RAID rebuild times (after one drive failing) get ever larger. Using RAID 6 significantly reduces the danger of data loss due to a double-RAID failure. For additional information, see 3.5.1, "RAID configurations" on page 82.
>
> RAID configuration information changes occasionally. Consult your IBM service support representative (SSR) for the latest information.

According to the sparing algorithm of the DS8880, 0 - 2 spares can be taken from the array site. For more information, see 3.5.9, "Spare creation" on page 89.

Figure 4-2 shows the creation of a RAID 5 array with one spare, which is also called a *6+P+S array.* It has a capacity of six drives for data, a capacity of one drive for parity, and a spare drive. According to the RAID 5 rules, parity is distributed across all seven drives in this example. On the right side of Figure 4-2, the terms, D1, D2, D3, and so on, stand for the set of data that is contained on one drive within a stripe on the array. For example, if 1 GB of data is written, it is distributed across all of the drives of the array.
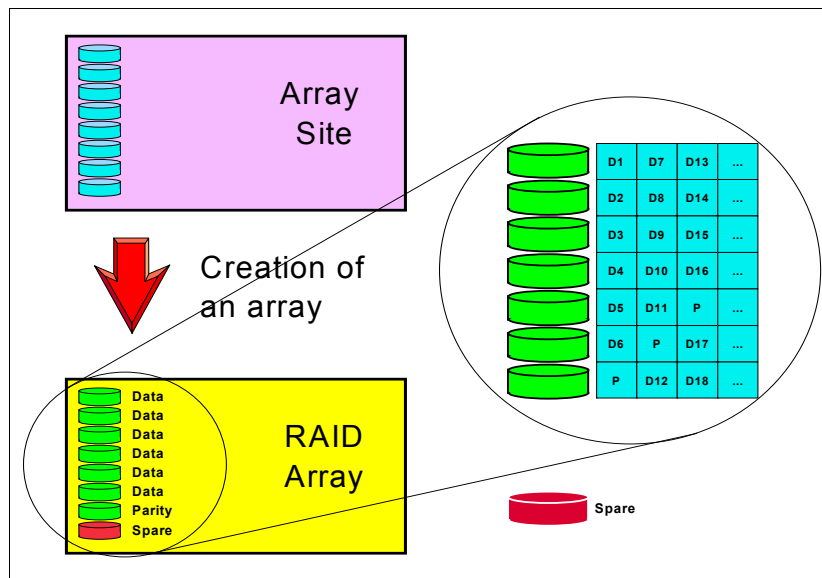
Figure 4-2   Creation of an array

Depending on the selected RAID level and sparing requirements, six types of arrays are possible, as shown in Figure 4-3.
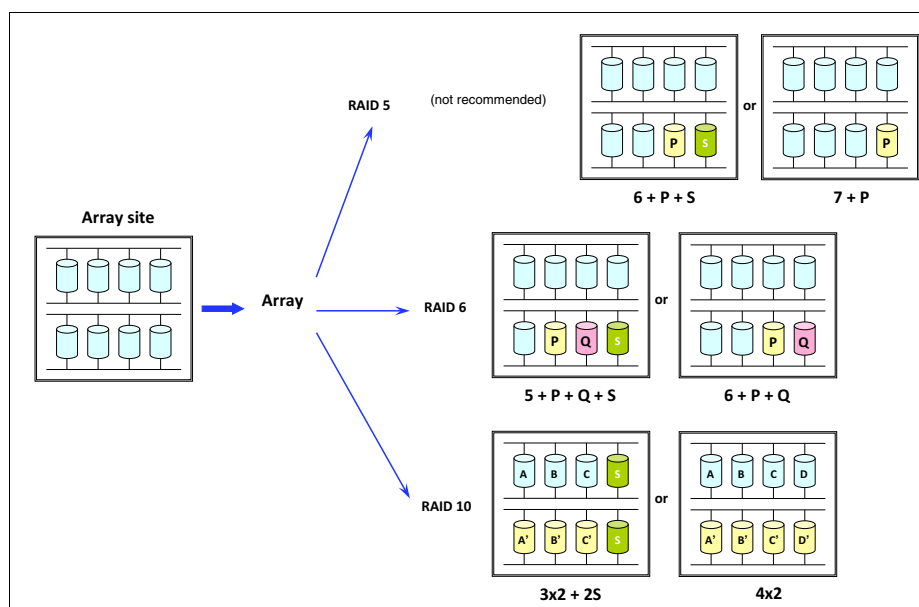
Figure 4-3   RAID array types

> **Tip:** Remember that larger drives have a longer rebuild time. Only RAID 6 can recover from a second error during a rebuild due to the extra parity. RAID 6 is the best choice for systems that require high availability, and is therefore the default.

4.3.3 Ranks

After the arrays are created, the next task is to define a *rank*. A rank is a logical representation of the physical array formatted for use as FB or CKD storage types. In the DS8880, ranks are defined in a one-to-one relationship to arrays. Before you define any ranks, you must decide whether you will encrypt the data.

Encryption group

All drives that are offered in the DS8880 are Full Disk Encryption (FDE)-capable to secure critical data. In the DS8880, the Encryption Authorization license is included in the Base Function license group.

If you plan to use encryption, you must define an *encryption group* before any ranks are created. The DS8880 supports only one encryption group. All ranks must be in this encryption group. The encryption group is an attribute of a rank. Therefore, your choice is to encrypt everything or nothing. If you want to enable encryption later (create an encryption group), all ranks must be deleted and re-created, which means that your data is also deleted.

For more information, see the latest version of *IBM DS8880 Data-at-rest Encryption*, REDP-4500.

Defining ranks

When a new rank is defined, its name is chosen by the data storage (DS) Management graphical user interface (GUI) or data storage command-line interface (DS CLI), for example, R1, R2, or R3. The rank is then associated with an array.

> **Important:** In all DS8000 series implementations, a rank is defined as using only one array. Therefore, rank and array can be treated as synonyms.

The process of defining a rank accomplishes the following objectives:

► The array is formatted for FB data for open systems, or CKD for z Systems data. This formatting determines the size of the set of data that is contained on one drive within a stripe on the array.

► The capacity of the array is subdivided into partitions, which are called *extents*. The extent size depends on the *extent type*, which is FB or CKD. The extents are the building blocks of the logical volumes. An extent is striped across all drives of an array, as shown in Figure 4-4 on page 103, and indicated by the small squares in Figure 4-5 on page 105.

► You can choose between large extents and small extents.

A fixed-block rank features an extent size of either 1 GB (more precisely, GiB, gibibyte, or binary gigabyte, which equals 2^{30} bytes), called *large extents*, or an extent size of 16 MiB, called *small extents*.

z Systems users or administrators typically do not deal with gigabytes or gibibytes. Instead, storage is defined in terms of the original 3390 volume sizes. A 3390 Model 3 is three times the size of a Model 1. A Model 1 features 1113 cylinders, which are about 0.946 GB. The extent size of a CKD rank is one 3390 Model 1, or 1113 cylinders. A 3390 Model 1, or 1113 cylinders is the large extent size for CKD ranks. The small CKD extent size is 21 cylinders. This size corresponds to the z/OS allocation unit for EAV volumes larger than 65520 cylinders where software changes addressing modes and allocates storage in 21 cylinder units.

An extent can be assigned to only one volume. Although you can define a CKD volume with a capacity that is an integral multiple of one cylinder or a fixed-block LUN with a capacity that is an integral multiple of 128 logical blocks (64 KB), if you define a volume this way, you might waste the unused capacity in the last extent that is assigned to the volume.

For example, the DS8880 theoretically supports a minimum CKD volume size of one cylinder. But the volume will still claim one full extent of 1113 cylinders if large extents are used or 21 cylinder for small extents. So, 1112 cylinders are wasted if large extents are used.

> **Note:** An important change in the DS8880 firmware is that all volumes now have a common metadata structure. All volumes now have the metadata structure of ESE volumes, whether the volumes are thin-provisioned or fully provisioned. ESE is described in 4.4.4, "Volume allocation and metadata" on page 117.

Figure 4-4 shows an example of an array that is formatted for FB data with 1-GB extents. The squares in the rank indicate that the extent consists of several blocks from separate drives.
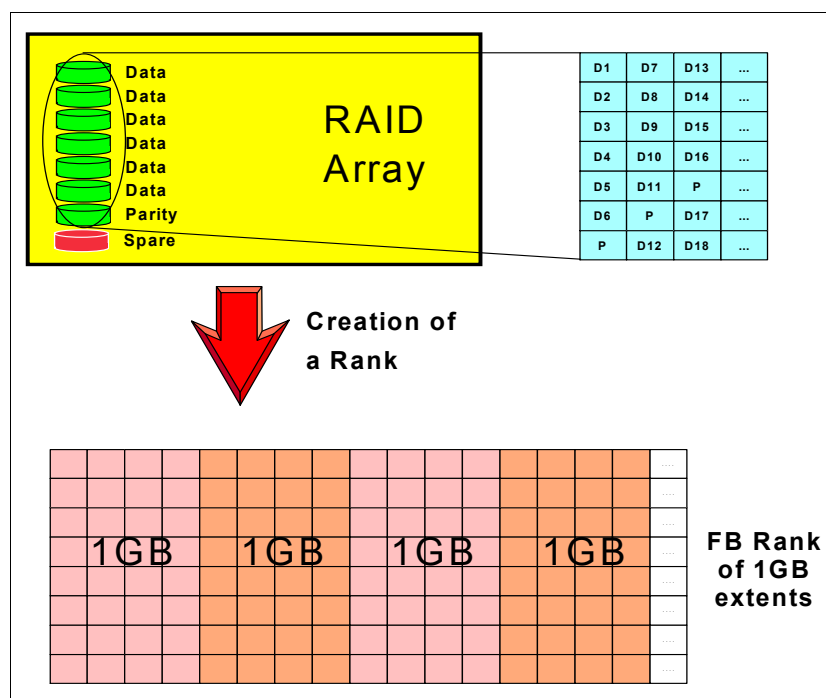
Figure 4-4   Forming an FB rank with 1-GB extents

Small or large extents

Whether to use small or large extents depends on the goals you want to achieve. Small extents provide a better capacity utilization, particularly for thin-provisioned storage. However, managing many small extents causes some small performance degradation during initial allocation. For example, a format write of 1 GB requires one storage allocation with large extents, but 64 storage allocations with small extents. Otherwise, host performance should not be adversely affected.

4.4 Extent pools

An *extent pool* is a logical construct to aggregate the extents from a set of ranks, and it forms a domain for extent allocation to a logical volume. Originally, extent pools were used to separate drives with different revolutions per minute (RPM) and capacity in different pools that have homogeneous characteristics. You still might want to use extent pools for this purpose.

No rank or array affinity to an internal server (central processor complex (CPC)) is predefined. The affinity of the rank (and its associated array) to a server is determined when it is assigned to an extent pool.

One or more ranks with the same extent type (FB or CKD) can be assigned to an extent pool.

> **Important:** Because a rank is formatted to have small or large extents, the first rank that is assigned to an extent pool determines whether the extent pool is a pool of all small or all large extents. You cannot have a pool with a mixture of small and large extents. You cannot change the extent size of an extent pool.

If you want Easy Tier to automatically optimize rank utilization, configure more than one rank in an extent pool. A rank can be assigned to only one extent pool. As many extent pools as ranks can exist.

Heterogeneous extent pools, with a mix of flash cards, flash drives, enterprise serial-attached Small Computer System Interface (SCSI) (SAS) drives, and nearline drives can take advantage of the capabilities of Easy Tier to optimize I/O throughput. Easy Tier moves data across different storage tiering levels to optimize the placement of the data within the extent pool.

With *storage pool striping*, you can create logical volumes that are striped across multiple ranks to enhance performance. To benefit from storage pool striping, more than one rank in an extent pool is required. For more information, see "Storage pool striping: Extent rotation" on page 113.

Storage pool striping can enhance performance significantly. However, in the unlikely event that a whole RAID array fails, the loss of the associated rank affects the entire extent pool because data is striped across all ranks in the pool. For data protection, consider mirroring your data to another DS8000 family storage system.

With the availability of Easy Tier, storage pool striping is somewhat irrelevant. It is a much better approach to let Easy Tier manage extents across ranks (arrays).

When an extent pool is defined, it must be assigned with the following attributes:

► Internal storage server affinity
► Extent type (FB or CKD)
► Encryption group

As with ranks, extent pools are also assigned to encryption group 0 or 1, where group 0 is non-encrypted, and group 1 is encrypted. The DS8880 supports only one encryption group, and all extent pools must use the same encryption setting as was used for the ranks.

A minimum of two extent pools must be configured to balance the capacity and workload between the two servers. One extent pool is assigned to internal server 0. The other extent pool is assigned to internal server 1. In a system with both FB and CKD volumes, four extent pools provide one FB pool for each server and one CKD pool for each server.

If you plan on using small extents for ESE volumes while retaining large extents for other volumes, you must create additional pools with small extents. Indeed, small and large extents cannot be in the same pool.

Figure 4-5 shows an example of a mixed environment that features CKD and FB extent pools.

Extent pools can be expanded by adding more ranks to the pool. All ranks that belong to extent pools with the same internal server affinity are called a *rank group*. Ranks are organized in two rank groups. Rank group 0 is controlled by server 0, and rank group 1 is controlled by server 1.
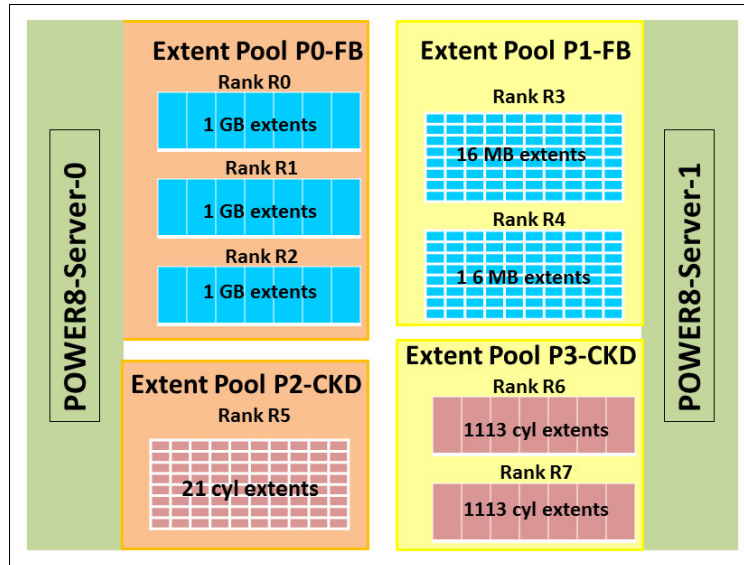
Figure 4-5   Extent pools

4.4.1 Dynamic extent pool merge

Dynamic extent pool merge is a capability that is provided by the Easy Tier manual mode facility. It allows one extent pool to be merged into another extent pool if they meet these criteria:

- ▸ Have extents of the same size
- ▸ Have the same storage type (FB or CKD)
- ▸ Have the same DS8880 internal server affinity

The logical volumes in both extent pools remain accessible to the host systems. Dynamic extent pool merge can be used for the following reasons:

- ▸ Consolidation of two smaller extent pools with the equivalent storage type (FB or CKD) and extent size into a larger extent pool. Creating a larger extent pool allows logical volumes to be distributed over a greater number of ranks, which improves overall performance in the presence of skewed workloads. Newly created volumes in the merged extent pool allocate capacity as specified by the selected extent allocation algorithm. Logical volumes that existed in either the source or the target extent pool can be redistributed over the set of ranks in the merged extent pool by using the Migrate Volume function.

- ▸ Consolidating extent pools with different storage tiers to create a merged extent pool with a mix of storage drive technologies. This type of an extent pool is called a *hybrid pool* and is a prerequisite for using the Easy Tier automatic mode feature.

The Easy Tier manual mode volume migration is shown in Figure 4-6.
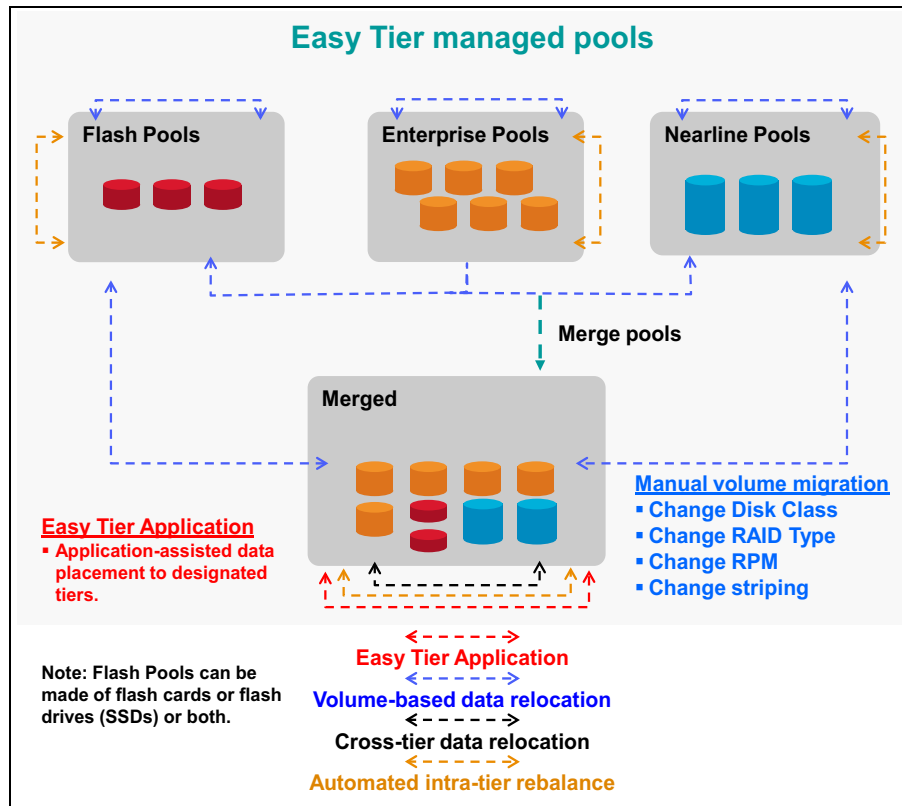
Figure 4-6   Easy Tier migration types

> **Important:** Volume migration (or Dynamic Volume Relocation) within the same extent pool is not supported in hybrid (or multi-tiered) pools. Easy Tier automatic mode rebalances the volumes' extents within the hybrid extent pool automatically based on activity. However, you can also use the Easy Tier application to manually place entire volumes in designated tiers. For more information, see *DS8870 Easy Tier Application*, REDP-5014.

Dynamic extent pool merge is allowed only among extent pools with the same internal server affinity or rank group. Additionally, the dynamic extent pool merge is not allowed in the following circumstances:

► If source and target pools feature different storage types (FB and CKD)

► If source and target pool feature different extent sizes

► If you selected an extent pool that contains volumes that are being migrated

► If the combined extent pools include 2 PB or more of ESE logical capacity (virtual capacity)

For more information about Easy Tier, see *IBM DS8000 EasyTier*, REDP-4667.

4.4.2 Logical volumes

A *logical volume* consists of a set of extents from one extent pool. The DS8880 supports up to 65,280 logical volumes (64 K CKD, or 64 K FB volumes, or a mixture of both, up to a maximum of 64 K total volumes). The abbreviation 64 K is used in this section, even though it is 65,536 minus 256, which is not quite 64 K in binary.

Fixed-block LUNs

A logical volume that is composed of fixed-block extents is called a LUN. A fixed-block LUN is composed of one or more 1 GiB (2^{30} bytes) large extents or one or more 16 MiB small extents from one FB extent pool. A LUN cannot span multiple extent pools, but a LUN can have extents from multiple ranks within the same extent pool. You can construct LUNs up to a size of 16 TiB (16 x 2^{40} bytes, or 2^{44} bytes) when using large extents.

> **Important:** DS8880 Copy Services do not support FB logical volumes larger than 2 TiB (2 x 2^{40} bytes). Do not create a LUN that is larger than 2 TiB if you want to use Copy Services for the LUN, unless the LUN is integrated as Managed Disks in an IBM SAN Volume Controller. Use SAN Volume Controller Copy Services instead.

LUNs can be allocated in binary GiB (2^{30} bytes), decimal GB (10^{9} bytes), or 512 or 520-byte blocks. However, the physical capacity that is allocated is a multiple of 1 GiB. For small extents, it is a multiple of 16 MiB. Therefore, it is a good idea to use LUN sizes that are a multiple of a gibibyte or a multiple of 16 MiB. If you define a LUN with a size that is not a multiple of 1 GiB (for example, 25.5 GiB), the LUN size is 25.5 GiB. However, 26 GiB are physically allocated, of which 0.5 GiB of the physical storage will be unusable. When you want to specify a LUN size that is not a multiple of 1 GiB, specify the number of blocks. A 16 MiB extent has 32768 blocks.

The allocation process for FB volumes is illustrated in Figure 4-7.
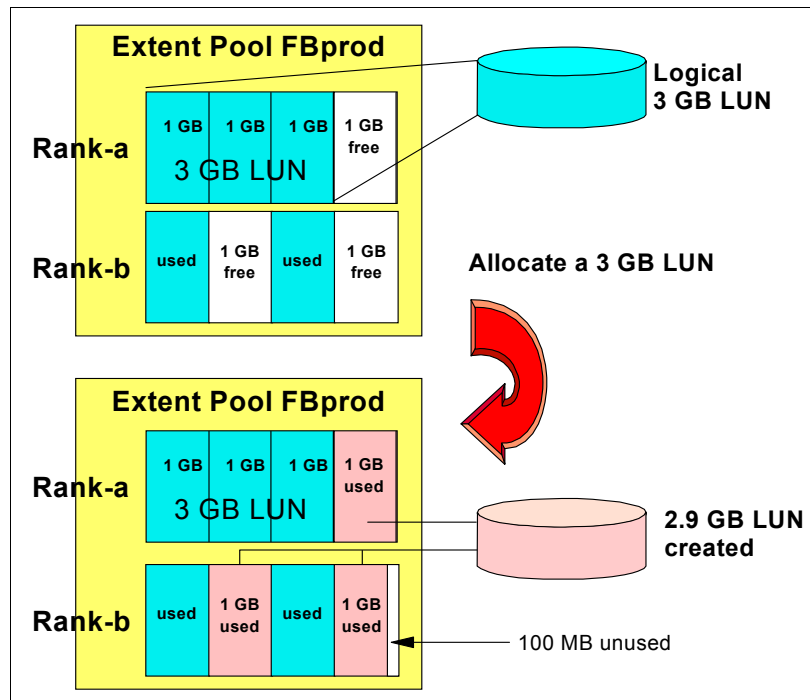
Figure 4-7   Creation of a fixed-block LUN

With small extents, waste of storage is not an issue.

A fixed-block LUN must be managed by an LSS. One LSS can manage up to 256 LUNs. The LSSs are created and managed by the DS8880, as required. A total of 255 LSSs can be created in the DS8880.

IBM i logical unit numbers

IBM i LUNs are also composed of fixed-block 1 GiB (or, 16 MiB) extents. However, special aspects need to be considered with IBM System i LUNs. LUNs that are created on a DS8880 are always RAID-protected. LUNs are based on RAID 10, RAID 6, or RAID 5 arrays. However, you might want to deceive IBM i and tell it that the LUN is not RAID-protected. This deception causes the IBM i to conduct its own mirroring. IBM i LUNs can have the `unprotected` attribute, in which case the DS8880 reports that the LUN is not RAID-protected. This selection of the `protected` or `unprotected` attribute does not affect the RAID protection that is used by the DS8880 on the open volume, however.

IBM i LUNs expose a 520-byte block to the host. The operating system uses eight of these bytes, so the usable space is still 512 bytes like other SCSI LUNs. The capacities that are quoted for the IBM i LUNs are in terms of the 512-byte block capacity, and they are expressed in GB (10^9). Convey these capacities to GiB (2^{30}) when the effective usage of extents that are 1 GiB (2^{30}) are considered.

> **Important:** The DS8880 supports IBM i variable volume (LUN) sizes, in addition to fixed volume sizes.

IBM i volume enhancement adds flexibility for volume sizes and can optimize the DS8880 capacity usage for IBM i environments. For instance, Table 4-1 shows the fixed volume sizes that were supported for the DS8000, and the amount of space that was wasted (when using large 1-GiB extents) because the fixed volume sizes did not match an exact number of GiB extents.

Table 4-1   IBM i fixed volume sizes

Model type		IBM i device size (GB)	Number of logical block addresses (LBAs)	Extents (large/ small extents)	Unusable space (GiB[a] / w/ large extents)	Usable space % (w/ large extents)
Unprotected	**Protected**					
2107-A81	2107-A01	8.6	16,777,216	8 / 512	0.00	100.00
2107-A82	2107-A02	17.5	34,275,328	17 / 1046	0.66	96.14
2107-A85	2107-A05	35.2	68,681,728	33 / 2096	0.25	99.24
2107-A84	2107-A04	70.6	137,822,208	66 / 4206	0.28	99.57
2107-A86	2107-A06	141.1	275,644,416	132 / 8412	0.56	99.57
2107-A87	2107-A07	282.3	551,288,832	263 / 16824	0.13	99.95

a. GiB represents "binary gigabytes" (2^{30} bytes), and GB represents "decimal gigabytes" (10^9 bytes).

The DS8880 supports IBM i variable volume data types A50, an unprotected variable size volume, and A99, a protected variable size volume. See Table 4-2.

Table 4-2   System i variable volume sizes

Model type		IBM i device size (GB)	Number of logical block addresses (LBAs)	Extents	Unusable space (GiB[1])	Usable space%
Unprotected	**Protected**					
2107-050	2107-099	**Variable**			0.00	**Variable**

Example 4-1 demonstrates the creation of both a protected and an unprotected IBM i variable size volume by using the DS CLI.

Example 4-1   Creating the System i variable size for unprotected and protected volumes

```
dscli> mkfbvol -os400 050 -extpool P4 -name itso_iVarUnProt1 -cap 10 5413
CMUC00025I mkfbvol: FB volume 5413 successfully created.

dscli> mkfbvol -os400 099 -extpool P4 -name itso_iVarProt1 -cap 10 5417
CMUC00025I mkfbvol: FB volume 5417 successfully created.
```

> **Important:** The creation of IBM i variable size volumes is only supported by using DS CLI commands. Currently, no support exists for this task in the GUI.

When you plan for new capacity for an existing IBM i system, the larger the LUN, the more data it might have, which causes more input/output operations per second (IOPS) to be driven to it. Therefore, mixing different drive sizes within the same system might lead to hot spots.

> **Note:** IBM i fixed volume sizes continue to be supported in current and future DS8880 code levels. Consider the best option for your environment between fixed and variable size volumes.

T10 data integrity field support

The ANSI T10 standard provides a way to check the integrity of data that is read and written from the application or the host bus adapter (HBA) to the drive and back through the SAN fabric. This check is implemented through the data integrity field (DIF) that is defined in the T10 standard. This support adds protection information that consists of cyclic redundancy check (CRC), LBA, and host application tags to each sector of FB data on a logical volume.

A T10 DIF-capable LUN uses 520-byte sectors instead of the common 512-byte sector size. Eight bytes are added to the standard 512-byte data field. The 8-byte DIF consists of 2 bytes of CRC data, a 4-byte Reference Tag (to protect against misdirected writes), and a 2-byte Application Tag for applications that might use it.

On a write, the DIF is generated by the HBA, which is based on the block data and LBA. The DIF field is added to the end of the data block, and the data is sent through the fabric to the storage target. The storage system validates the CRC and Reference Tag and, if correct, stores the data block and DIF on the physical media. If the CRC does not match the data, the data was corrupted during the write. The write operation is returned to the host with a write error code. The host records the error and retransmits the data to the target. In this way, data corruption is detected immediately on a write, and the corrupt data is never committed to the physical media.

On a read, the DIF is returned with the data block to the host, which validates the CRC and Reference Tags. This validation adds a small amount of latency for each I/O, but it might affect overall response time on smaller block transactions (less than 4 KB I/Os).

The DS8880 supports the T10 DIF standard for FB volumes that are accessed by the Fibre Channel Protocol (FCP) channels that are used by Linux on z Systems, or AIX. You can define LUNs with an option to instruct the DS8880 to use the CRC-16 T10 DIF algorithm to store the data.

You can also create T10 DIF-capable LUNs for operating systems that do not yet support this feature (except for IBM System i). However, active protection is available only for Linux on z Systems and AIX on Power Systems.

When you create an FB LUN with the `mkfbvol` DS CLI command, add the option `-t10dif`. If you query a LUN with the `showfbvol` command, the data type is shown to be FB 512T instead of the standard FB 512 type.

> **Important:** Because the DS8880 internally always uses 520-byte sectors (to be able to support IBM i volumes), no capacity is considered when standard or T10 DIF-capable volumes are used.
>
> **Target LUN:** When FlashCopy for a T10 DIF LUN is used, the target LUN must also be a T10 DIF-type LUN. This restriction does not apply to mirroring.

Count key data volumes

A z Systems CKD volume is composed of one or more extents from one CKD extent pool. CKD extents are of the size of 3390 Model 1, which features 1113 cylinders for large extents or 21 cylinders for small extents. When you define a z Systems CKD volume, specify the size of the volume as a multiple of 3390 Model 1 volumes or the number of cylinders that you want for the volume.

Before a CKD volume can be created, an LCU must be defined that provides up to 256 possible addresses that can be used for CKD volumes. Up to 255 LCUs can be defined. For more information about LCUs, which also are called LSSs, see 4.4.5, "Logical subsystems" on page 122.

On a DS8880, you can define CKD volumes with up to 1,182,006 cylinders, or about 1 TB. This volume capacity is called an EAV, and it is supported by the 3390 Model A.

A CKD volume cannot span multiple extent pools, but a volume can have extents from different ranks in the same extent pool. You also can stripe a volume across the ranks. For more information, see "Storage pool striping: Extent rotation" on page 113.

Figure 4-8 shows an example of how a logical volume is allocated with a CKD volume.
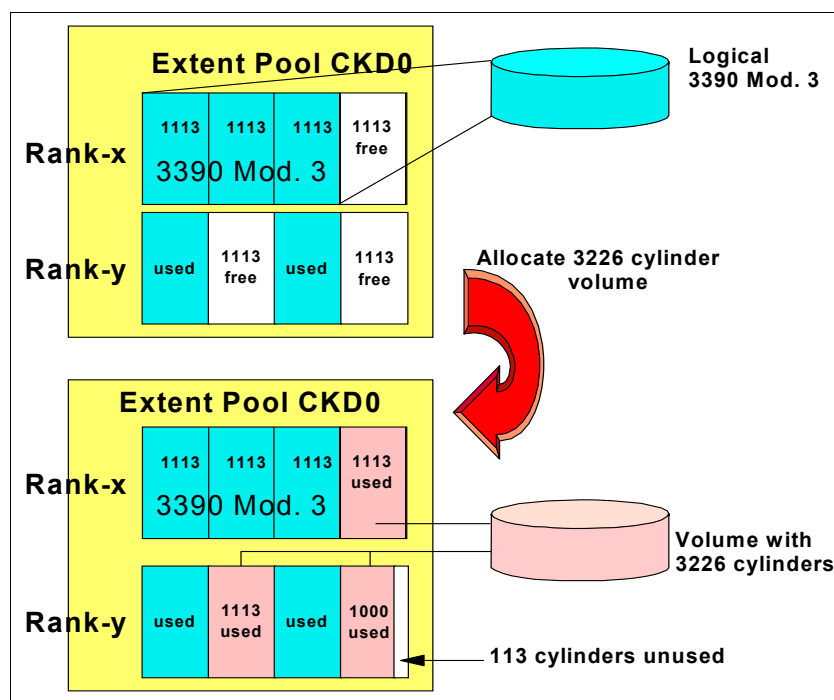
Figure 4-8   Allocation of a count-key-data logical volume

CKD alias volumes

Another type of CKD volume is the PAV *alias volume*. PAVs are used by z/OS to send parallel I/Os to the same base CKD volume. Alias volumes are defined within the same LCU as their corresponding base volumes. Although they have no size, each alias volume needs an address, which is linked to a base volume. The total of base and alias addresses cannot exceed the maximum of 256 for an LCU.

SuperPAV

With HyperPAV, the alias can be assigned to access a base device within the same LCU on an I/O basis. However, the system can exhaust the alias resources that are needed within a LCU and cause an increase in IOS queue time (queue time in software waiting for a device to issue to the I/O). In this case, you need more aliases to access hot volumes.

Classically, to start an I/O to a base volume, z/OS could select any alias address only from the same LCU as the base address to perform the I/O. Starting with DS8880 R8.1 code, the concept of *SuperPAV* or cross LCU HyperPAV alias device is introduced. With SuperPAV support in DS8880 and in z/OS (check for SuperPAV APARs, support is introduced with APARs OA49090, OA49110), the operating system can use alias addresses from *other LCUs* to perform an I/O for a base address.

The restriction is that the LCU of the alias address belongs to the same DS8880 server. In other words, if the base address is from an even/odd LCU, the alias address that z/OS can select must also be from an even/odd LCU. In addition, the LCU of the base volume and the LCU of the alias volume must be in the same path group. z/OS prefers alias addresses from the same LCU as the base address, but if no alias address is free, z/OS looks for free alias addresses in LCUs of the same *Alias Management Group*.

An Alias Management Group is all LCUs that have affinity to the same DS8880 internal server and that have the same paths to the DS8880. SMF and IBM RMF™ (APAR OA49415) are enhanced to provide reports at the Alias Management Group level.

Figure 4-9 shows how SuperPAV works.
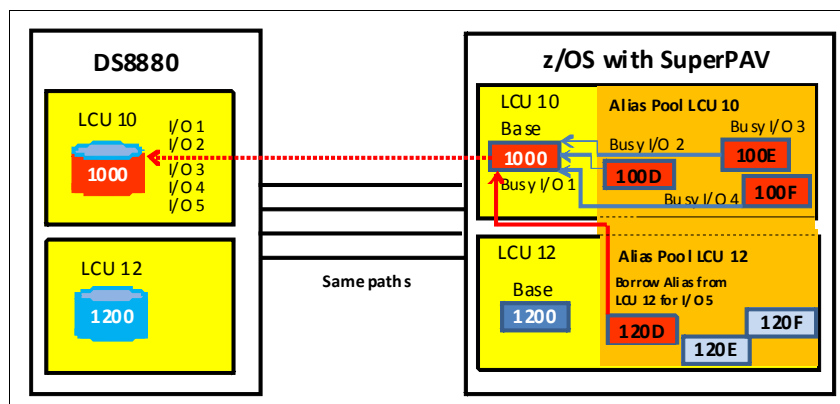
Figure 4-9   SuperPAV

It is still a requirement that initially each alias address is assigned to a base address. Therefore, it is not possible to define an LCU with only alias addresses.

As with PAV and HyperPAV, SuperPAV must be enabled. SuperPAV is enabled by the HYPERPAV=XPAV statement in the IECIOSxx parmlib member or by the SETIOS HYPERPAV=XPAV command. The `D M=DEV(address)` and the `D M=CU(address)` display commands show whether XPAV is enabled or not. With the `D M=CU(address)` command, you can check whether aliases from other LCUs are currently being used (see Example 4-2).

Example 4-2   Display information about SuperPAV (XPAV)

```
D M=CU(4E02)
IEE174I 11.26.55 DISPLAY M 511
CONTROL UNIT 4E02
CHP                   65   5D   34   5E
ENTRY LINK ADDRESS    98   ..   434B ..
DEST LINK ADDRESS     FA   0D   200F 0D
CHP PHYSICALLY ONLINE Y    Y    Y    Y
PATH VALIDATED        Y    Y    Y    Y
MANAGED               N    N    N    N
ZHPF - CHPID          Y    Y    Y    Y
ZHPF - CU INTERFACE   Y    Y    Y    Y
MAXIMUM MANAGED CHPID(S) ALLOWED = 0
DESTINATION CU LOGICAL ADDRESS = 56
CU ND             = 002107.981.IBM.75.0000000FXF41.0330
CU NED            = 002107.981.IBM.75.0000000FXF41.5600
TOKEN NED         = 002107.900.IBM.75.0000000FXF41.5600
FUNCTIONS ENABLED = ZHPF, XPAV
XPAV CU PEERS     = 4802, 4A02, 4C02, 4E02
DEFINED DEVICES
  04E00-04E07
DEFINED PAV ALIASES
  14E40-14E47
```

With cross-LCU HyperPAV, called SuperPAV, the number of alias addresses can further be reduced while the pool of available alias addresses to handle I/O bursts to volumes is increased.

4.4.3 Allocation, deletion, and modification of LUNs and CKD volumes

Extents of ranks that are assigned to an extent pool are independently available for allocation to logical volumes. The extents for a LUN or volume are logically ordered, but they do not have to come from one rank. The extents do not have to be contiguous on a rank.

This construction method of using fixed extents to form a logical volume in the DS8880 allows flexibility in the management of the logical volumes. You can delete LUNs or CKD volumes, resize LUNs or volumes, and reuse the extents of those LUNs to create other LUNs or volumes, including ones of different sizes. One logical volume can be removed without affecting the other logical volumes that are defined on the same extent pool.

The extents are cleaned after you delete a LUN or CKD volume. The reformatting of the extents is a background process, and it can take time until these extents are available for reallocation.

Two extent allocation methods (EAMs) are available for the DS8000: *Storage pool striping* (rotate extents) and *rotate volumes*.

Storage pool striping: Extent rotation

The storage allocation method is chosen when a LUN or volume is created. The extents of a volume can be striped across several ranks. An extent pool with more than one rank is needed to use this storage allocation method.

The DS8880 maintains a sequence of ranks. The first rank in the list is randomly picked at each power-on of the storage system. The DS8880 tracks the rank in which the last allocation started. The allocation of the first extent for the next volume starts from the next rank in that sequence.

The next extent for that volume is taken from the next rank in sequence, and so on. Therefore, the system rotates the extents across the ranks, as shown in Figure 4-10.

Figure 4-10   Rotate extents

> **Note:** Although the preferred storage allocation method was storage pool striping, it is now a better choice to let Easy Tier manage the storage pool extents. This chapter describes rotate extents for the sake of completeness, but it is now mostly irrelevant.

Rotate volumes allocation method

Extents can be allocated sequentially. In this case, all extents are taken from the same rank until enough extents are available for the requested volume size or the rank is full. In this case, the allocation continues with the next rank in the extent pool.

If more than one volume is created in one operation, the allocation for each volume starts in another rank. When several volumes are allocated, rotate through the ranks, as shown in Figure 4-11.
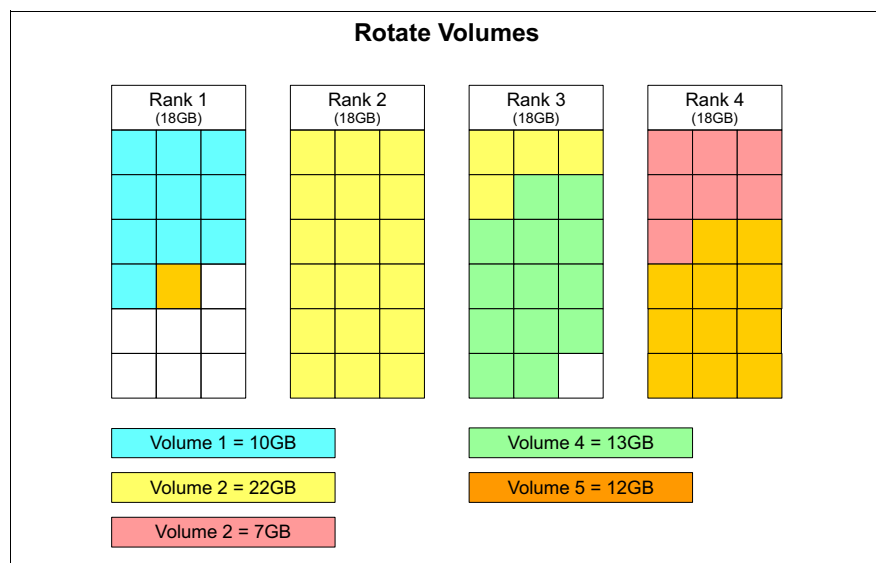
Figure 4-11   Rotate volumes

You might want to consider this allocation method when you prefer to manage performance manually. The workload of one volume is going to one rank. This configuration makes the identification of performance bottlenecks easier. However, by putting all the volumes' data onto one rank, you might introduce a bottleneck, depending on your actual workload.

> **Important:** Rotate extents and rotate volume EAMs provide distribution of volumes over ranks. Rotate extents run this distribution at a granular (1 GiB or 16 MiB extent) level, and is a better method to minimize hot spots and improve overall performance.
>
> However, as previously stated, Easy Tier is really the preferred choice for managing the storage pool extents.

In a mixed-drive-characteristic (or *hybrid*) extent pool that contains different *classes* (or *tiers*) of ranks, the storage pool striping EAM is used independently of the requested EAM, and EAM is set to `managed`.

For extent pools that contain flash ranks, extent allocation is performed initially on enterprise or nearline ranks while space remains available. Easy Tier algorithms migrate the extents as needed to flash ranks. For extent pools that contain a mix of enterprise and nearline ranks, initial extent allocation is performed on enterprise ranks first.

This is the default allocation rule for DS8000 code. Since the DS8880 R8.1 code, there is an option for the whole Storage Facility to change the allocation preference to start the allocation at the highest flash/SSD tier. The `chsi` command can be used to change the `ETTierOrder` parameter to *High Utilization* or to *High Performance*.

When you create striped volumes and non-striped volumes in an extent pool, a rank might be filled before the others. A full rank is skipped when you create new striped volumes.

By using striped volumes, you distribute the I/O load of a LUN or CKD volume to more than one set of eight drives, which can enhance performance for a logical volume. In particular, operating systems that do not include a volume manager with striping capability benefit most from this allocation method.

Small extents can increase the parallelism of sequential writes. While the system stays within one rank until 1 GiB has been written, with small extents, we jump to the next rank after 16 MiB. This configuration uses more disk drives when performing sequential writes.

> **Important:** If you must add capacity to an extent pool because it is nearly full, it is better to add several ranks at the same time, not just one rank. This method allows new volumes to be striped across the newly added ranks.
>
> With the Easy Tier manual mode facility, if the extent pool is a non-hybrid pool, the user can request an extent pool merge followed by a volume relocation with striping to run the same function. For a hybrid-managed extent pool, extents are automatically relocated over time, according to performance needs. For more information, see *IBM DS8000 EasyTier*, REDP-4667.
>
> **Rotate volume EAM:** The rotate volume EAM is not allowed if one extent pool is composed of flash drives and configured for virtual capacity.

Dynamic volume expansion

The DS8880 supports expanding the size of a LUN or CKD volume without data loss, by adding extents to the volume. The operating system must support resizing.

A logical volume includes the attribute of being striped across the ranks or not. If the volume was created as striped across the ranks of the extent pool, the extents that are used to increase the size of the volume are striped. If a volume was created without striping, the system tries to allocate the additional extents within the same rank that the volume was created from originally.

> **Important:** Before you can expand a volume, you must delete any Copy Services relationship that involves that volume.

Because most operating systems have no means of moving data from the end of the physical drive off to unused space at the beginning of the drive, and because of the risk of data corruption, IBM does not support shrinking a volume. The DS8880 DS CLI and DS GUI interfaces will not allow reducing the size of a volume.

Dynamic volume migration

Dynamic volume migration or dynamic volume relocation (DVR) is a capability that is provided as part of the Easy Tier manual mode facility.

DVR allows data that is stored on a logical volume to be migrated from its currently allocated storage to newly allocated storage while the logical volume remains accessible to attached hosts.

The user can request DVR by using the Migrate Volume function that is available through the DS Storage Manager GUI (DS GUI) or DS CLI. DVR allows the user to specify a target extent pool and an EAM. The target extent pool can be a separate extent pool than the extent pool where the volume is located. It can also be the same extent pool, but only if it is a non-hybrid or single-tier pool. However, the target extent pool must be managed by the same DS8880 internal server.

> **Important:** DVR in the same extent pool is not allowed in the case of a managed pool. In managed extent pools, Easy Tier automatic mode automatically relocates extents within the ranks to allow performance rebalancing. The Easy Tier application can also be used to manually place volumes in designated tiers within a managed pool. For more information, see *DS8870 Easy Tier Application*, REDP-5014.
>
> In the current implementation (DS8880 R8.2 code), you can only move volumes between pools of the *same extent size*.

Dynamic volume migration provides the following capabilities:

► The ability to change the extent pool in which a logical volume is provisioned. This ability provides a mechanism to change the underlying storage characteristics of the logical volume to include the drive class (Flash, Enterprise, or Nearline), drive RPM, and RAID array type. Volume migration can also be used to migrate a logical volume into or out of an extent pool.

► The ability to specify the extent allocation method for a volume migration that allows the extent allocation method to be changed between the available extent allocation methods any time after volume creation. Volume migration that specifies the rotate extents EAM can also be used (in non-hybrid extent pools) to redistribute a logical volume's extent allocations across the existing ranks in the extent pool if more ranks are added to an extent pool.

Each logical volume has a configuration state. To begin a volume migration, the logical volume initially must be in the normal configuration state.

More functions are associated with volume migration that allow the user to pause, resume, or cancel a volume migration. Any or all logical volumes can be requested to be migrated at any time if available capacity is sufficient to support the reallocation of the migrating logical volumes in their specified target extent pool. For more information, see *IBM DS8000 EasyTier*, REDP-4667.

4.4.4 Volume allocation and metadata

> **Important:** Compared to earlier DS8000 models, the DS8880 introduces a new internal data layout for all volumes.

In earlier DS8000 systems, metadata allocation differed between fully provisioned or thin provisioned (extent space efficient (ESE)) volumes:

► For fully provisioned volumes, metadata was contained in system-reserved areas of the physical arrays, outside of the client data logical extent structure.

Figure 4-12 illustrates the extent layout in the DS8870 and earlier DS8000 models, for standard, fully provisioned volumes. Note the size of the reserved area (not to scale), which includes volume metadata. The rest of the space is composed of 1 GB extents.
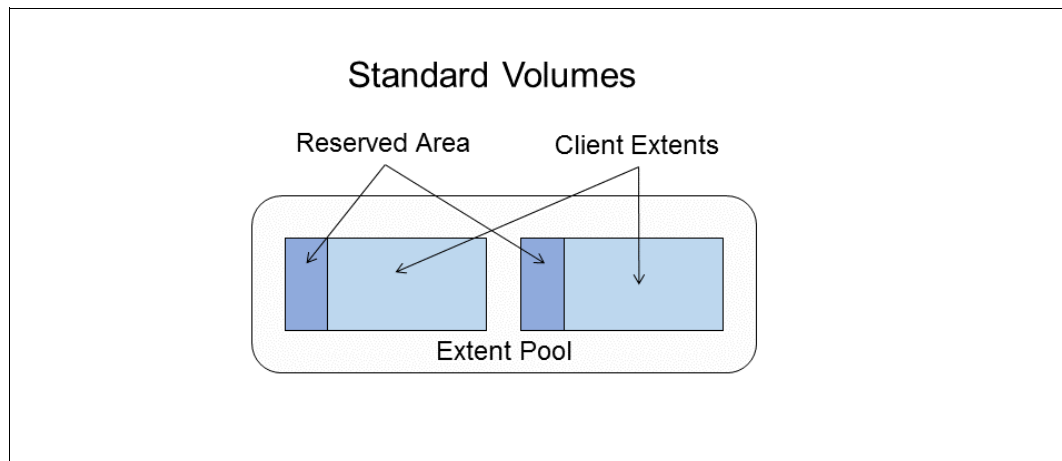
Figure 4-12   DS8000 standard volumes

► For ESE volumes, several logical extents, which are designated as *auxiliary rank extents*, are allocated to contain volume metadata.

Figure 4-13 shows the extent layout in the DS8870 and earlier, for ESE volumes. In addition to the reserved area, auxiliary 1 GB extents are also allocated to store metadata for the ESE volumes.
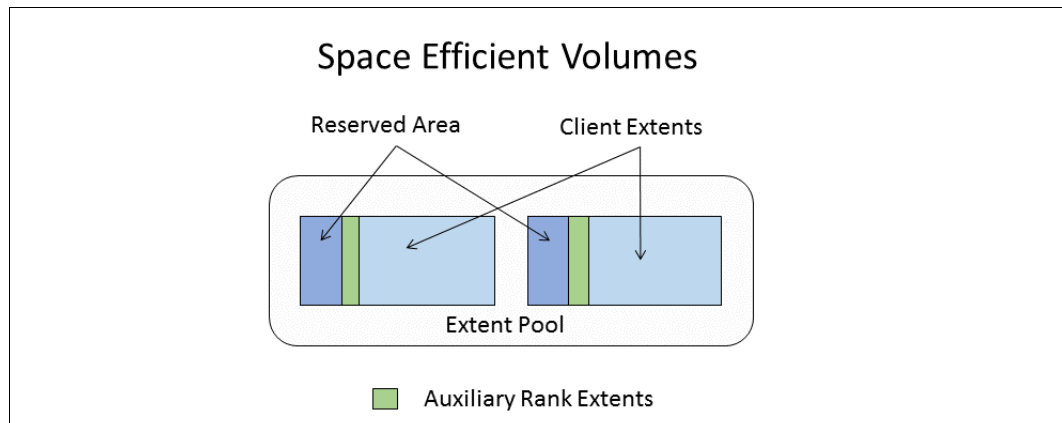
Figure 4-13   DS8870 extent space-efficient volumes

In the DS8880, the logical extent structure changed to move to a single volume structure for all volumes. All volumes now use an improved metadata extent structure, similar to what was previously used for ESE volumes. This unified extent structure greatly simplifies the internal management of logical volumes and their metadata. No area is fixed and reserved for volume metadata, and this capacity is added to the space that is available for use.

The metadata is allocated in the Storage Pool when volumes are created, and the space that is used by metadata is referred to as the *Pool Overhead*. Pool Overhead means that the amount of space that can be allocated by volumes is variable and depends on both the number of volumes and the logical capacity of these volumes.

> **Tips:** The following characteristics apply to the DS8880:
>
> ► Metadata is variable in size. Metadata is used only as logical volumes are defined.
>
> ► The size of the metadata for a logical volume depends on the number and size of the logical volumes that are being defined.
>
> ► As volumes are defined, the metadata is immediately allocated and reduces the size of the available user extents in the storage pool.
>
> ► As you define new logical volumes, you must allow for the fact that this metadata space continues to use user extent space.

For storage pools with large extents, metadata is also allocated as large extents (1 GiB for FB pools or 1113 cylinder for CKD pools). Large extents that are allocated for metadata are subdivided into 16 MiB subextents, which are also referred to as *metadata extents*, for FB volumes, or 21 cylinders for CKD. For extent pools with small extents, metadata extents are also small extents. Sixty-four metadata subextents are in each large metadata extent for FB, and 53 metadata subextents are in each large metadata extent for CKD.

For each FB volume that is allocated, an initial 16 MiB metadata subextent or metadata small extent is allocated, and an additional 16 MiB metadata subextent or metadata small extent is allocated for every 10 GiB of allocated capacity or portion of allocated capacity.

For each CKD volume that is allocated, an initial 21 cylinders metadata subextent or metadata small extent is allocated, and an additional 21 cylinders metadata subextent or metadata small extent is allocated for every 11130 cylinders (or 10 3390 Model 1) of allocated capacity or portion of allocated capacity.

For example, a 3390-3 (that is, 3339 cylinders or about 3 GB) or 3390-9 (that is, 10,017 cylinders or 10 GB) volume takes two metadata extents (one metadata extent for the volume and another metadata extent for any portion of the first 10 GB). A 128 GB FB volume takes 14 metadata extents (one metadata extent for the volume and another 13 metadata extents to account for the 128 GB).

Figure 4-14 shows an illustration of 3 GB and 12 GB FB volumes for a storage pool with large extents. In an extent pool with small extents, there is no concept of sub-extents. You just have user extents, unused extents, and metadata extents.
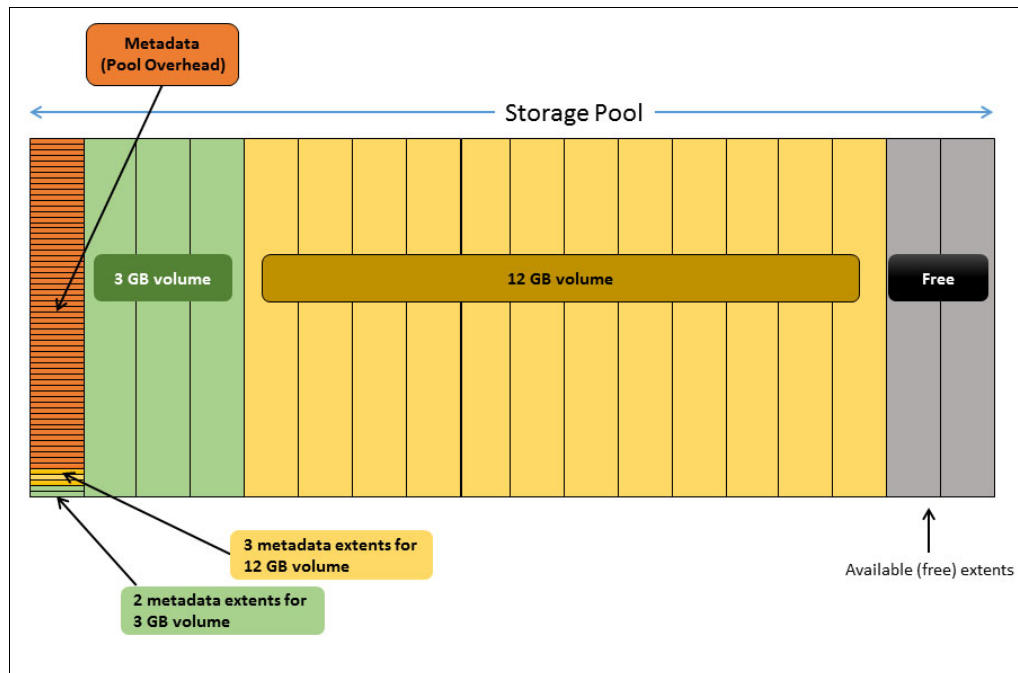
Figure 4-14   Metadata allocation

Metadata extents with free space can be used for metadata by any volume in the extent pool.

> **Important:** In a multitier extent pool, volume metadata is created in the highest available tier. Generally, have some flash/SSD storage for the volume metadata.
>
> Typically, ten percent of flash storage in an extent pool results in all volume metadata being stored on flash, assuming no overprovisioning. With overprovisioning, a corresponding larger percentage of flash would be required.

With this concept of metadata extents and user extents within an extent pool, some planning and calculations are required. This is particularly true in a mainframe environment where often thousands of volumes are defined and often the whole capacity is allocated during the initial configuration. You have to calculate the capacity used up by the metadata to get the capacity that can be used for user data. This is only important when fully provisioned volumes will be used. Thin-provisioned volumes consume no space when created, only metadata and space will be consumed when data is actually written.

For extent pools with small extents, the number of available user data extents can be estimated as follows:

```
(user extents) = (pool extents) - (number of volumes) - (total virtual
capacity)/10
```

For extent pools with regular 1 GiB extents, when the details of the volume configuration are not known, you can estimate the number of metadata extents based on many volumes only. The calculations are performed as shown:

► FB Pool Overhead = (Number of volumes × 2 + total volume extents/10)/64 and rounded up to the nearest integer

► CKD Pool Overhead = (Number of volumes × 2 + total volume extents/10)/53 and rounded up to the nearest integer

This formula overestimates the space that is used by the metadata by a small amount because it assumes wastage on every volume. However, the precise size of each volume does not need to be known.

Two examples are explained:

► A CKD storage pool has 6,190 extents in which you expect to allocate all of the space on 700 volumes. Your Pool Overhead is 39 extents by using this calculation:

`(700 x 2 + 6190/10)/53 = 38.09`

► An FB storage pool has 6,190 extents in which you expect to use thin provisioning and allocate up to 12,380 extents (2:1 overprovisioning) on 100 volumes. Your Pool Overhead is 23 extents by using this calculation:

`(100 x 2 + 12380/10)/64 = 22.46`

Space for Easy Tier

The Easy Tier function of DS8000 monitors access to extents transparently to host applications. Easy Tier moves extents to higher or lower tiers depending on the access frequency or heat of an extent. To be able to move around extents, some free space or free extents must be available in an extent pool. You can manually have an eye on this, but a better option is to let the system reserve some space for Easy Tier extent movements. The `chsi` command can be used to change the `ETSRMode` parameter to `Enabled`. This setting causes the storage system to reserve some capacity for Easy Tier. Obviously, when this mode is enabled, capacity for new volume allocations is reduced.

Extent space efficient volumes

Volumes can be created as thin-provisioned or fully provisioned.

Space for a thin-provisioned volume is allocated when a write occurs. More precisely, it is allocated when a destage from the cache occurs and insufficient free space is left on the currently allocated extent.

Therefore, thin provisioning allows a volume to exist that is larger than the physical space in the extent pool that it belongs to. This approach allows the "host" to work with the volume at its defined capacity, even though insufficient physical space might exist to fill the volume with data.

The assumption is that either the volume will never be filled, or as the DS8880 runs low on physical capacity, more will be added. This approach also assumes that the DS8880 is not at its maximum capacity.

Thin-provisioned volume support is contained in the Base Function license group. This provisioning is an attribute that you specify when creating a volume.

Important: The DS8880 does not support track space efficient (TSE) volumes.

If thin provisioning is used, the metadata is allocated for the entire volume (virtual capacity) when the volume is created not when extents are used.

CKD Thin Provisioning

Thin Provisioning is also possible with CKD volumes. See *IBM DS8880 Thin Provisioning, REDP-5343,* and check for APARs that are needed for certain z/OS versions.

Extent space efficient capacity controls for thin provisioning

Use of thin provisioning can affect the amount of storage capacity that you choose to order. Use ESE capacity controls to allocate storage correctly.

With the mixture of thin-provisioned (ESE) and fully provisioned (non-ESE) volumes in an extent pool, a method is needed to dedicate part of the extent-pool storage capacity for ESE user data usage, and to limit the ESE user data usage within the extent pool.

Also, you need to detect when the available storage space within the extent pool for ESE volumes is running out of space.

ESE capacity controls provide extent pool attributes to limit the maximum extent pool storage that is available for ESE user data usage. They also guarantee that a proportion of the extent pool storage will be available for ESE user data usage.

The following controls are available to limit the use of extents in an extent pool:

► Reserve capacity in an extent pool by enabling the extent pool limit function with the `chextpool -extentlimit enable -limit extent_limit_percentage pool_ID` command.

► You can reserve space for the sole use of ESE volumes by creating a repository with the `mksestg -repcap capacity pool_id` command.

Figure 4-15 shows the areas in an extent pool.
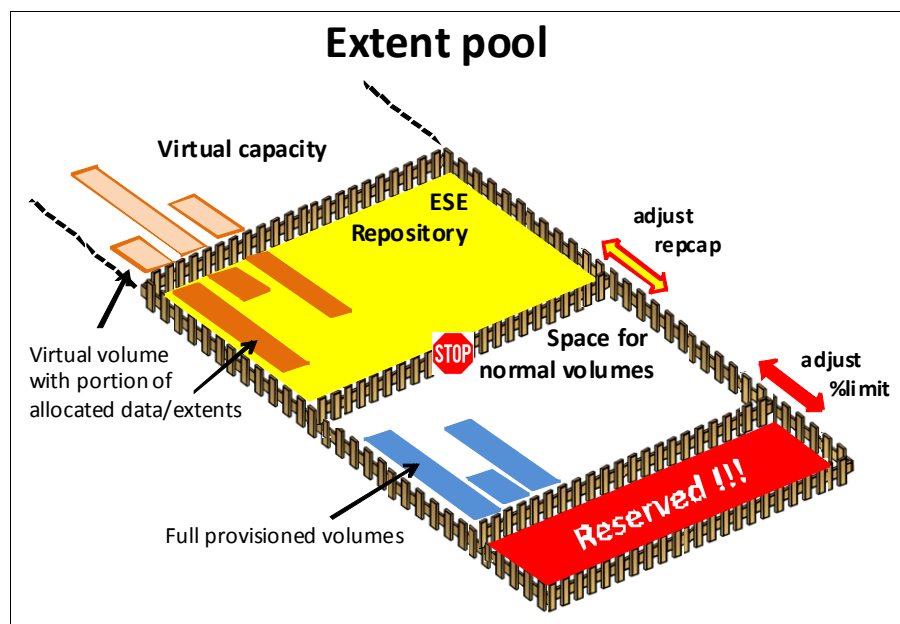
Figure 4-15   Spaces in an extent pool

Capacity controls exist for an extent pool and also for a repository, if it is defined. There are system defined warning thresholds at 15% and 0% free capacity left, and you can set your own user-defined threshold for the whole extent pool or for the ESE repository. Thresholds for an *extent pool* are set with the DSCLI `chextpool -threshold` (or `mkextpool`) command and for a *repository* with the `chsestg -repcapthreshold` (or `mksestg`) command.

The *threshold exceeded* status refers to the user-defined threshold.

The extent pool or repository status shows one of the three following values:

► 0: The percent of available capacity is greater than the extent pool / repository threshold.

► 1: The percent of available capacity is greater than zero but less than or equal to the extent pool / repository threshold.

► 10: The percent of available capacity is zero.

An SNMP trap is associated with the extent pool / repository capacity controls that notifies you when the extent usage in the pool exceeds a user-defined threshold, and when the extent pool is out of extents for user data.

When the size of the extent pool remains fixed or when it is increased, the available physical capacity remains greater than or equal to the allocated physical capacity. However, a reduction in the size of the extent pool can cause the *available* physical capacity to become less than the *allocated* physical capacity in certain cases.

For example, if the user requests that one of the ranks in an extent pool is depopulated, the data on that rank is moved to the remaining ranks in the pool. This process causes the rank to become not allocated and removed from the pool. The user is advised to inspect the limits and threshold on the extent pool after any changes to the size of the extent pool to ensure that the specified values are still consistent with the user's intentions.

4.4.5 Logical subsystems

An LSS is another logical construct. It can also be referred to as an LCU. In reality, it is licensed internal code (LIC) that is used to manage up to 256 logical volumes. The term *LSS* is used in association with FB volumes. The term *LCU* is used in association with CKD volumes. A maximum of 255 LSSs can exist in the DS8880. They each have an identifier from 00 - FE. An individual LSS must manage either FB or CKD volumes.

All even-numbered LSSs (X'00', X'02', X'04', up to X'FE') are handled by internal server 0, and all odd-numbered LSSs (X'01', X'03', X'05', up to X'FD') are handled by internal server 1. LSS X'FF' is reserved. This configuration allows both servers to handle host commands to the volumes in the DS8880 if the configuration takes advantage of this capability. If either server is not available, the remaining operational server handles all LSSs. LSSs are also placed in address groups of 16 LSSs, except for the last group that has 15 LSSs. The first address group is 00 - 0F, and so on, until the last group, which is F0 - FE.

Because the LSSs manage volumes, an individual LSS must manage the same type of volumes. An address group must also manage the same type of volumes. The first volume (either FB or CKD) that is assigned to an LSS in any address group sets that group to manage those types of volumes. For more information, see "Address groups" on page 124.

Volumes are created in extent pools that are associated with either internal server 0 or 1. Extent pools are also formatted to support either FB or CKD volumes. Therefore, volumes in any internal server 0 extent pools can be managed by any even-numbered LSS, if the LSS and extent pool match the volume type. Volumes in any internal server 1 extent pools can be managed by any odd-numbered LSS, if the LSS and extent pool match the volume type.

Volumes also have an identifier that ranges from 00 - FF. The first volume that is assigned to an LSS has an identifier of 00. The second volume is 01, and so on, up to FF, if 256 volumes are assigned to the LSS.

For FB volumes, the LSSs that are used to manage them are not significant if you spread the volumes between odd and even LSSs. When the volume is assigned to a host (in the DS8880 configuration), a LUN is assigned to it that includes the LSS and Volume ID. This LUN is sent to the host when it first communicates with the DS8880, so it can include the LUN in the "frame" that is sent to the DS8880 when it wants to run an I/O operation on the volume. This method is how the DS8880 knows which volume to run the operation on.

Conversely, for CKD volumes, the LCU *is* significant. The LCU must be defined in a configuration that is called the input/output configuration data set (IOCDS) on the host. The LCU definition includes a control unit address (CUADD). This CUADD must match the LCU ID in the DS8880. A device definition for each volume, which has a unit address (UA) that is included, is also included in the IOCDS. This UA needs to match the volume ID of the device. The host must include the CUADD and UA in the "frame" that is sent to the DS8880 when it wants to run an I/O operation on the volume. This method is how the DS8880 knows which volume to run the operation on.

For both FB and CKD volumes, when the "frame" that is sent from the host arrives at a host adapter port in the DS8880, the adapter checks the LSS or LCU identifier to know which internal server to pass the request to inside the DS8880. For more information about host access to volumes, see 4.4.6, "Volume access" on page 125.

Fixed-block LSSs are created automatically when the first fixed-block logical volume on the LSS is created. Fixed-block LSSs are deleted automatically when the last fixed-block logical volume on the LSS is deleted. CKD LCUs require user parameters to be specified and must be created before the first CKD logical volume can be created on the LCU. They must be deleted manually after the last CKD logical volume on the LCU is deleted.

Certain management actions in Metro Mirror, Global Mirror, or Global Copy operate at the LSS level. For example, the freezing of pairs to preserve data consistency across all pairs, in case a problem occurs with one of the pairs, is performed at the LSS level.

The option to put all or most of the volumes of a certain application in one LSS makes the management of remote copy operations easier, as shown in Figure 4-16.
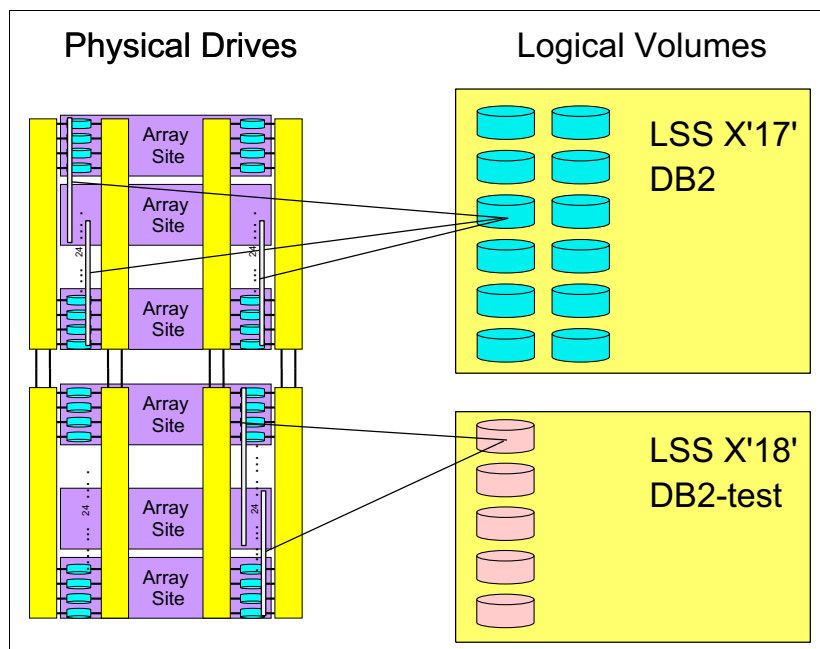
Figure 4-16   Grouping of volumes in LSSs

Address groups

Address groups are created automatically when the first LSS that is associated with the address group is created. The groups are deleted automatically when the last LSS in the address group is deleted.

All devices in an LSS must be CKD or FB. This restriction goes even further. LSSs are grouped into address groups of 16 LSSs. LSSs are numbered X'*ab*', where *a* is the address group and *b* denotes an LSS within the address group. For example, X'10' to X'1F' are LSSs in address group 1.

All LSSs within one address group must be of the same type (CKD or FB). The first LSS that is defined in an address group sets the type of that address group.

> **Important:** z Systems users who still want to use IBM Enterprise Systems Connection (ESCON) to attach hosts through ESCON/Fibre Channel connection (FICON) converters to the DS8880 need to be aware that ESCON supports only the 16 LSSs of address group 0 (LSS X'00' to X'0F'). Therefore, reserve this address group for CKD devices that attach to ESCON in this case, and do not use it for FB LSSs. The DS8880 does not support native ESCON channels.

Figure 4-17 shows the concept of LSSs and address groups.
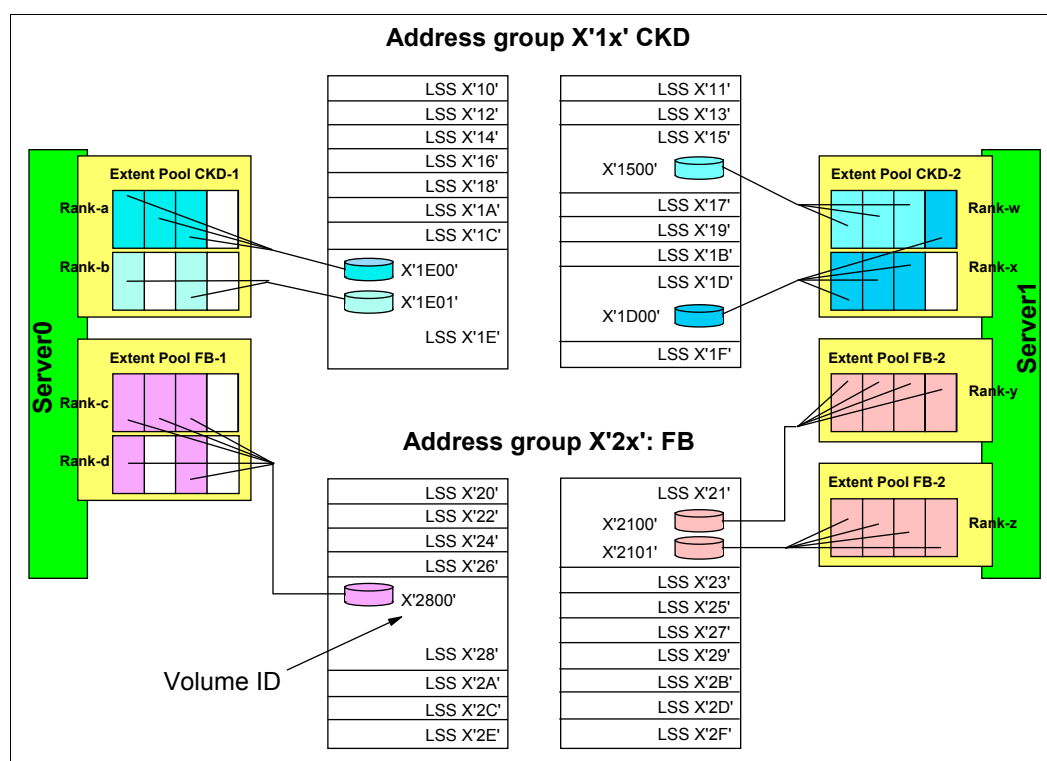
Figure 4-17   Logical storage subsystems

The LUN identifications X'*gabb*' are composed of the address group X'*g*', and the LSS number within the address group X'*a*', and the ID of the LUN within the LSS X'*bb*'. For example, FB LUN X'2101' denotes the second (X'01') LUN in LSS X'21' of address group 2.

An extent pool can have volumes that are managed by multiple address groups. The example in Figure 4-17 shows only one address group that is used with each extent pool.

4.4.6 Volume access

The DS8880 provides mechanisms to control host access to LUNs. In most cases, a host system features two or more HBAs and the system needs access to a group of LUNs. For easy management of system access to logical volumes, the DS8880 introduced the concept of host attachments and volume groups.

Host attachment

HBAs are identified to the DS8880 in a host attachment construct that specifies the worldwide port names (WWPNs) of a host's HBAs. A set of host ports can be associated through a port group attribute that allows a set of HBAs to be managed collectively. This port group is referred to as a *host attachment* within the configuration.

Each host attachment can be associated with a volume group to define the LUNs that host is allowed to access. Multiple host attachments can share the volume group. The host attachment can also specify a port mask that controls the DS8880 I/O ports that the host HBA is allowed to log in to. Whichever ports the HBA logs in on, it sees the same volume group that is defined on the host attachment that is associated with this HBA.

The maximum number of host attachments on a DS8880 is 8,192. This host definition is only required for open systems hosts. Any z Systems servers can access any CKD volume in a DS8880 if its IOCDS is correct.

Volume group

A *volume group* is a named construct that defines a set of logical volumes. A volume group is only required for FB volumes. When a volume group is used with CKD hosts, a default volume group contains all CKD volumes. Any CKD host that logs in to a FICON I/O port has access to the volumes in this volume group. CKD logical volumes are automatically added to this volume group when they are created and are automatically removed from this volume group when they are deleted.

When a host attachment object is used with open systems hosts, a host attachment object that identifies the HBA is linked to a specific volume group. You must define the volume group by indicating the FB volumes that are to be placed in the volume group. Logical volumes can be added to or removed from any volume group dynamically.

> **Important:** Volume group management is available only with the DS CLI. In the DS GUI, users define hosts and assign volumes to hosts. A volume group is defined in the background. No volume group object can be defined in the DS GUI.

Two types of volume groups are used with open systems hosts. The type determines how the logical volume number is converted to a host addressable LUN_ID in the Fibre Channel SCSI interface. A *SCSI map volume group* type is used with FC SCSI host types that poll for LUNs by walking the address range on the SCSI interface. This type of volume group can map any FB logical volume numbers to 256 LUN IDs that have zeros in the last 6 bytes and the first 2 bytes in the range of X'0000' to X'00FF'.

A *SCSI mask volume group* type is used with FC SCSI host types that use the `Report LUNs` command to determine the LUN IDs that are accessible. This type of volume group can allow any FB logical volume numbers to be accessed by the host where the mask is a bitmap that specifies the LUNs that are accessible. For this volume group type, the logical volume number X'*abcd*' is mapped to LUN_ID X'40ab40cd00000000'. The volume group type also controls whether 512-byte block LUNs or 520-byte block LUNs can be configured in the volume group.

When a host attachment is associated with a volume group, the host attachment contains attributes that define the logical block size and the Address Discovery Method (LUN Polling or Report LUNs) that is used by the host HBA. These attributes must be consistent with the volume group type of the volume group that is assigned to the host attachment. This consistency ensures that HBAs that share a volume group have a consistent interpretation of the volume group definition and have access to a consistent set of logical volume types. The DS Storage Manager GUI typically sets these values for the HBA based on your specification of a host type. You must consider what volume group type to create when a volume group is set up for a particular HBA.

FB logical volumes can be defined in one or more volume groups. This definition allows a LUN to be shared by host HBAs that are configured to separate volume groups. An FB logical volume is automatically removed from all volume groups when it is deleted.

The DS8880 supports a maximum of 8,320 volume groups.

Figure 4-18 shows the relationships between host attachments and volume groups. Host AIXprod1 has two HBAs, which are grouped in one host attachment and both HBAs are granted access to volume group DB2-1. Most of the volumes in volume group DB2-1 are also in volume group DB2-2, which is accessed by the system AIXprod2.

However, one volume in each group is not shared in Figure 4-18. The system in the lower-left part of the figure features four HBAs, and they are divided into two distinct host attachments. One HBA can access volumes that are shared with AIXprod1 and AIXprod2. The other HBAs have access to a volume group that is called `docs`.
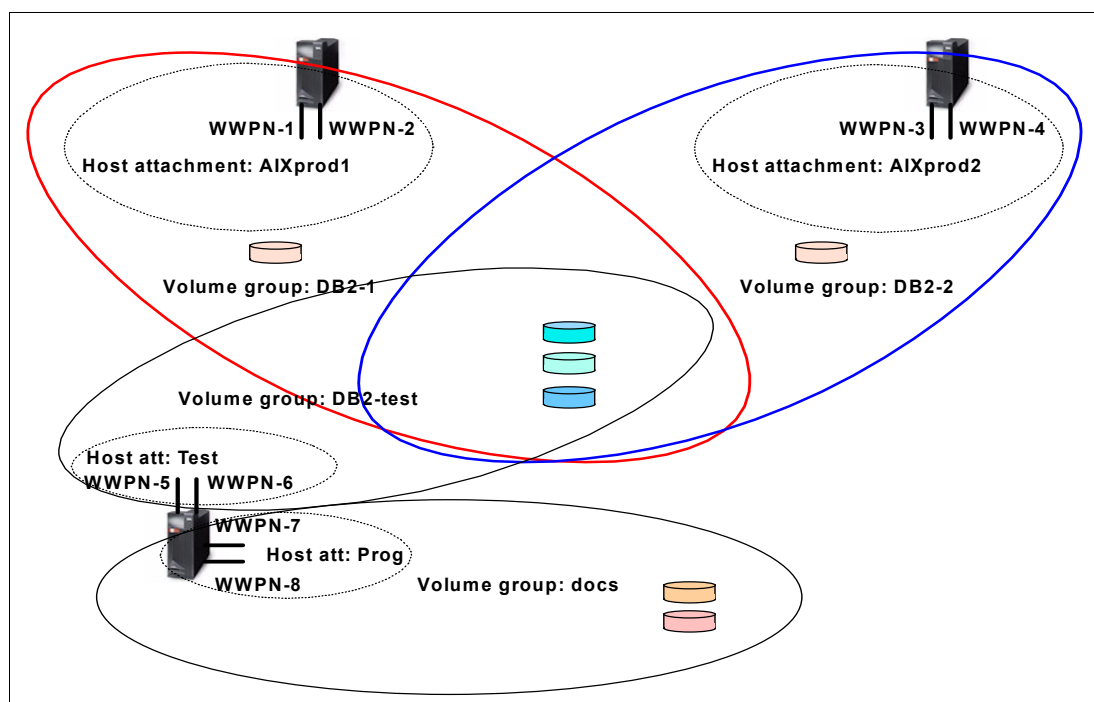
Figure 4-18   Host attachments and volume groups

When working with Open Systems clusters, for defining volumes, use the Create Cluster function of the DS Storage Manager GUI to easily define volumes. In general, the GUI hides the complexity of certain DS8000-internal definition levels like volume groups, array sites, and ranks. It helps save time by directly processing these definitions internally in the background, without presenting them to the administrator.

4.4.7 Virtualization hierarchy summary

As you view the virtualization hierarchy (as shown in Figure 4-19), start with many drives that are grouped in array sites. The array sites are created automatically when the drives are installed. Complete the following steps as a user:

1. Associate an array site with a RAID array, and allocate spare drives, as required.

2. Associate the array with a rank with small or large extents formatted for FB or CKD data.

3. Add the extents from the selected ranks to an extent pool. The combined extents from the ranks in the extent pool are used for subsequent allocation for one or more logical volumes. Within the extent pool, you can reserve space for ESE volumes. ESE volumes require virtual capacity to be available in the extent pool.

4. Create logical volumes within the extent pools (by default, striping the volumes), and assign them a logical volume number that determines the logical subsystem that they are associated with and the internal server that manages them. The LUNs are assigned to one or more volume groups.

5. Configure the host HBAs into a host attachment associated with a volume group.

Figure 4-19 shows the virtualization hierarchy.
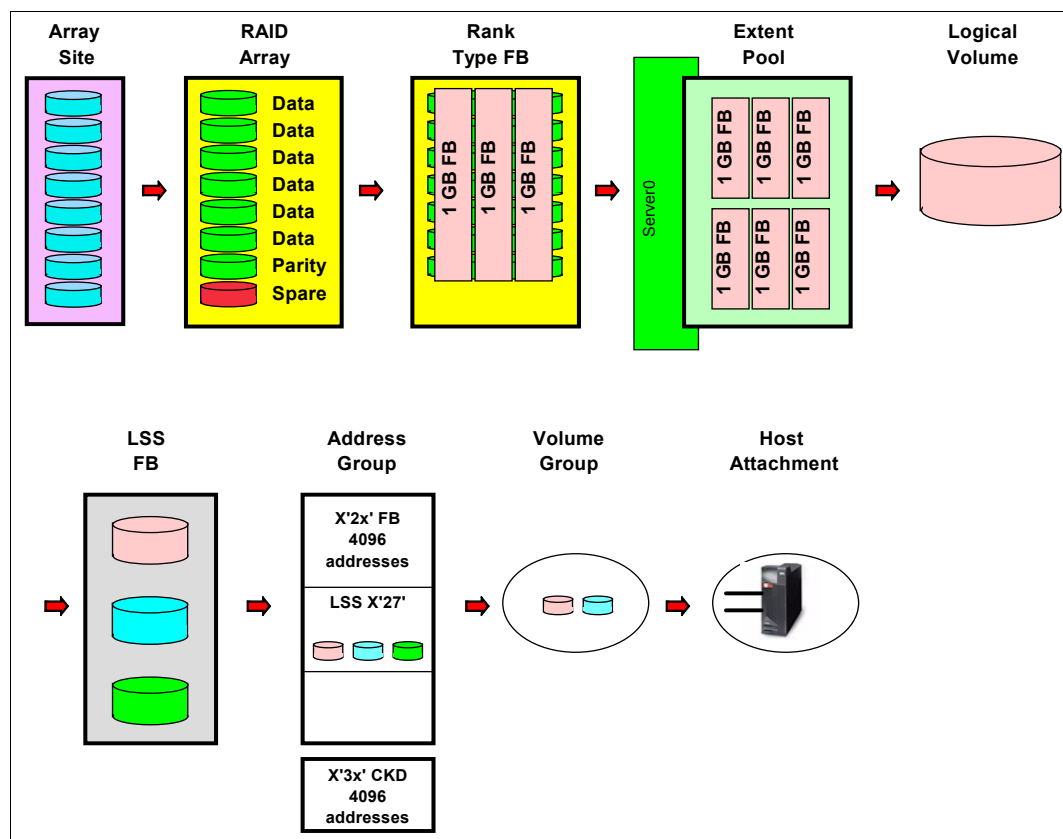
Figure 4-19   Virtualization hierarchy

This virtualization concept provides a high degree of flexibility. Logical volumes can be dynamically created, deleted, and resized. They can also be grouped logically to simplify storage management. Large LUNs and CKD volumes reduce the total number of volumes, which contributes to the reduction of management effort.

Tip: The DS GUI helps save administration steps by handling some of these virtualization levels in the background, and automatically processing them for the administrator.

5

DS8000 Copy Services overview

This chapter provides an overview of the Copy Services functions that are available with the DS8000 series models, including Remote Mirror and Copy, and Point-in-Time Copy functions.

These functions make the DS8000 series a key component for disaster recovery solutions, data migration activities, and for data duplication and backup solutions.

This chapter covers the following topics:

► Introduction to Copy Services
► FlashCopy
► Remote Mirror and Copy

5.1 Introduction to Copy Services

Copy Services is a collection of functions that provide disaster recovery, data migration, and data duplication functions. With the Copy Services functions, for example, you can create backup data with little or no disruption to your application. You also can back up your application data to the remote site for disaster recovery.

The Copy Services functions run on the DS8880 storage unit and support open systems and z Systems environments. They are also supported on other DS8000 family models.

5.1.1 DS8000 Copy Services functions

Copy Services in the DS8000 include the following optional licensed functions:

► IBM System Storage FlashCopy, a point-in-time copy function

  The FlashCopy Space Efficient function is not supported on the DS8880. Instead, you can define your FlashCopy targets as Extent Space Efficient volumes in an extent pool by using small extents.

► Remote mirror and copy functions:

  – IBM System Storage Metro Mirror.

  – IBM System Storage Global Copy.

  – IBM System Storage Global Mirror.

  – IBM System Storage Metro/Global Mirror, which is a three-site solution to meet the most rigorous business resiliency needs.

  – IBM System Storage Multiple Target Peer-to-Peer Remote Copy (PPRC), which is another three-site solution that concurrently replicates a primary volume to two secondary volumes. It can be implemented as either two Metro Mirror relationships off the same primary volume or one Metro Mirror relationship and an asynchronous replication from the same primary volume concurrently.

► For IBM z Systems users, the following options are available:

  – z/OS Global Mirror, which was known as Extended Remote Copy (XRC).

  – z/OS Metro/Global Mirror, which is a three-site solution that combines z/OS Global Mirror and Metro Mirror.

Many design characteristics of the DS8000, its data copy and mirror capabilities, and features contribute to the full-time protection of your data.

Other design characteristics have improved. The FlashCopy SE (former Space-Efficient based on Track Space Efficient) function is not supported on the DS8880. Instead, you can define your FlashCopy targets as Extent Space-Efficient volumes in an extent pool using small extents.

The information that is provided in this chapter is only an overview. Copy Services are covered in more detail in the following materials:

► *DS8000 Copy Services*, SG24-8367
► *IBM DS8870 Multiple Target Peer-to-Peer Remote Copy*, REDP-5151

5.1.2 Copy Services management interfaces

You control and manage the DS8000 Copy Services functions by using the following interfaces:

► Data storage command-line interface (DS CLI), which provides a set command that covers all Copy Services functions and options.

► IBM Copy Services Manager (CSM), formerly known as *Tivoli Storage Productivity Center for Replication*, to manage large Copy Services implementations easily and to provide data consistency across multiple systems. With DS8000 Release 8.1 and later, CSM is preinstalled on the Hardware Management Console (HMC).

► DS Storage Manager, the graphical user interface (GUI) of the DS8000 (DS GUI).

► DS Open application programming interface (DS Open API).

z Systems users can also use the following interfaces:

► Time Sharing Option (TSO) commands
► Device Support Facilities (ICKDSF) utility commands
► ANTRQST application programming interface (API)
► Data Facility Storage Management Subsystem data set services (DFSMSdss) utility
► DSCLI on z/OS (see 12.1.4, "Installing the DS CLI on z/OS System" on page 351)

5.2 FlashCopy

FlashCopy creates a point-in-time copy of the data. For open systems, FlashCopy creates a copy of the logical unit number (LUN). The target LUN must exist before you can use FlashCopy to copy the data from the source LUN to the target LUN.

> **Note:** FlashCopy Space Efficient (Track Space Efficient) is not supported by the DS8880. Instead, you can define your FlashCopy targets as Extent Space Efficient (ESE) volumes in an extent pool using small extents.

When a FlashCopy operation is started, it takes less than a second to establish the FlashCopy relationship, which consists of the source and target volume pairing and the necessary control bitmaps. Thereafter, a copy of the source volume is available as though all the data was copied. When the pair is established, you can read and write to the source and target volumes.

The basic concepts of FlashCopy are explained in the following section and are shown in Figure 5-1.
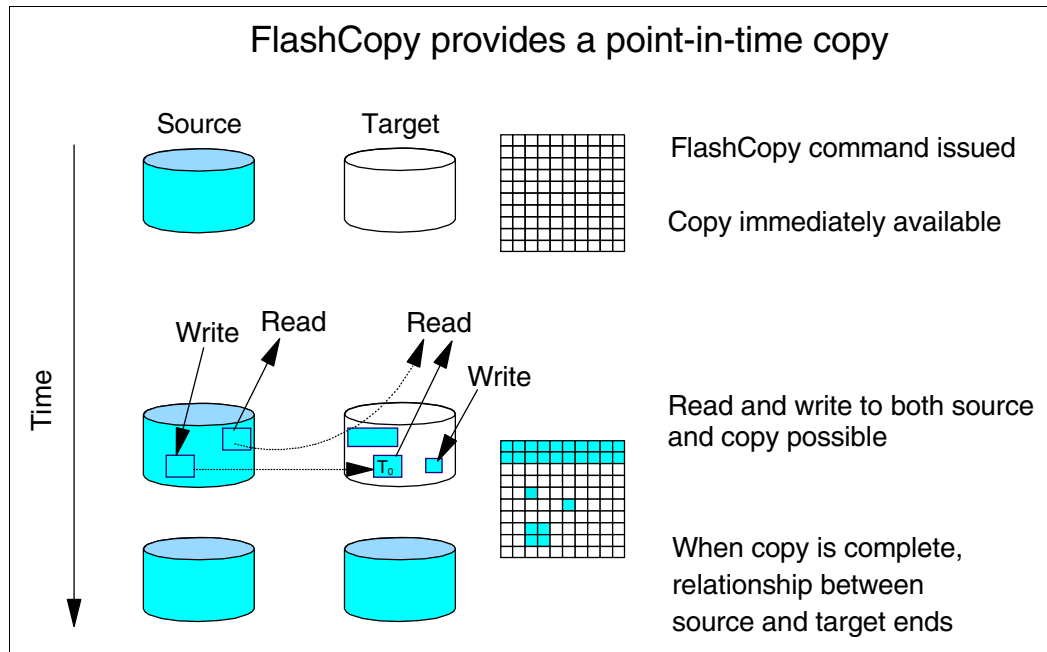
Figure 5-1   FlashCopy concepts

If you access the source or the target volumes while the FlashCopy relationship exists, I/O requests are handled in the following manner:

► Read from the source volume

When a read request goes to the source, data is directly read from there.

► Read from the target volume

When a read request goes to the target volume, FlashCopy checks the bitmap and takes one of the following actions:

– If the requested data has been copied to the target, it is read from there.

– If the requested data has not yet been copied, it is read from the source.

► Write to the source volume

When a write request goes to the source, the data is first written to the cache and persistent memory (write cache). Later, when the data is destaged to the physical extents of the source volume, FlashCopy checks the bitmap for the location that is to be overwritten and takes one of the following actions:

– If the point-in-time data was already copied to the target, the update is written to the source directly.

– If the point-in-time data was not yet copied to the target, it is now copied immediately and only then is the update written to the source.

► Write to the target volume

Whenever data is written to the target volume while the FlashCopy relationship exists, the storage system checks the bitmap and updates it, if necessary. FlashCopy does not overwrite data that was written to the target with point-in-time data.

5.2.1 The FlashCopy background copy

By default, standard FlashCopy (also called *FULL copy*) starts a background copy process that copies all point-in-time data to the target volume. After the completion of this process, the FlashCopy relationship ends and the target volume becomes independent of the source.

The background copy can slightly affect application performance because the physical copy needs storage resources. The impact is minimal because host I/O always has higher priority than the background copy.

5.2.2 No background copy option

A standard FlashCopy relationship can also be established by using the NOCOPY option. With this option, FlashCopy does not start a background copy. Point-in-time data is copied only when a copy is required due to an update to the source or target. This option eliminates the impact of the background copy.

The NOCOPY option is useful in the following situations:

► When the target is not needed as an independent volume
► When repeated FlashCopy operations to the same target are expected

5.2.3 Benefits and use

The point-in-time copy that is created by FlashCopy often is used when you need to produce a copy of the production data with little or no application downtime. Use cases for the point-in-time copy that is created by FlashCopy include online backup, testing new applications, or creating a copy of transactional data for data mining purposes. To the host or application, the target looks exactly like the original source. It is an instantly available binary copy.

5.2.4 FlashCopy options

FlashCopy provides many more options and functions. These options and capabilities are described in this section:

► Persistent FlashCopy
► Incremental FlashCopy (refresh target volume)
► Multiple relationship FlashCopy
► Data Set FlashCopy
► Consistency Group FlashCopy
► FlashCopy on existing Metro Mirror or Global Copy primary
► Inband commands over remote mirror link

Persistent FlashCopy

Persistent FlashCopy allows the FlashCopy relationship to remain even after the (FULL) copy operation completes. You must explicitly delete the relationship to end it.

Incremental FlashCopy (refresh target volume)

Incremental FlashCopy requires that the background copy and the Persistent FlashCopy options are enabled. The first full copy must be complete.

Refresh target volume refreshes a FlashCopy relationship without copying all data from the source to the target again. When a subsequent FlashCopy operation is initiated, only the changed tracks must be copied from the source to the target. The direction of the refresh also can be reversed from the (former) target to the source.

In many cases, only a small percentage of the entire data is changed in a day. In this situation, you can use this function for daily backups and save the time that is required for the physical copy of FlashCopy.

Multiple relationship FlashCopy

FlashCopy allows a source to have relationships with up to 12 targets simultaneously. A usage case for this feature is the creation of regular point-in-time copies as online backups or time stamps.

With DS8000 Licensed Machine Code (LMC) 7.7.40.*xx*, FlashCopy was enhanced to allow all FlashCopy source relationships to be incremental by introducing t*ype 2 incremental* FlashCopy relationships.

Two potential uses for multiple incremental FlashCopy relationships are shown:

► Database backup

   When more than one backup is made with FlashCopy each day, they can all be incremental copies.

► Global Mirror test copy

   The Global Mirror journal FlashCopy is an incremental copy. Multiple incremental copies now allow the test copy to also be incremental.

Data Set FlashCopy

By using Data Set FlashCopy, you can create a point-in-time copy of individual data sets instead of complete volumes in an IBM z Systems environment.

That process can be done by using an z/OS utility like DFSMSdss, for example.

Consistency Group FlashCopy

By using Consistency Group FlashCopy, you can freeze and temporarily queue I/O activity to a volume. Consistency Group FlashCopy helps you to create a consistent point-in-time copy without quiescing the application across multiple volumes, and even across multiple storage units.

Consistency Group FlashCopy ensures that the order of dependent writes is always maintained. Therefore, Consistency Group FlashCopy creates host-consistent copies, not application-consistent copies. The copies have power-fail or crash-level consistency. To recover an application from Consistency Group FlashCopy target volumes, you must perform the same recovery as after a system crash or power outage.

FlashCopy on existing Metro Mirror or Global Copy primary

By using this option, you establish a FlashCopy relationship where the target is a Metro Mirror or Global Copy primary volume. Through this relationship, you create full or incremental point-in-time copies at a local site and then use remote mirroring to copy the data to the remote site.

> **Important:** You cannot perform a FlashCopy from a source to a target if the target also is a Global Mirror primary volume.

For more information about Metro Mirror and Global Copy, see 5.3.2, "Metro Mirror" on page 136, and 5.3.3, "Global Copy" on page 137.

Inband commands over remote mirror link

In a remote mirror environment, commands to manage FlashCopy at the remote site can be issued from the local or intermediate site and transmitted over the remote mirror Fibre Channel links. This ability eliminates the need for a network connection to the remote site solely for the management of FlashCopy.

5.2.5 Remote-Pair FlashCopy (Preserve Mirror)

Remote Pair FlashCopy (also referred to as Preserve Mirror) transmits the `FlashCopy` command to the remote site if the target volume is mirrored with Metro Mirror. As the name implies, Preserve Mirror preserves the existing Metro Mirror status of FULL DUPLEX.

For more information about Remote Pair FlashCopy, see *IBM System Storage DS8000: Remote Pair FlashCopy (Preserve Mirror)*, REDP-4504.

5.2.6 Remote-Pair FlashCopy with Multiple-Target PPRC

Multiple-Target PPRC allows a single Metro Mirror, Global Mirror, or Global Copy primary to have two secondary volumes. Additional considerations are involved to use Remote-Pair FlashCopy in a Multiple-Target PPRC environment. For more information, see *IBM DS8870 Multiple Target Peer-to-Peer Remote Copy*, REDP-5151.

5.2.7 FlashCopy with Thin Provisioning

FlashCopy can be used with thin-provisioned volumes. All volume combinations are possible: Thick-to-thick, thick-to-thin, thin-to-thin and so on. Therefore, it is possible for an instance to have the source volume as a standard thick volume, for example in an extent pool with large extents, and the FlashCopy target volume being thinly provisioned, and in an extent pool with small extents.

5.3 Remote Mirror and Copy

The Remote Mirror and Copy functions of the DS8000 are a set of flexible data mirroring solutions that allow replication between volumes on two or more disk storage systems. These functions are used to implement remote data backup and disaster recovery solutions.

The following Remote Mirror and Copy functions are optional licensed functions of the DS8000:

► Metro Mirror
► Global Copy
► Global Mirror
► Metro/Global Mirror

Remote Mirror functions can be used in open systems and z Systems environments.

In addition, z Systems users can use the DS8000 for the following functions:

► z/OS Global Mirror
► z/OS Metro/Global Mirror
► IBM Geographically Dispersed Parallel Sysplex (IBM GDPS)

5.3.1 Thin Provisioning with Remote Replication

Thin provisioning can also be used with Remote Mirror and Copy. On establishing a PPRC relation, space is released on the target volume. The DS8000 can replicate between standard thick volumes (thick-to-thick), or from an ESE thinly provisioned volume to another thin volume.

Starting with DS8000 code levels R8.2 and later and when both primary and secondary DS8000 have at least this level, replicating thick-to-thin is also possible.

When using the option to have both primary and secondary volumes as ESE thin volumes, some bandwidth for the replication can be saved because only the physically allocated extents are copied.

5.3.2 Metro Mirror

Metro Mirror provides real-time mirroring of logical volumes between two DS8880s, or any other combination of DS8870, DS8100, DS8300, DS8700, DS8800, DS6800, and ESS800. These volumes can be located up to 300 km (186.4 miles) from each other. It is a synchronous copy solution in which a write operation must be carried out on both copies, at the local and remote sites, before it is considered complete.

The basic operational characteristics of Metro Mirror are shown in Figure 5-2.
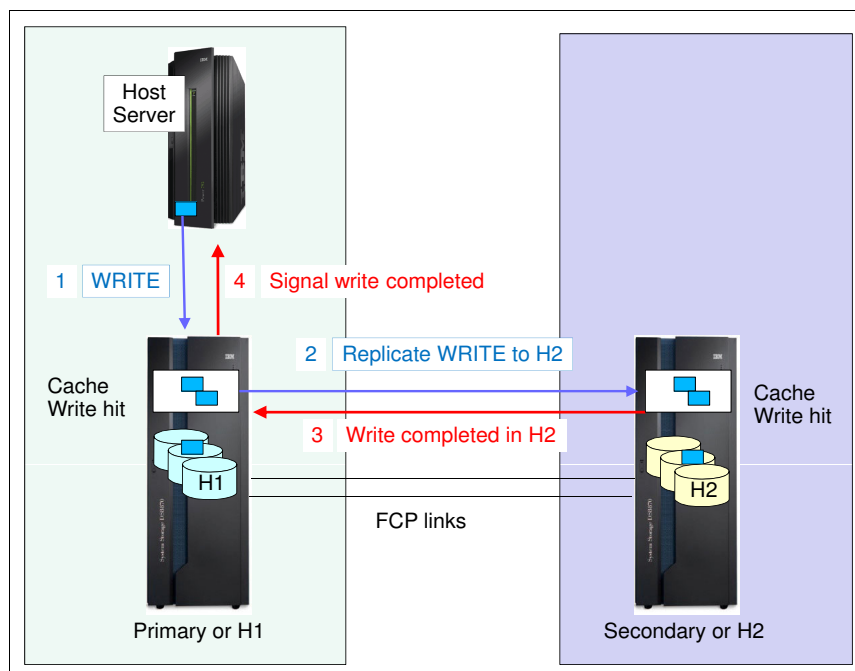
Figure 5-2   Metro Mirror basic operation

5.3.3 Global Copy

Global Copy copies data asynchronously and over longer distances than is possible with Metro Mirror. Global Copy is included in the Metro Mirror or Global Mirror license. When you are operating in Global Copy mode, the source does not wait for copy completion on the target before a host write operation is acknowledged. Therefore, the host is not affected by the Global Copy operation. Write data is sent to the target because the connecting network allows an independent of the order of the host writes. This configuration makes the target data lag behind and is inconsistent during normal operation.

You must take extra steps to make Global Copy target data usable at specific point in time. The steps that are used depend on the purpose of the copy:

► Data migration

You can use Global Copy to migrate data over long distances. When you want to switch from old to new data, you must stop the applications on the old site, tell Global Copy to synchronize the data, and wait until it is finished.

► Asynchronous mirroring

Global Copy is also used to create a full-copy of data from an existing system to a new system without affecting client performance. If the Global Copy is incomplete, the data at the remote system is not consistent. When the Global Copy completes, you can stop it and then start with the Copy relationship (Metro Mirror or Global Mirror) starting with a full resynchronization so the data is consistent.

5.3.4 Global Mirror

Global Mirror is a two-site, long distance, asynchronous, remote copy technology. This solution integrates the Global Copy and FlashCopy technologies. With Global Mirror, the data that the host writes at the local site is asynchronously mirrored to the storage unit at the remote site. With special management steps (under control of the local master storage unit), a consistent copy of the data is automatically maintained and periodically updated by using FlashCopy on the storage unit at the remote site.

You need extra storage at the remote site for these FlashCopies. However, starting with DS8880 Release 8.1, you can define your FlashCopy targets as Extent Space Efficient ESE volumes in an extent pool using small extents.

The Global Mirror Journal volumes are a key use case for ESE to reduce the additional space that is required for this configuration.

Global Mirror performs automatic space releases on such space-efficient Journal targets. This process will not occur on every consistency group formation, but often and in a regular manner.

Global Mirror features the following benefits:

► Support for almost unlimited distances between the local and remote sites, with the distance typically limited only by the capabilities of the network and the channel extension technology. You can use this unlimited distance to choose your remote site location based on business needs. This capability enables site separation to add protection from globalized disasters.

► A consistent and restartable copy of the data at the remote site, which is created with minimal impact to applications at the local site.

► Data currency where, for many environments, the remote site lags behind the local site typically 3 - 5 seconds, which minimizes the amount of data exposure if an unplanned outage occurs. The actual lag in data currency that you experience depends on many factors, including specific workload characteristics and bandwidth between the local and remote sites.

► Dynamic selection of the recovery point objective (RPO) that you want, based on business requirements and optimization of available bandwidth.

► Session support in which data consistency at the remote site is internally managed across up to eight storage units that are at the local site and the remote site.

► Efficient synchronization of the local and remote sites with support for failover and failback operations. This configuration helps to reduce the time that is required to switch back to the local site after a planned or unplanned outage.

The basic operational characteristics of Global Mirror are shown in Figure 5-3.
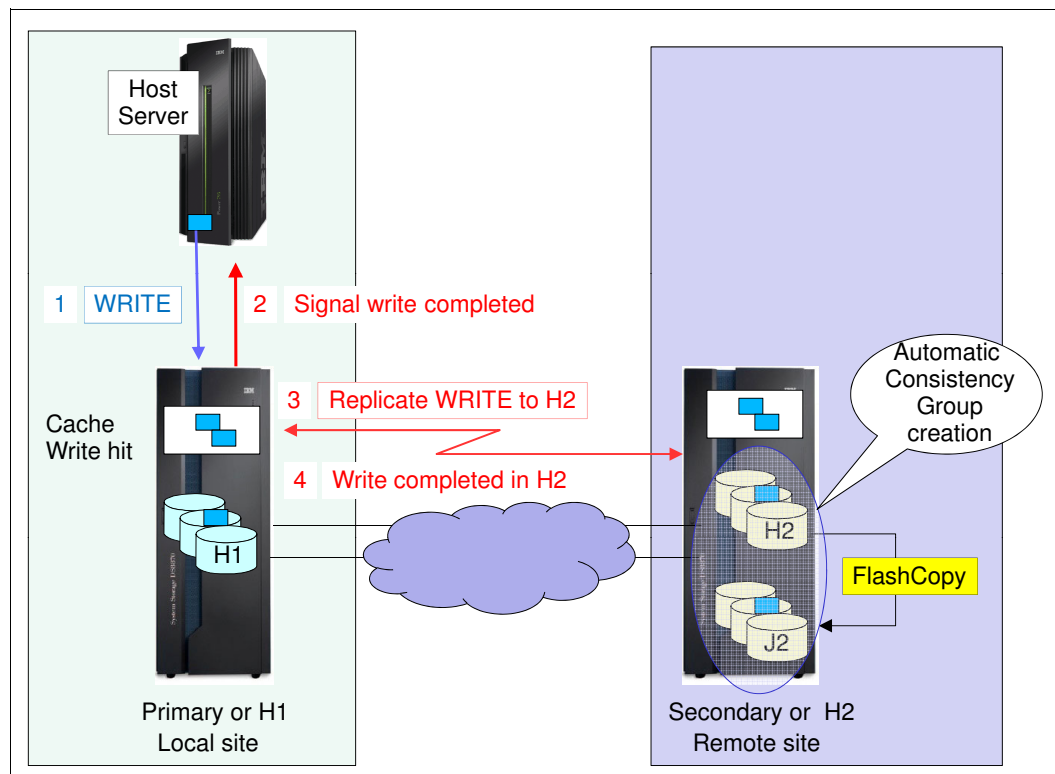
Figure 5-3   Global Mirror basic operation

The H1 volumes at the local site are the production volumes. They are used as Global Copy primaries. The data from the H1 volumes is replicated to the H2 volumes by using Global Copy. At a certain point, a consistency group is created from all of the H1 volumes, even if

they are in separate storage units. This creation has little impact on applications because the creation of the consistency group is quick (often a few milliseconds).

After the consistency group is created, the application writes can continue updating the H1 volumes. The missing increment of the consistent data is sent to the H2 volumes by using the existing Global Copy relationships. After all data reaches the H2 volumes, Global Copy is halted for a brief period while Global Mirror creates a FlashCopy from the H2 to the J2 volumes. These volumes now contain a consistent set of data at the secondary site.

The data at the remote site typically is current within 3 - 5 seconds, but this recovery point depends on the workload and bandwidth that is available to the remote site.

With its efficient and autonomic implementation, Global Mirror is a solution for disaster recovery implementations where a consistent copy of the data must always be available at a remote location that is separated by a long distance from the production site.

5.3.5 Metro/Global Mirror

Metro/Global Mirror is a three-site, multi-purpose, replication solution. Metro Mirror from local site H1 to intermediate site H2 provides high availability replication. Global Mirror from intermediate site H2 to remote site H3 provides long-distance disaster recovery replication with Global Mirror (Figure 5-4). This cascaded approach for a three-site solution does not burden the primary storage system with sending out the data twice.
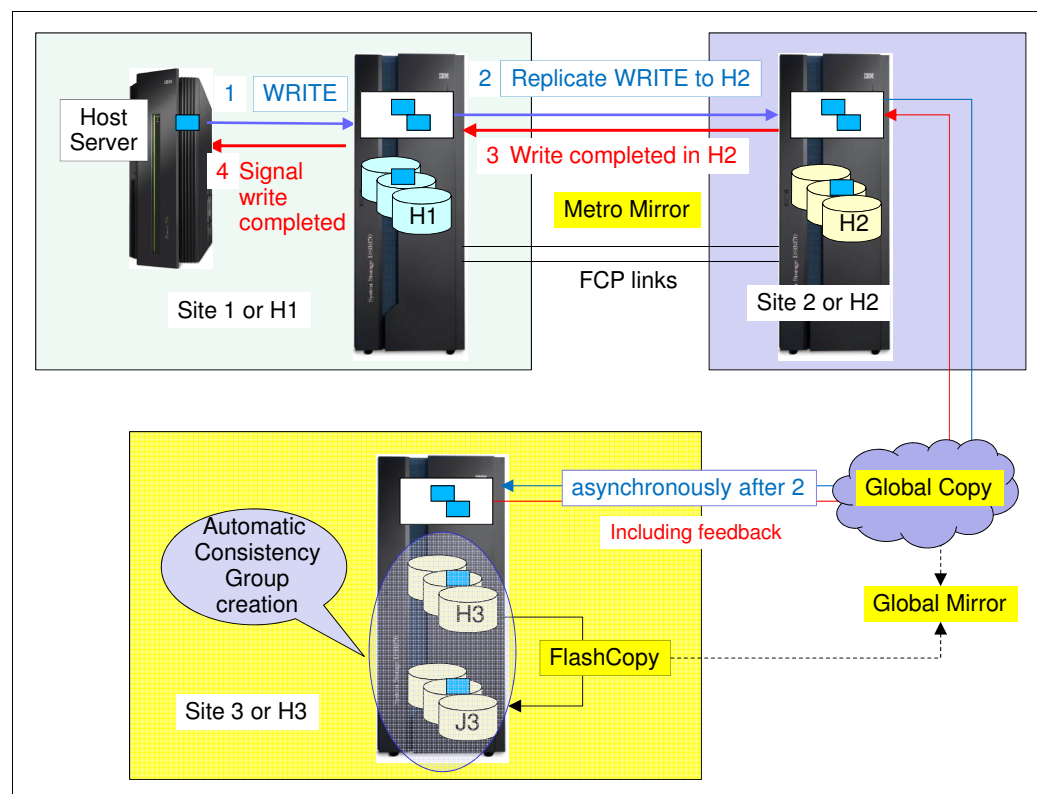
Figure 5-4   Cascaded Metro/Global Mirror elements

Metro Mirror and Global Mirror are well-established replication solutions. Metro/Global Mirror combines Metro Mirror and Global Mirror to incorporate the following best features of the two solutions:

- ► Metro Mirror:
  - – Synchronous operation supports zero data loss.
  - – The opportunity to locate the intermediate site disk systems close to the local site allows the use of intermediate site disk systems in a high-availability configuration.
- ► Global Mirror:
  - – Asynchronous operation supports long-distance replication for disaster recovery.
  - – The Global Mirror methodology has no effect on applications at the local site.
  - – This solution provides a recoverable, restartable, and consistent image at the remote site with an RPO, typically within 3 - 5 seconds.

Multiple Global Mirror sessions

The DS8870 and DS8880 supports several Global Mirror (GM) sessions within a storage system (storage facility image (SFI)). Up to 32 Global Mirror hardware sessions can be supported within the same DS8870 or DS8880. The basic management of a GM session does not change. The GM session builds on the existing Global Mirror technology and LMC of the DS8000.

For details, see the "Global Mirror Overview" chapter in *DS8000 Copy Services*, SG24-8367.

GM and MGM collision avoidance

Global Copy and GM are asynchronous functions that are suited for long distances between a primary and a secondary DS8000 storage system. At a long distance, it is important to allow hosts to complete an I/O operation, even if the transaction on the remote site is incomplete.

During high activity (for example, long-running batch jobs), multiple writes might update the same track or block, which can result in a collision. To avoid such collisions, Global Mirror locks tracks in the consistency group (CG) on the primary DS8000 at the end of the CG formation window.

5.3.6 Multi-Target PPRC

IBM Multi-Target PPRC enhances a multi-site disaster recovery environment by providing the capability to have two PPRC relationships on a single primary volume, allowing use of another remote site for extra data protection.

Multi-Target PPRC provides the following enhancements:

- ► Mirrors data from a single primary (local) site to two secondary (remote) sites
- ► Provides an increased capability and flexibility in the following disaster recovery solutions:
  - – Synchronous replication
  - – Asynchronous replication
  - – Combination of synchronous replication and asynchronous replication configurations
- ► Improves a cascaded Metro/Global Mirror configuration and simplifies certain procedures

Before Multi-Target PPRC, it was possible for a primary volume to mirror data to only one secondary volume. With Multi-Target PPRC, the same primary volume can have more than one target, allowing data to be mirrored from a single primary site to two target sites.

Figure 5-5 shows a general Multi-Target PPRC topology where a single primary site is replicated to two secondary sites. Host I/O is directed to the H1 site and Multi-Target PPRC mirrors the updates to both H2 and H3.
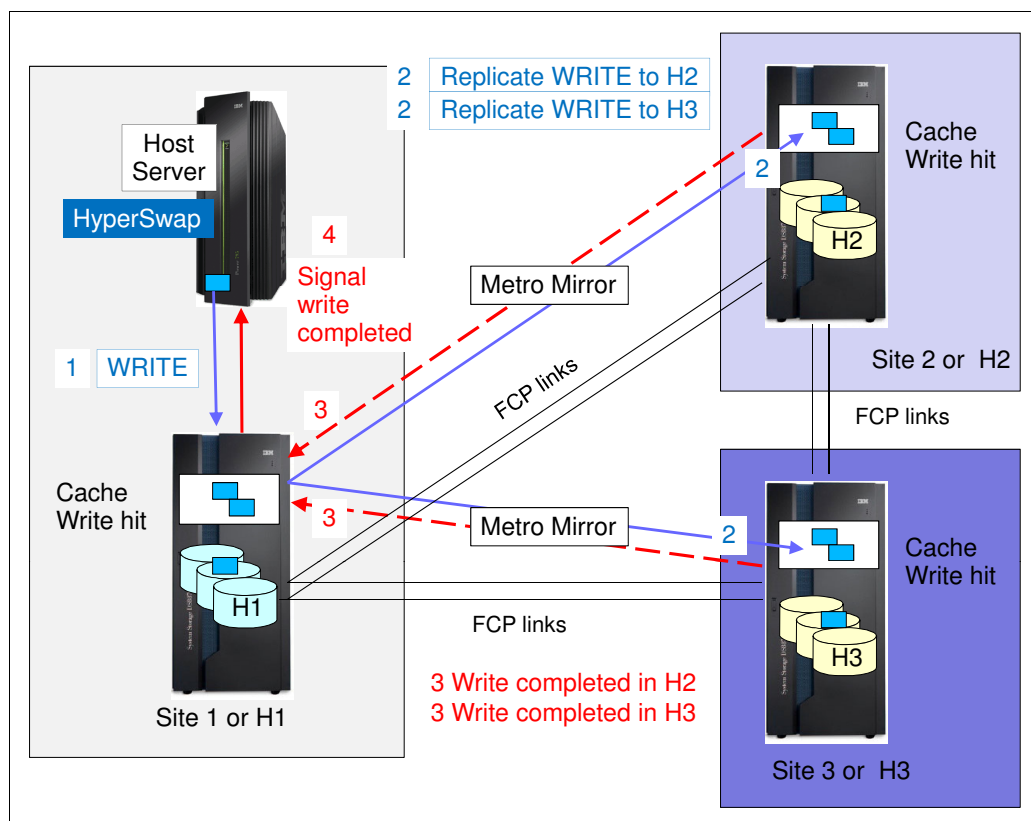
Figure 5-5   Basic Multi-Target PPRC configuration

A primary volume can have any combination of two Metro Mirror, Global Copy, or Global Mirror relationships, with a restriction that a primary volume can belong to only one Global Mirror session at a time.

> **Note:** A volume can be in only one Global Mirror session.

Little extra host response time impact occurs with two relationships compared to one relationship of the same type. Multi-Target PPRC is available in both open systems and z Systems environments.

For more information, see *IBM DS8870 Multiple Target Peer-to-Peer Remote Copy*, REDP-5151.

5.3.7 z/OS Global Mirror

z/OS Global Mirror, which was known as Extended Remote Copy (XRC), is a copy function that is available for the z/OS operating system. The basic operational characteristics of z/OS Global Mirror are shown in Figure 5-6.
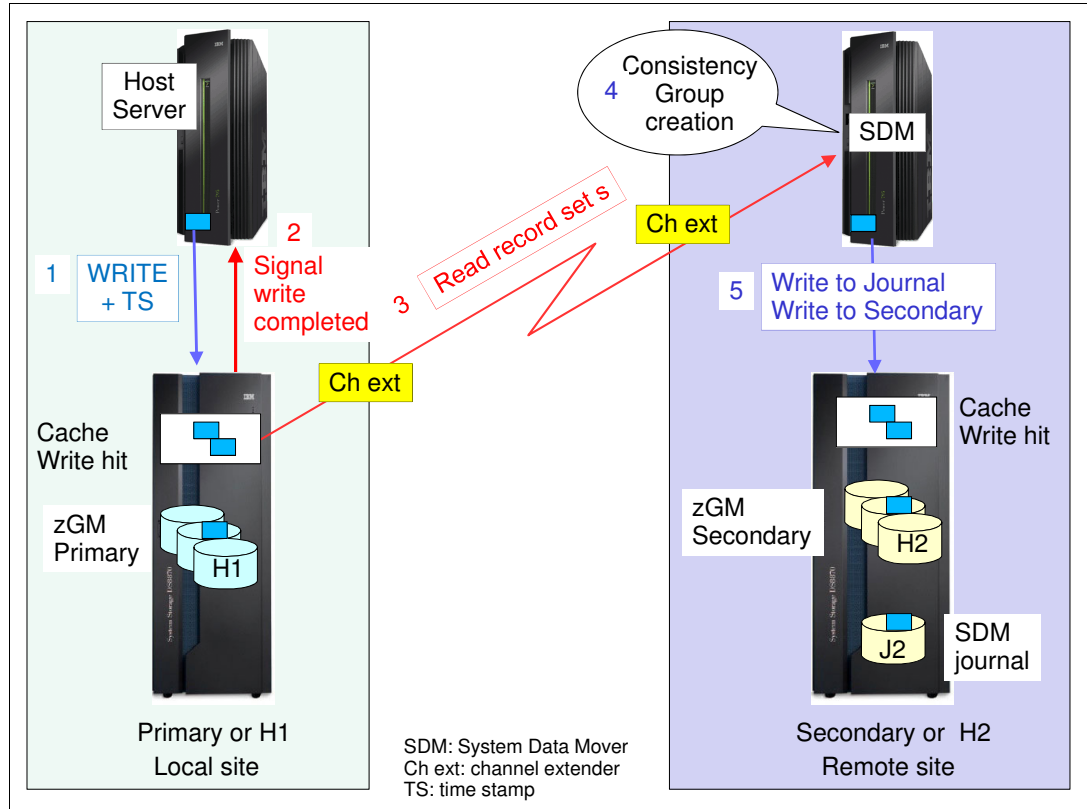
Figure 5-6 z/OS Global Mirror basic operations

It involves the System Data Mover (SDM), which is part of DFSMSdfp. DFSMSdfp in turn is a part of z/OS. z/OS Global Mirror maintains a consistent copy of the data asynchronously at a remote location. It can be implemented over unlimited distances. It is a combined hardware and software solution that offers data integrity and data availability that can be used as part of business continuance solutions, for workload movement, and for data migration.

The z/OS Global Mirror function is an optional licensed function (called Remote Mirroring for z Systems (RMZ)) of the DS8000 that enables the SDM to communicate with the primary DS8000. No z/OS Global Mirror license is required for the storage system in site H2. It can be any storage system that is supported by z/OS. However, consider that you might want to reverse the mirror, in which case, your auxiliary storage system needs a z/OS Global Mirror license as well.

5.3.8 z/OS Metro/Global Mirror

This mirroring configuration uses Metro Mirror to replicate primary site data to a location within the metropolitan area. It also uses z/OS Global Mirror to replicate primary site data to a location that is a long distance away. This configuration enables a z/OS three-site high-availability and disaster recovery solution for even greater protection against unplanned outages.

The basic operational characteristics of a z/OS Metro/Global Mirror implementation are shown in Figure 5-7.
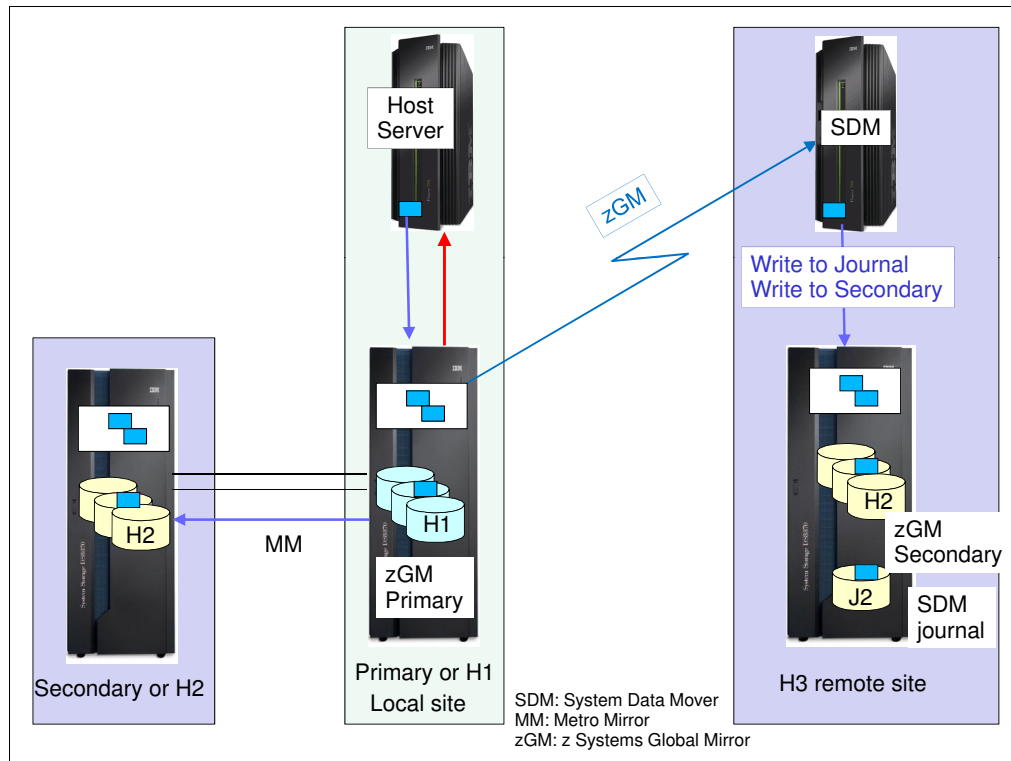
Figure 5-7   z/OS Metro/Global Mirror

With DS8880 R8.2, thanks to the Enhanced FICON Information Unit (IU) Pacing that uses the persistent pacing feature, a latency reduction factor of up to four times can be expected with certain channel programs like CCW chains of DB2 log writes in GDPS multi-site workload environment or other CCW chains with more than 16 commands. This feature applies only to 16 Gbps FC Host Adapters.

Important: Results can vary depending on the network capacity and the write workload.

5.3.9 GDPS on z/OS environments

GDPS is the IBM solution for managing large and complex environments and to keep the client data safe and consistent. It provides an easy interface to manage multiple sites with MGM pairs.

With its HyperSwap capability, GDPS is the ideal solution if you want 99.9999% availability.

GDPS easily monitors and manages your MGM pairs, and it also allows clients to run disaster recovery tests without affecting production. These features lead to faster recovery from real disaster events.

Using GDPS/Active-Active, the DS8880 R8.2 and later levels provide z/OS with a read-only access capability to a Metro Mirror secondary device, allowing a fast failover capability with the Zero Data Loss option. Reads are allowed if the device is full-duplex or suspended, but prevented if the device is in a pending state.

GDPS/Active-Active provides monitoring and management of the Zero Data Loss environment.

IBM InfoSphere® Data Replicator for DB2 enables you to read DB2 log data from a Metro Mirror secondary device.

GDPS functions include the following examples:

► The option to hot swap between primary and secondary Metro Mirror is managed concurrently with client operations. Operations can continue if a disaster or planned outage occurs.

► Disaster recovery management in a disaster at the primary site allows operations to restart at the remote site quickly and safely while data consistency is continuously monitored.

► GDPS freezes the Metro Mirror pairs if a problem with mirroring occurs. It restarts the copy process to secondaries after the problem is evaluated and solved, maintaining data consistency on all pairs.

For more information about GDPS, see *IBM GDPS Family: An Introduction to Concepts and Capabilities*, SG24-6374.

6

Designed for performance

This chapter describes the performance characteristics of the IBM DS8880 that relates to the physical and logical configuration. The considerations that are presented in this chapter can help you plan the physical and logical setup of the DS8880.

This chapter covers the following topics:

- ► DS8880 hardware: Performance characteristics
- ► Software performance: Synergy items
- ► Performance considerations for disk drives
- ► DS8000 superior caching algorithms
- ► Performance considerations for logical configurations
- ► I/O Priority Manager
- ► IBM Easy Tier
- ► Host performance and sizing considerations

6.1 DS8880 hardware: Performance characteristics

The IBM DS8880 is designed to support the most demanding business applications with its exceptional all-around performance and data throughput. These features are combined with world-class business resiliency and encryption capabilities to deliver a unique combination of high availability, performance, scalability, and security.

The DS8880 features IBM POWER8 processor-based server technology and uses a PCI Express (PCIe) I/O infrastructure. The DS8880 includes options for 6-core, 8-core, 16-core, 24-core, or 48-core processors for each processor complex. Up to 2 TB of system memory are available in the DS8880 for increased performance.

> **Note:** Six-core processors are only used in the DS8884. DS8886 uses the 8-core, 16-core, and 24-core processors. DS8888 uses 24-core or 48-core processors.

This section reviews the architectural layers of the DS8880 and describes the performance characteristics that differentiate the DS8880 from other disk systems.

6.1.1 Vertical growth and scalability

The DS8880 offers nondisruptive memory upgrades, and processor upgrades for the DS8886 for increased performance. For connectivity, 8 Gb and 16 Gb host adapters can be added without disruption. A wide range of storage options are available by using hard disk drives (HDDs), flash drives, or solid-state drives (SSDs), and high-performance flash enclosures (HPFEs). Expansion frames can also be added to the base 984, 985, 986, or 988 frame for increased capacity. Other advanced-function software features, such as IBM Easy Tier, I/O Priority Manager, and storage pool striping, contribute to performance potential.

For more information about hardware and architectural scalability, see Chapter 2, "IBM DS8880 hardware components and architecture" on page 23.

Figure 6-1 shows an example of how the DS8886 performance scales as the configuration changes from eight cores to 24 cores in an open systems database environment.
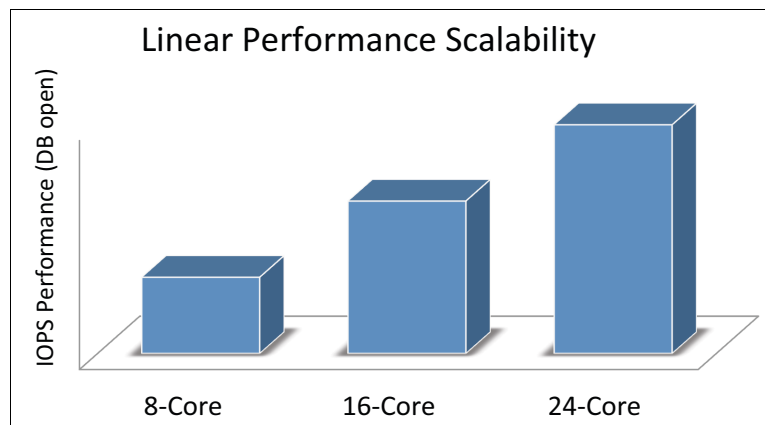
Figure 6-1   DS8886 linear performance scalability

6.1.2 POWER8

The DS8884 model 984 systems use 2U 8284-22A servers, which are based on 3.89 GHz POWER8 processors.

The DS8886 model 985 and 986 systems use 4U 8286-42A servers, which are based on 4.15 or 3.52 GHz POWER8 processors.

The DS8888 model 988 systems use 4U 8408-E8E servers, which are based on 3.03 GHz POWER8 processors.

Figure 6-2 shows the Hardware Management Console (HMC) server view.
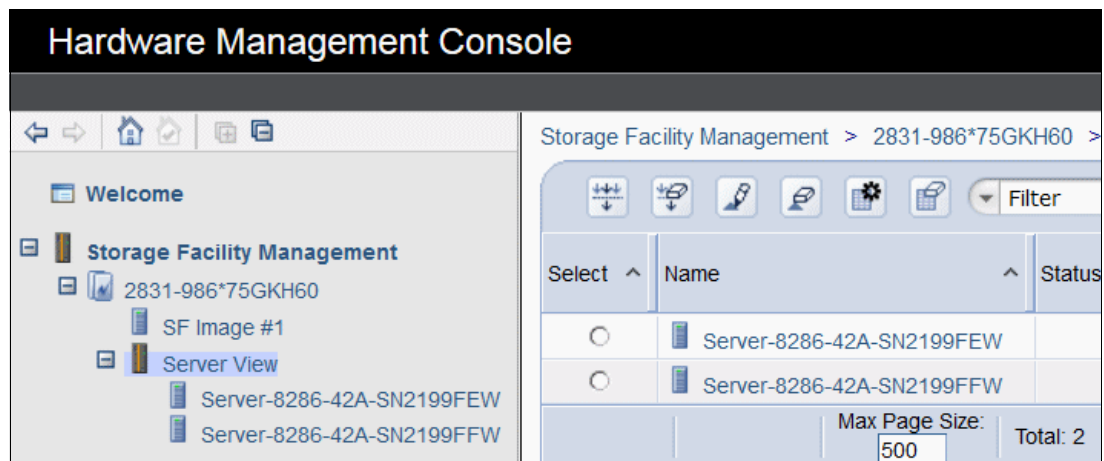
Figure 6-2   Power S824 servers 8286-42A (DS8886 based on POWER8)

The POWER8 is based on 22-nm technology, which allows higher density packaging than the 32-nm processor size on which the earlier POWER7 models were based. The number of transistors for each chip doubled from 2.1 bn to 4.2 bn. The POWER8 processor chip contains 1.5x the number of processor cores that were available in POWER7+.

Each POWER8 processor core contains 32 KB of Level 1 (L1) instruction cache, 64 KB of L1 data cache, and 512 KB of L2 cache, which is double the L2 cache that is available in POWER7+. The POWER8 processor chip contains 96 MB of embedded dynamic random access memory (eDRAM) L3 cache, which is divided into multiple 8 MB regions for each core. These regions can be dynamically shared between processor cores. New to POWER8 is an additional L4 cache, which is contained within the memory buffer chips on the custom memory dual inline memory modules (CDIMMs).

Simultaneous multithreading (SMT) allows a single physical processor core to simultaneously dispatch instructions from more than one hardware thread context. With its larger and faster caches, and faster memory access, POWER8 processors can provide eight concurrent SMT threads for each core, with up to 192 threads for each internal server.

Figure 6-3 shows the POWER8 physical processor chip layout.
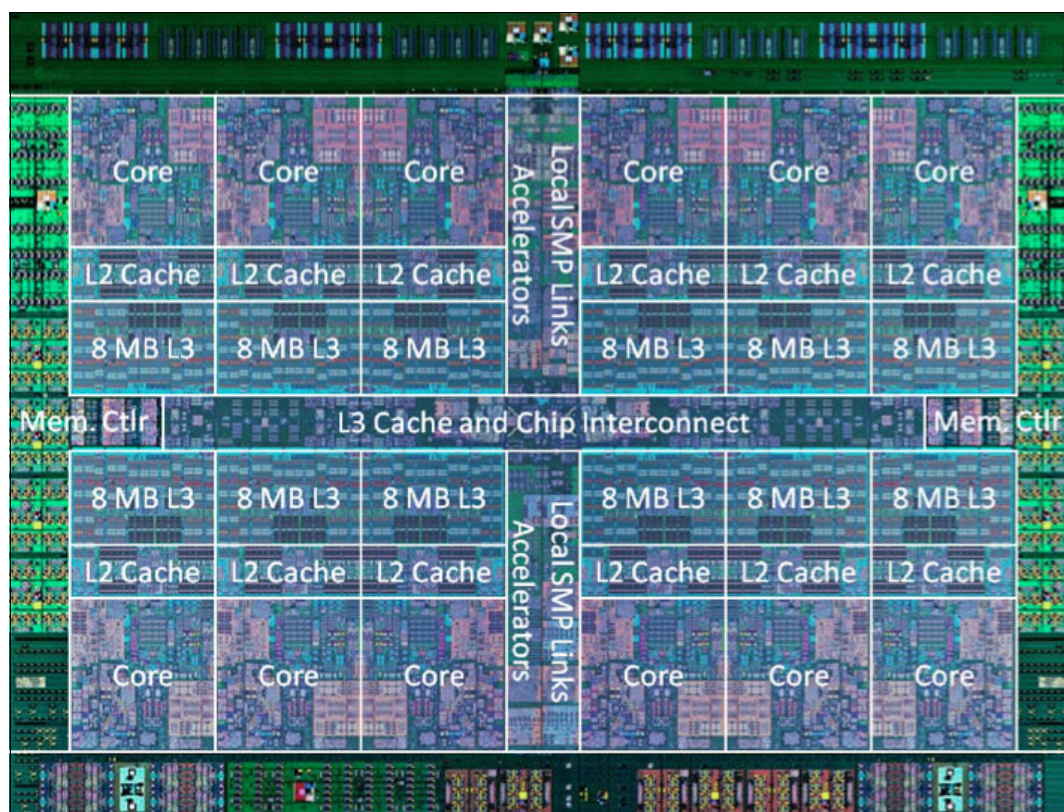
Figure 6-3   Twelve-way POWER8 processor module

PCI Express Gen3

POWER 8 implements PCIe Gen3 system bus architecture. Each POWER8 processor module has 48 PCIe Gen3 lanes, each running at 8 Gbps full-duplex, for an aggregate internal bandwidth of 192 GBps full-duplex for each 8286-42A internal server.

PCIe Gen3 I/O enclosures

To complement the improved I/O bandwidth that is provided by POWER8, the DS8880 also includes new PCIe Gen3-based I/O enclosures. Each I/O enclosure connects to both internal servers over a pair of x8 PCIe Gen3 cables, each providing 8 GBps connectivity. An extra pair of PCIe connectors on the I/O enclosure provides PCIe connections to the HPFE Microbays. Each pair of I/O enclosures supports up to two pairs of HPFE Gen2 enclosures, and up to four enclosures for certain all-flash models.

6.1.3 High-Performance Flash Enclosure

The High-Performance Flash Enclosure architecture is specifically designed for flash. Announced with DS8000 Release 8.2, HPFE Gen2 provides the highest standard of flash performance that is available in the DS8880. The Microbays in each HPFE Gen2 pair are attached through two PCI Gen-3 connections to the DS8880 processor complexes.

Each Microbay contains a Flash RAID adapter, which is dual-core PowerPC based and hence can do RAID parity processing with speed and amount of I/O that go far beyond what a usual Device Adapter could handle. Each contains, among other components, a PCIe switch card that carries the signal forward from the Flash RAID adapters over through the PCIe Gen-3

into the processor complexes. The Flash RAID adapters go with 8 SAS connections per Microbay pair to a pair of the specific flash enclosures that hold SAS expanders and the flash cards.

All components are optimized for the performance capabilities of flash-based storage.

The flash cards that are used in the HPFE are encryption-capable, and are packaged in a 2.5-inch form factor. They feature the Enterprise Multi-Level Cell (eMLC) technology. IBM was the first server vendor to provide this flash technology option, which blends enterprise class performance and reliability characteristics with the more cost-effective characteristics of MLC flash storage.

The 2.5-inch NAND flash cards that are used in the HPFE build on these base use advances in device controller flash memory management and advances in eMLC technology. Like the earlier IBM eMLC SSDs, flash cards are designed to provide high sustained performance levels, and extended endurance or reliability. The eMLC flash modules were designed to provide 24×7×365 usage even running write-intensive levels for at least five years. Typical client usage is expected to be much lower, especially for the average percentage of writes. Therefore, drive lifespan can be much longer.

HPFE arrays can coexist with Fibre Channel-attached Flash Drive arrays within the same extent pool. Both storage types are treated by Easy Tier as the highest performance tier, which is Tier 0. However, Easy Tier is able to differentiate between the performance capabilities of the two and run intra-tier rebalancing. So, "hotter" I/O-intensive extents of the volumes are moved to the high-performance flash arrays within Tier 0. For more information, see *IBM DS8000 EasyTier*, REDP-4667.

6.1.4 Switched Fibre Channel Arbitrated Loops

Standard drive enclosures connect to the device adapters (DAs) by using Fibre Channel Arbitrated Loop (FC-AL) topology. Ordinarily, this configuration creates arbitration issues within the loops. To overcome the arbitration issues, IBM employs a switch-based FC-AL topology in the DS8880. By using this approach, individual switched loops are created for each drive interface, providing isolation from the other drives, and routing commands and data only to the destination drives.

The drive enclosure interface cards contain switch logic that receives the FC-AL protocol from the device adapters, and attaches to each of the drives by using a point-to-point connection. The arbitration message for the drive is captured in the switch, processed, and then propagated to the intended drive without routing it through all of the other drives in the loop.

Each drive has two switched point-to-point connections to both enclosure interface cards, which in turn each connect to both DAs. This configuration has these benefits:

► This architecture doubles the bandwidth over conventional FC-AL implementations because no arbitration competition exists and no interference occurs between one drive and all the other drives. Each drive effectively has two dedicated logical paths to each of the DAs that allow two concurrent read operations and two concurrent write operations.

► In addition to superior performance, the reliability, availability, and serviceability (RAS) are improved in this setup when compared to conventional FC-AL. The failure of a drive is detected and reported by the switch. The switch ports distinguish between intermittent failures and permanent failures. The ports understand intermittent failures, which are recoverable, and collect data for predictive failure statistics. If one of the switches fails, a disk enclosure service processor detects the failing switch and reports the failure by using the other loop. All drives can still connect through the remaining switch.

FC-AL switched loops are shown in Figure 6-4.

DS8000 Storage Enclosure with Switched Dual Loops
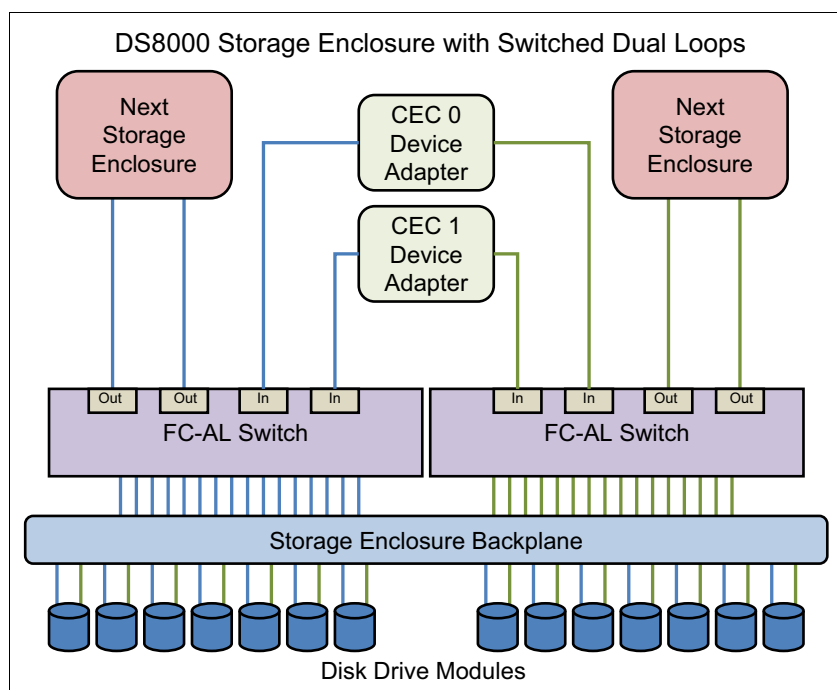
Figure 6-4   Switched FC-AL drive loops

A virtualization approach that is built on top of the high-performance architectural design contributes even further to enhanced performance, as described in Chapter 4, "Virtualization concepts" on page 97.

6.1.5 Fibre Channel device adapter

The DS8880 supports standard drive enclosures that connect to the internal servers by using a pair of DAs. The device adapters provide RAID-6, RAID-10, and RAID-5 capabilities by using PowerPC technology, with an application-specific integrated circuit (ASIC). Device adapters are PCIe Gen2-based. Each adapter provides four 8 Gbps Fibre Channel loops, and the adapter pair provides two independent fabrics to the drive enclosures.

The DS8880 uses the DAs in *Split Affinity* mode. Split Affinity means that each central processor complex (CPC) unit uses one device adapter in every I/O enclosure. This configuration allows both device adapters in an I/O enclosure to communicate concurrently because each DA uses a different PCIe connection between the I/O enclosure and the CPC. This design significantly improves performance when compared to the approach that was used before the DS8870.

6.1.6 16-Gbps and 8-Gbps Fibre Channel host adapters

The DS8880 supports up to 32 Fibre Channel host adapters (HAs), with either four or eight FC ports for each adapter. Each port can be independently configured to support Fibre Channel connection (FICON) or Fibre Channel Protocol (FCP).

The DS8880 offers three Fibre Channel host adapter types: 16 Gbps 4-port, 8 Gbps 8-port, and 8 Gbps 4-port. Each adapter type is available in both longwave (LW) and shortwave (SW) versions. The DS8880 I/O bays support up to four HAs for each bay, allowing up to 128 ports maximum for each storage system. This configuration results in a theoretical aggregated host I/O bandwidth of 128 x 16 Gbps. Each port provides industry-leading throughput and I/O rates for FICON and FCP.

The host adapters that are available in the DS8880 include the following characteristics:

▶ 8 Gbps HBAs:

 – Four or eight Fibre Channel ports
 – Gen2 PCIe interface
 – Dual-core PowerPC processor
 – Negotiation to 8, 4, or 2 Gbps (1 Gbps is not possible.)

▶ 16 Gbps HBAs:

 – Four Fibre Channel ports
 – Gen2 PCIe interface
 – Quad-core PowerPC processor
 – Negotiation to 16, 8, or 4 Gbps (1 and 2 Gbps are not possible.)

The DS8880 supports the intermixture of 16 Gbps and 8 Gbps HBAs. Hosts with slower FC speeds are still supported if their HBAs are connected through a switch.

With FC adapters that are configured for FICON, the DS8000 series provides the following configuration capabilities:

▶ Fabric or point-to-point topologies

▶ A maximum of 128 host adapter ports, depending on the DS8880 system memory and processor features

▶ A maximum of 509 logins for each FC port

▶ A maximum of 8192 logins for each storage unit

▶ A maximum of 1280 logical paths on each FC port

▶ Access to all 255 control-unit images (65,280 CKD devices) over each FICON port

▶ A maximum of 512 logical paths for each control unit image

IBM z13 servers support 32,000 devices per FICON host channel, whereas IBM zEnterprise® EC12 and IBM zEnterprise BC12 servers support 24,000 devices per FICON host channel. Earlier z Systems servers support 16,384 devices per FICON host channel. To fully access 65,280 devices, it is necessary to connect multiple FICON host channels to the storage system. You can access the devices through a Fibre Channel switch or FICON director to a single storage system FICON port.

The 16 Gbps FC host adapter doubles the data throughput of 8 Gbps links. Compared to 8 Gbps adapters, the 16 Gbps adapters provide improvements in full adapter I/O per second (IOPS) and reduce latency.

6.2 Software performance: Synergy items

Many performance features in the DS8880 work together with the software on IBM hosts and are collectively referred to as *synergy items*. These items allow the DS8880 to cooperate with the host systems to benefit the overall performance of the systems.

6.2.1 Synergy with Power Systems

The IBM DS8880 can work in cooperation with Power Systems to provide the following performance enhancement functions.

End-to-end I/O priority: Synergy with AIX and DB2 on Power Systems

End-to-end I/O priority is an IBM-requested addition to the Small Computer System Interface (SCSI) T10 standard. This feature allows trusted applications to override the priority that is given to each I/O by the operating system. This feature is only applicable to raw volumes (no file system) and with the 64-bit kernel. Currently, AIX supports this feature with DB2. The priority is delivered to the storage subsystem in the FCP Transport Header.

The priority of an AIX process can be 0 (no assigned priority) or any integer value from 1 (highest priority) to 15 (lowest priority). All I/O requests that are associated with a process inherit its priority value. However, with end-to-end I/O priority, DB2 can change this value for critical data transfers. At the DS8880, the host adapter gives preferential treatment to higher priority I/O, which improves performance for specific requests that are deemed important by the application, such as requests that might be prerequisites for other requests (for example, DB2 logs).

Cooperative caching: Synergy with AIX and DB2 on Power Systems

Another software-related performance item is cooperative caching, a feature that provides a way for the host to send cache management hints to the storage facility. Currently, the host can indicate that the information that was recently accessed is unlikely to be accessed again soon. This status decreases the retention period of the data that is cached at the host, which allows the subsystem to conserve its cache for data that is more likely to be reaccessed. Therefore, the cache hit ratio is improved.

With the implementation of cooperative caching, the AIX operating system allows trusted applications, such as DB2, to provide cache hints to the DS8000. This ability improves the performance of the subsystem by keeping more of the repeatedly accessed data cached in the high-performance flash at the host. Cooperative caching is supported in IBM AIX for Power Systems with the Multipath I/O (MPIO) Path Control Module (PCM) that is provided with the Subsystem Device Driver (SDD). It is only applicable to raw volumes (no file system) and with the 64-bit kernel.

Long busy wait host tolerance: Synergy with AIX on Power Systems

The SCSI T10 standard includes support for *SC*SI long busy wait, which provides a way for the target system to specify that it is busy and how long the initiator must wait before an I/O is tried again.

This information, provided in the FCP status response, prevents the initiator from trying again too soon. This delay, in turn, reduces unnecessary requests and potential I/O failures that can be caused by exceeding a set threshold for the number of times it is tried again. IBM AIX for Power Systems supports SCSI long busy wait with MPIO. It is also supported by the DS8880.

6.2.2 Synergy with z Systems

The DS8880 is able to work in cooperation with z Systems to provide several performance enhancement functions. The following section gives a brief overview of these synergy items.

Parallel access volume and HyperPAV

Parallel access volume (PAV) is included in the DS8880 z-Synergy Services license group, for the z/OS and z/VM operating systems. With PAV, you can run multiple I/O requests to a volume at the same time. With dynamic PAV, the z/OS Workload Manager (WLM) controls the assignment of alias addresses to base addresses. The number of alias addresses defines the parallelism of I/Os to a volume. HyperPAV is an extension to PAV where any alias address from a pool of addresses can be used to drive the I/O, without WLM involvement.

Cross Control Unit PAV or SuperPAV

Starting with DS8880 release 8.1, HyperPAV has been further enhanced. If all alias addresses are busy, z/OS can now use an alias address from another logical control unit (LCU), providing these conditions are met:

► The alias devices are assigned to the same DS8000 server (odd or even LSS)
► The alias devices share a path group on the z/OS host

The system now has a lager pool of alias addresses for bursts of I/O to a base volume and thus reduce queuing of I/O. This feature is known as Cross-CU PAV or SuperPAV. See "SuperPAV" on page 111.

DS8000 I/O Priority Manager with z/OS Workload Manager

I/O Priority Manager, together with z/OS Workload Manager (zWLM), enable more effective storage consolidation and performance management when different workloads share a common disk pool (extent pool). This function, tightly integrated with zWLM, is intended to improve disk I/O performance for important workloads. It drives I/O prioritization to the disk system by allowing WLM to give priority to the system's resources (disk arrays) automatically when higher priority workloads are not meeting their performance goals. The I/O of less prioritized workloads to the same extent pool are slowed down to give the higher prioritized workload a higher share of the resources, mainly the disk drives. Integration with zWLM is exclusive to the DS8000 and z Systems. For more information about I/O Priority Manager, see *DS8000 I/O Priority Manager*, REDP-4760.

Easy Tier support

IBM Easy Tier is an intelligent data placement algorithm in the DS8000 that is designed to support both open systems and z Systems workloads.

Specifically for z Systems, IBM Easy Tier provides an application programming interface (API) through which z Systems applications (zDB2 initially) can communicate performance requirements for optimal data set placement. The application hints set the intent, and Easy Tier then moves the data set to the correct tier. For more information, see 6.7, "IBM Easy Tier" on page 173.

Extended Address Volumes

This capability can help relieve address constraints to support large storage capacity needs by addressing the capability of z Systems environments to support volumes that can scale up to approximately 1 TB (1,182,006 cylinders).

High Performance FICON for z Systems

High Performance FICON for z Systems (zHPF) is an enhancement of the FICON channel architecture. You can reduce the impact of FICON channel I/O traffic by using zHPF with the FICON channel, the z/OS operating system, and the storage system. zHPF allows the storage system to stream the data for multiple commands back in a single data transfer section for I/Os that are initiated by various access methods, This process improves the channel throughput on small block transfers.

zHPF is an optional feature of the DS8880, included in the z-Synergy Services license group. Recent enhancements to zHPF include Extended Distance capability, zHPF List Pre-fetch, Format Write, and zHPF support for sequential access methods. The DS8880 with zHPF and z/OS V1.13 offers I/O performance improvements for certain I/O transfers for workloads that use these methods:

► Queued sequential access method (QSAM)
► Basic partitioned access method (BPAM)
► Basic sequential access method (BSAM)

With 16 Gbps host adapters on the DS8880 and z Systems server IBM z13/z13s™ (and later) channels, zHPF Extended Distance II supports heavy write I/Os over an extended distance of up to 100 km (62 miles). The result is an increase of the write throughput by 50% or better.

DB2 list prefetch

zHPF is enhanced to support DB2 list prefetch. The enhancements include a new cache optimization algorithm that can greatly improve performance and hardware efficiency. When combined with the latest releases of z/OS and DB2, it can demonstrate up to a 14x - 60x increase in sequential or batch-processing performance. All DB2 I/Os, including format writes and list prefetches, are eligible for zHPF. In addition, DB2 can benefit from the new caching algorithm at the DS8000 level, which is called List Prefetch Optimizer (LPO). For more information about list prefetch, see *DB2 for z/OS and List Prefetch Optimizer*, REDP-4862.

DB2 castout acceleration

DB2 offers a data sharing feature to improve scalability and availability over separate independent DB2 systems. DB2 uses the concept of a group buffer pool (GBP). It is a z/OS coupling facility structure to cache data that is accessed by multiple applications and ensure consistency.

In environments with write intensive activity, the group buffer pool fills up quickly and must be destaged to storage. This process is bursty and can cause performance problems on read threads. The process of writing pages from the GBP to disk is known as *castout*.

During a castout, DB2 will typically generate long chains of writes, resulting in multiple IOs. In a Metro Mirror environment, it is imperative that the updates go to the secondary in order, and before availability of the DS8000 Release 8.1 code, each IO in the chain was synchronized individually with the secondary.

With Release 8.1 (bundles 88.11.25.0) and later, the DS8000 can be notified through the DB2 Media Manager that the multiple IOs in a castout can be treated as a single logical IO, although there are multiple embedded IOs. In other words, the data hardening requirement is for the entire IO chain. This new process brings significant response time reduction.

This enhancement only applies for zHPF IO.

Quick initialization (z Systems)

The DS8880 supports quick volume initialization for z Systems environments, which can help clients who frequently delete volumes, allowing capacity to be reconfigured without waiting for initialization. Quick initialization initializes the data logical tracks or block within a specified extent range on a logical volume with the appropriate initialization pattern for the host.

Normal read and write access to the logical volume is allowed during the initialization process. Depending on the operation, the quick initialization can be started for the entire logical volume or for an extent range on the logical volume.

Quick initialization improves device initialization speeds and allows a Copy Services relationship to be established after a device is created.

zHyperwrite

In a Metro Mirror environment, all writes (including DB2 log writes) are mirrored synchronously to the secondary device, which increases transaction response times. *zHyperwrite* enables DB2 log writes to be performed to the primary and secondary volumes in parallel, which reduces DB2 log write response times. Implementation of zHyperwrite requires that HyperSwap is enabled through either IBM Geographically Dispersed Parallel Sysplex (IBM GDPS) or IBM Tivoli Storage Productivity Center for Replication.

Further DS8880 synergy items with z Systems

The available 16 Gbps adapters offer additional end-to-end z Systems and DS8880 synergy items:

► Forward Error Correction (FEC) is a protocol that is designed to capture errors that are generated during data transmission. Both the z Systems z13 (and later) and DS8880 extend the use of FEC to complete end-to-end coverage for 16 Gbps links and preserve data integrity with more redundancy.

► Fibre Channel Read Diagnostic Parameters (RDP) improve the end-to-end link fault isolation for 16 Gbps links on the z Systems z13/z13s and DS8880. RDP data provides the optical signal strength, error counters, and other critical information that is crucial to determine the quality of the link.

► FICON Dynamic Routing (FIDR) is a Systems z13 (and later) and DS8880 feature that supports the use of dynamic routing policies in the switch to balance loads across inter-switch links (ISLs) on a per I/O basis.

► Fabric I/O Priority provides end-to-end quality of service (QoS). It is a unique synergy feature between the z Systems z/OS Workload Manager, Brocade SAN Fabric, and the DS8880 system that is designed to manage QoS on a single I/O level.

For more information, see *IBM DS8870 and IBM z Systems Synergy*, REDP-5186, and *Get More Out of Your IT Infrastructure with IBM z13 I/O Enhancements*, REDP-5134.

6.3 Performance considerations for disk drives

When you are planning your system, determine the number and type of ranks that are required based on the needed capacity and on the workload characteristics in terms of access density, read-to-write ratio, and cache hit rates. These factors are weighed against the performance and capacity characteristics of the physical storage.

Current 15K revolutions per minute (RPM) disks, for example, provide an average seek time of approximately 3 ms and an average latency of 2 ms. For transferring only a small block, the transfer time can be neglected. This time is an average 5 ms for each random disk I/O operation or 200 IOPS. Therefore, a combined number of eight disks, which is the case for a DS8000 array, potentially sustain 1,600 IOPS when they are spinning at 15K RPM. Reduce the number by 12.5% (1,400 IOPS) when you assume a spare drive in the array site.

On the host side, consider an example with 1,000 IOPS from the host, a read-to-write ratio of 70/30, and 50% read cache hits. This configuration leads to the following IOPS numbers:

► 700 read IOPS.

► 350 read I/Os must be read from disk (based on the 50% read cache hit ratio).

► 300 writes, assuming RAID 5, result in 1,200 disk operations because of the RAID 5 write penalty (read old data and parity, write new data and parity).

► A total of 1,550 disk I/Os.

With 15K RPM disk drive modules (DDMs) running 1,000 random IOPS from the server, you complete 1,550 I/O operations on disk compared to a maximum of 1,600 operations for 7+P configurations or 1,400 operations for 6+P+S configurations. Therefore, in this scenario, 1,000 random I/Os from a server with a certain read-to-write ratio and a certain cache hit ratio saturate the disk drives. This scenario assumes that server I/O is purely random. When sequential I/Os exist, track-to-track seek times are much lower and higher I/O rates are possible. It also assumes that reads have a cache-hit ratio of only 50%. With higher hit ratios, higher workloads are possible. These considerations show the importance of intelligent caching algorithms as used in the DS8000. These algorithms are described in 6.4, "DS8000 superior caching algorithms" on page 159.

> **Important:** When a storage system is sized, consider the capacity and the number of disk drives that are needed to satisfy the space requirements, and also the performance capabilities that will satisfy the IOPS requirements.

For a single disk drive, various disk vendors provide the disk specifications on their websites. Because the access times for the disks are the same for the same RPM speeds, but they have different capacities, the I/O density is different. A 300 GB 15K RPM disk drive can be used for access densities up to, and slightly over, 0.5 I/O per GB. For 600-GB drives, it is approximately 0.25 I/O per GB. Although this discussion is theoretical in approach, it provides a first estimate.

After the speed of the disk is decided, the capacity can be calculated based on your storage capacity needs and the effective capacity of the RAID configuration that you use.

6.3.1 Flash storage

From a performance point of view, the best performing choice for your DS8880 storage is flash storage. Flash has no moving parts (no spinning platters and no actuator arm) and a lower energy consumption. The performance advantages are the fast seek time and average access time. Flash storage is targeted at applications with heavy IOPS, bad cache hit rates, and random access workload that necessitate fast response times. Database applications with their random and intensive I/O workloads are prime candidates for deployment on flash.

Using DS8880, either an all-flash strategy is possible, or the DS8000 can also accommodate a hybrid configuration that combines different disk drive tiers. Within flash, you can use either classical flash drives (SSDs) that can also have large capacities today or HPFE with flash cards for performance-optimized solutions.

Flash cards

Flash cards are available in the HPFE feature. Flash cards offer the highest performance option that is available in the DS8880. Integrated dual Flash RAID Adapters with native x8 PCIe Gen-3 attachments provide high-bandwidth connectivity, without the protocol impact of Fibre Channel. The DS8880 offers these maximums:

- ► Up to 384 flash cards in the DS8888
- ► Up to 192 flash cards in the DS8886
- ► Up to 96 flash cards in the DS8884

Flash drives

Flash drives offer extra flash storage capacity by using small-form-factor (SFF) drives that are installed in Fibre Channel-attached standard drive enclosures. The DS8880 supports up to 1536 flash drives (DS8886 model 985), in addition to the HPFE flash cards.

6.3.2 Enterprise drives

Enterprise drives provide high performance and cost-effective storage for various workloads. Enterprise drives rotate at 15,000 or 10,000 RPM. If an application requires high-performance data throughput and continuous, I/O-intensive operations, enterprise drives provide the best price-performance option.

> **Tip:** The 1.8-TB drives, introduced with release 8.1, offer a 33% increase in capacity over 1.2-TB models, with equivalent or better 4 KB random and sequential read/write performance. Rebuild rates are equivalent to the 1.2-TB drives.

6.3.3 Nearline drives

When disk capacity is a priority, nearline drives are the largest of the drives that are available for the DS8880. Given their large capacity, and lower (7,200 RPM) rotational speed, as compared to enterprise drives, nearline drives are not intended to support high-performance or I/O-intensive applications that demand fast random data access. For sequential workloads, nearline drives can be an attractive storage option. These nearline drives offer a cost-effective option for lower-priority data, such as fixed content, data archival, reference data, and nearline applications that require large amounts of storage capacity for lighter workloads. These drives are meant to complement, not compete with, enterprise drives.

The 6-TB nearline drives offer a 33% increase in capacity, over 4-TB models, with equivalent or better 4 KB random read and write performance, and significantly better sequential performance. Rebuild rates are also equivalent or better than 4-TB nearline drives.

6.3.4 RAID level

The DS8000 series offers RAID 6, RAID 10, and RAID 5. By default, the DS Storage Manager GUI suggests a configuration with RAID 6.

RAID 5

RAID 5 has often been used because it provides good performance for random and sequential workloads, and it does not need much more storage for redundancy (one parity drive). The DS8000 series can detect sequential workload. When a complete stripe is in cache for destage, the DS8000 series switches to a RAID 3-like algorithm. Because a complete stripe must be destaged, the old data and parity do not need to be read. Instead, the new parity is calculated across the stripe, and the data and parity are destaged to disk. This configuration provides good sequential performance. A random write causes a cache hit, but the I/O is not complete until a copy of the write data is put in non-volatile storage (NVS). When data is destaged to disk, a write in RAID 5 causes the following four disk operations, the so-called *write penalty*:

► Old data and the old parity information must be read.
► New parity is calculated in the device adapter.
► Data and parity are written to disk.

Most of this activity is hidden to the server or host because the I/O is complete when data enters cache and NVS.

However, due to the much higher availability that RAID 6 offers, the usage of RAID 5 is discouraged, and is disabled on drives larger than 1 TB because these also have larger rebuild times.

RAID 6

RAID 6 increases data fault tolerance. It can tolerate two disk failures in the same rank, as compared to RAID 5, which is single disk fault tolerant. RAID 6 uses a second independent distributed parity scheme (dual parity). RAID 6 provides a read performance that is similar to RAID 5, but RAID 6 has a larger write penalty than RAID 5 because it must calculate and write a second parity stripe.

Consider RAID 6 in situations where you might consider RAID 5, but need increased reliability. RAID 6 was designed for protection during longer rebuild times on larger capacity drives to cope with the risk of a second drive failure within a rank during rebuild. RAID 6 has the following characteristics:

► Sequential Read: About 99% x RAID 5 rate
► Sequential Write: About 65% x RAID 5 rate
► Random 4 K 70% R /30% W IOPS: About 55% x RAID 5 rate

If two disks fail in the same rank, the performance degrades during the rebuild.

> **Important:** Configure large-capacity drives as RAID 6 arrays. RAID 6 is advised as the default for all drive sizes now.

RAID 10

A workload that consists of mainly *random writes*, which is usually the case when you have a higher write ratio, benefits from RAID 10. Here, data is striped across several disks and mirrored to another set of disks. A write causes only two disk operations when compared to four operations of RAID 5, or six of RAID 6. RAID 10 requires nearly twice as many disk drives (and twice the cost) for the same capacity when compared to RAID 5, but it can achieve four times greater random write throughput. Therefore, it is definitely worth considering the use of RAID 10 for high-performance random write workloads.

> **Tip:** Consider RAID 10 for random-write ratios over 35%.

The decision to configure capacity as RAID 10, RAID 6, or (for small drives) RAID 5, and the amount of capacity to configure for each type, can be made at any time. RAID 6, RAID 5, and RAID 10 arrays can be intermixed within a single system. In addition, the physical capacity can be logically reconfigured later (for example, RAID 6 arrays can be reconfigured into RAID 10 arrays). However, the arrays must first be emptied. Changing the RAID level is not permitted when logical volumes exist.

> **Important:** For more information about important restrictions on DS8880 RAID configurations, see 3.5.1, "RAID configurations" on page 82.

6.4 DS8000 superior caching algorithms

Most, if not all, high-end disk systems have an internal cache that is integrated into the system design. The DS8880 can be equipped with up to 2,048 GB of system memory, most of which is configured as cache. The DS8880 offers twice the available system memory than was available in the DS8870.

With its POWER8 processors, the server architecture of the DS8880 makes it possible to manage such large caches with small cache segments of 4 KB (and large segment tables). The POWER8 processors have the power to support sophisticated caching algorithms, which contribute to the outstanding performance that is available in the IBM DS8880. These algorithms and the small cache segment size optimize cache hits and cache utilization. Cache hits are also optimized for different workloads, such as sequential workloads and transaction-oriented random workloads, which might be active at the same time. Therefore, the DS8880 provides excellent I/O response times.

Write data is always protected by maintaining a second copy of cached write data in the NVS of the other internal server until the data is destaged to disks.

6.4.1 Sequential Adaptive Replacement Cache

The DS8000 series uses the Sequential Adaptive Replacement Cache (SARC) algorithm, which was designed by IBM Storage Development in partnership with IBM Research. It is a self-tuning, self-optimizing solution for a wide range of workloads with a varying mix of sequential and random I/O streams. SARC is inspired by the Adaptive Replacement Cache (ARC) algorithm and inherits many of its features. For more information about ARC, see "Outperforming LRU with an adaptive replacement cache algorithm" by N. Megiddo, et al., in IEEE Computer, volume 37, number 4, pages 58 - 65, 2004. For more information about SARC, see "SARC: Sequential Prefetching in Adaptive Replacement Cache" by Binny Gill, et al., in "Proceedings of the USENIX 2005 Annual Technical Conference", pages 293 - 308.

SARC attempts to determine the following cache characteristics:

► When data is copied into the cache
► Which data is copied into the cache
► Which data is evicted when the cache becomes full
► How the algorithm dynamically adapts to different workloads

The DS8000 series cache is organized in 4-KB pages that are called *cache pages* or *slots*. This unit of allocation (which is smaller than the values that are used in other storage systems) ensures that small I/Os do not waste cache memory.

The decision to copy data into the DS8880 cache can be triggered from the following policies:

► Demand paging

Eight disk blocks (a 4 K cache page) are brought in only on a cache miss. Demand paging is always active for all volumes and ensures that I/O patterns with locality discover at least recently used data in the cache.

► Prefetching

Data is copied into the cache speculatively even before it is requested. To prefetch, a prediction of likely data accesses is needed. Because effective, sophisticated prediction schemes need an extensive history of page accesses (which is not feasible in real systems), SARC uses prefetching for sequential workloads. Sequential access patterns naturally arise in video-on-demand, database scans, copy, backup, and recovery. The goal of sequential prefetching is to detect sequential access and effectively prefetch the likely cache data to minimize cache misses. Today, prefetching is ubiquitously applied in web servers and clients, databases, file servers, on-disk caches, and multimedia servers.

For prefetching, the cache management uses tracks. A track is a set of 128 disk blocks (16 cache pages). To detect a sequential access pattern, counters are maintained with every track to record whether a track was accessed together with its predecessor. Sequential prefetching becomes active only when these counters suggest a sequential access pattern. In this manner, the DS8880 monitors application read I/O patterns and dynamically determines whether it is optimal to stage into cache the following I/O elements:

► Only the page requested
► The page that is requested plus the remaining data on the disk track
► An entire disk track (or a set of disk tracks) that was not requested

The decision of when and what to prefetch is made in accordance with the Adaptive Multi-stream Prefetching (AMP) algorithm. This algorithm dynamically adapts the amount and timing of prefetches optimally on a per-application basis (rather than a system-wide basis). For more information about AMP, see 6.4.2, "Adaptive Multi-stream Prefetching" on page 162.

To decide which pages are evicted when the cache is full, sequential and random (non-sequential) data is separated into separate lists. The SARC algorithm for random and sequential data is shown in Figure 6-5.
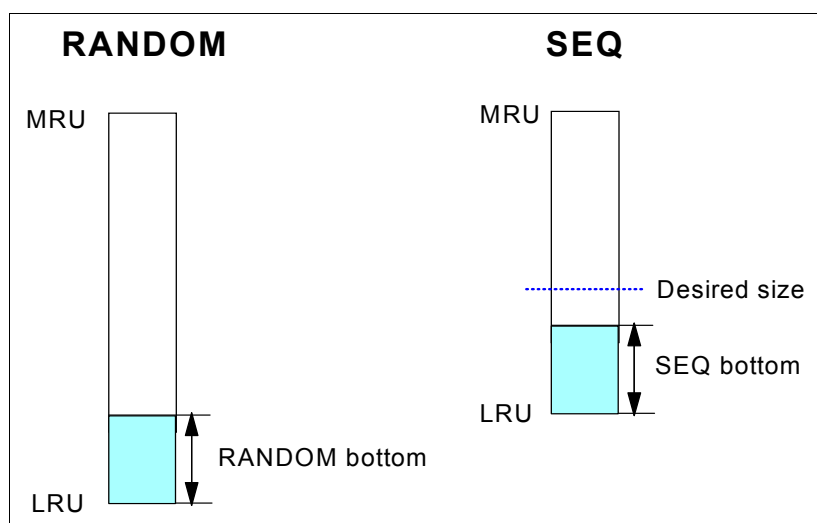
Figure 6-5   Sequential Adaptive Replacement Cache

A page that was brought into the cache by simple demand paging is added to the head of Most Recently Used (MRU) section of the RANDOM list. Without further I/O access, it goes down to the bottom of Least Recently Used (LRU) section. A page that was brought into the cache by a sequential access or by sequential prefetching is added to the head of MRU of the SEQ list and then goes in that list. Other rules control the migration of pages between the lists so that the system does not keep the same pages in memory twice.

To follow workload changes, the algorithm trades cache space between the RANDOM and SEQ lists dynamically and adaptively. This function makes SARC scan resistant so that one-time sequential requests do not pollute the whole cache. SARC maintains a wanted size parameter for the sequential list. The wanted size is continually adapted in response to the workload. Specifically, if the bottom portion of the SEQ list is more valuable than the bottom portion of the RANDOM list, the wanted size is increased. Otherwise, the wanted size is decreased. The constant adaptation strives to make optimal use of limited cache space and delivers greater throughput and faster response times for a specific cache size.

Additionally, the algorithm dynamically modifies the sizes of the two lists and the rate at which the sizes are adapted. In a steady state, pages are evicted from the cache at the rate of cache misses. A larger (or smaller) rate of misses leads to a faster (or slower) rate of adaptation.

Other implementation details account for the relationship of read and write (NVS) cache, efficient destaging, and the cooperation with Copy Services. In this manner, the DS8880 cache management goes far beyond the usual variants of the LRU/Least Frequently Used (LFU) approaches, which are widely used in other storage systems on the market.

6.4.2 Adaptive Multi-stream Prefetching

SARC dynamically divides the cache between the RANDOM and SEQ lists, where the SEQ list maintains pages that are brought into the cache by sequential access or sequential prefetching.

In the DS8880, AMP, an algorithm that was developed by IBM Research, manages the SEQ list. AMP is an autonomic, workload-responsive, self-optimizing prefetching technology that adapts the amount of prefetch and the timing of prefetch on a per-application basis to maximize the performance of the system. The AMP algorithm solves the following problems that plague most other prefetching algorithms:

► Prefetch wastage occurs when prefetched data is evicted from the cache before it can be used.

► Cache pollution occurs when less useful data is prefetched instead of more useful data.

By wisely choosing the prefetching parameters, AMP provides optimal sequential read performance and maximizes the aggregate sequential read throughput of the system. The amount that is prefetched for each stream is dynamically adapted according to the application's needs and the space that is available in the SEQ list. The timing of the prefetches is also continuously adapted for each stream to avoid misses and any cache pollution.

SARC and AMP play complementary roles. SARC carefully divides the cache between the RANDOM and the SEQ lists to maximize the overall hit ratio. AMP manages the contents of the SEQ list to maximize the throughput that is obtained for the sequential workloads. SARC affects cases that involve both random and sequential workloads. However, AMP helps any workload that has a sequential read component, including pure sequential read workloads.

AMP dramatically improves performance for common sequential and batch processing workloads. It also provides excellent performance synergy with DB2 by preventing table scans from being I/O-bound and improves performance of index scans and DB2 utilities, such as `Copy` and `Recover`. Furthermore, AMP reduces the potential for array hot spots that result from extreme sequential workload demands.

For more information about AMP and the theoretical analysis for its optimal usage, see "AMP: Adaptive Multi-stream Prefetching in a Shared Cache" by Gill. For a more detailed description, see "Optimal Multistream Sequential Prefetching in a Shared Cache" by Gill, et al.

6.4.3 Intelligent Write Caching

Another cache algorithm, which is referred to as Intelligent Write Caching (IWC), was implemented in the DS8000 series. IWC improves performance through better write cache management and a better destaging order of writes. This algorithm is a combination of CLOCK, a predominantly read cache algorithm, and CSCAN, an efficient write cache algorithm. Out of this combination, IBM produced a powerful and widely applicable write cache algorithm.

The CLOCK algorithm uses temporal ordering. It keeps a circular list of pages in memory, with the clock hand that points to the oldest page in the list. When a page must be inserted in the cache, then an R (recency) bit is inspected at the clock hand's location. If R is zero, the new page is put in place of the page the clock hand points to and R is set to 1. Otherwise, the R bit is cleared and set to zero. Then the clock hand moves one step clockwise forward and the process is repeated until a page is replaced.

The CSCAN algorithm uses *spatial* ordering. The CSCAN algorithm is the circular variation of the SCAN algorithm. The SCAN algorithm tries to minimize the disk head movement when the disk head services read and write requests. It maintains a sorted list of pending requests with the position on the drive of the request. Requests are processed in the current direction of the disk head until it reaches the edge of the disk. At that point, the direction changes. In the CSCAN algorithm, the requests are always served in the same direction. After the head arrives at the outer edge of the disk, it returns to the beginning of the disk and services the new requests in this one direction only. This process results in more equal performance for all head positions.

The basic idea of IWC is to maintain a sorted list of write groups, as in the CSCAN algorithm. The smallest and the highest write groups are joined, forming a circular queue. The new idea is to maintain a recency bit for each write group, as in the CLOCK algorithm. A write group is always inserted in its correct sorted position and the recency bit is set to zero at the beginning. When a write hit occurs, the recency bit is set to one. The destage operation proceeds, where a destage pointer is maintained that scans the circular list and looks for destage victims. Now this algorithm allows destaging of only write groups whose recency bit is zero. The write groups with a recency bit of one are skipped and the recent bit is then turned off and reset to zero. This process gives an extra life to those write groups that were hit since the last time the destage pointer visited them. The concept of how this mechanism is illustrated is shown in Figure 6-6 on page 165.

In the DS8000 implementation, an IWC list is maintained for each rank. The dynamically adapted size of each IWC list is based on workload intensity on each rank. The rate of destage is proportional to the portion of NVS that is occupied by an IWC list. The NVS is shared across all ranks in a cluster. Furthermore, destages are smoothed out so that write bursts are not translated into destage bursts.

Another enhancement to IWC is an update to the cache algorithm that increases the residency time of data in NVS. This improvement focuses on maximizing throughput with good average response time.

In summary, IWC has better or comparable peak throughput to the best of CSCAN and CLOCK across a wide variety of write cache sizes and workload configurations. In addition, even at lower throughputs, IWC has lower average response times than CSCAN and CLOCK.

6.4.4 List Prefetch Optimizer

The Adaptive List Prefetch (ALP) algorithm, or List Prefetch Optimizer (LPO), enables prefetch of a list of non-sequential tracks, providing improved performance for DB2 workloads. ALP, when working together with AMP, dramatically improves the performance of record identifier (RID) list scans. These features are an essential part of Data Warehouse and Business Intelligence workloads, and help analytics applications benefit from being run on DS8880.

6.5 Performance considerations for logical configurations

To determine the optimal DS8880 layout, define the I/O performance requirements of the servers and applications up front because they play a large part in dictating the physical and logical configuration of the disk system. Before the disk system is designed, the disk space requirements of the application must be understood.

6.5.1 Workload characteristics

The answers to "How many host connections do I need?" and "How much cache do I need?" always depend on the workload requirements, such as how many I/Os per second for each server, and the I/Os per second for each gigabyte of storage.

The following information must be considered for a detailed modeling:

► Number of I/Os per second
► I/O density
► Megabytes per second
► Relative percentage of reads and writes
► Random or sequential access characteristics
► Cache-hit ratio
► Response time

6.5.2 Data placement in the DS8000

After you determine the disk subsystem throughput, disk space, and the number of disks that are required by your hosts and applications, determine the data placement.

To optimize the DS8880 resource utilization, use the following guidelines:

► Balance the ranks and extent pools between the two DS8880 internal servers to support the corresponding workloads on them.

► Spread the logical volume workload across the DS8880 internal servers by allocating the volumes equally on rank groups 0 and 1.

► Use as many disks as possible. Avoid idle disks, even if all storage capacity is not to be used initially.

► Distribute capacity and workload across device adapter pairs.

► Use multi-rank extent pools.

► Stripe your logical volume across several ranks (the default for multi-rank extent pools).

► Consider placing specific database objects (such as logs) on separate ranks.

► For an application, use volumes from even-numbered and odd-numbered extent pools. Even-numbered pools are managed by server 0, and odd-numbered pools are managed by server 1.

► For large, performance-sensitive applications, consider the use of two dedicated extent pools. One extent pool is managed by server 0, and the other extent pool is managed by server 1.

► Consider mixed extent pools with multiple tiers that use flash drives as the highest tier and that are managed by IBM Easy Tier.

In a typical DS8880 configuration with equally distributed workloads on two servers, the two extent pools (Extpool 0 and Extpool 1) are created, each with half of the ranks, as shown in Figure 6-6. The ranks in each extent pool are spread equally on each DA pair.
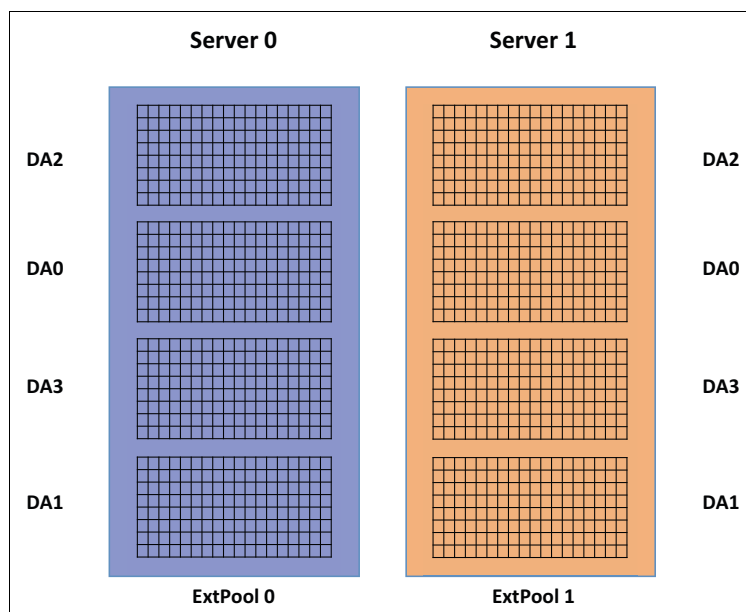
Figure 6-6   Ranks in a multi-rank extent pool configuration and balanced across the DS8000 servers

All disk arrays in a storage system have roughly equivalent utilization. Arrays that are used more than other arrays can reduce overall performance. IBM Easy Tier, with storage pool striping or host-level striping of logical volumes, can improve performance by distributing workload across multiple arrays. Easy Tier auto-rebalancing provides further benefits by automatically migrating data from overused extents to less-used areas of the extent pool.

Data striping

For optimal performance, spread your data across as many hardware resources as possible. RAID 10, RAID 6, or RAID 5 already spreads the data across the drives of an array, but this configuration is not always enough. The following approaches can be used to spread your data across even more disk drives:

► Storage pool striping that is combined with single-tier or multitier extent pools
► Striping at the host level

Easy Tier auto-rebalancing

With multitier extent pools, Easy Tier can automatically migrate the extents with the highest workload to faster storage tiers, and also to faster storage within a tier, as workloads dictate. *Auto-rebalance* migrates extents to achieve the best workload distribution within the pools and reduce hotspots. Together with intra-tier rebalancing, Easy Tier provides the optimum performance from each tier, while also reducing performance skew within a storage tier. Furthermore, auto-rebalance automatically populates new ranks that are added to the pool. Auto-rebalance can be enabled for hybrid and homogeneous extent pools.

Important: A preferred practice is to use IBM Easy Tier to balance workload across all ranks even when only a single tier of disk is installed in an extent pool. Use the options that are shown in Figure 6-7 to auto-rebalance all pools.
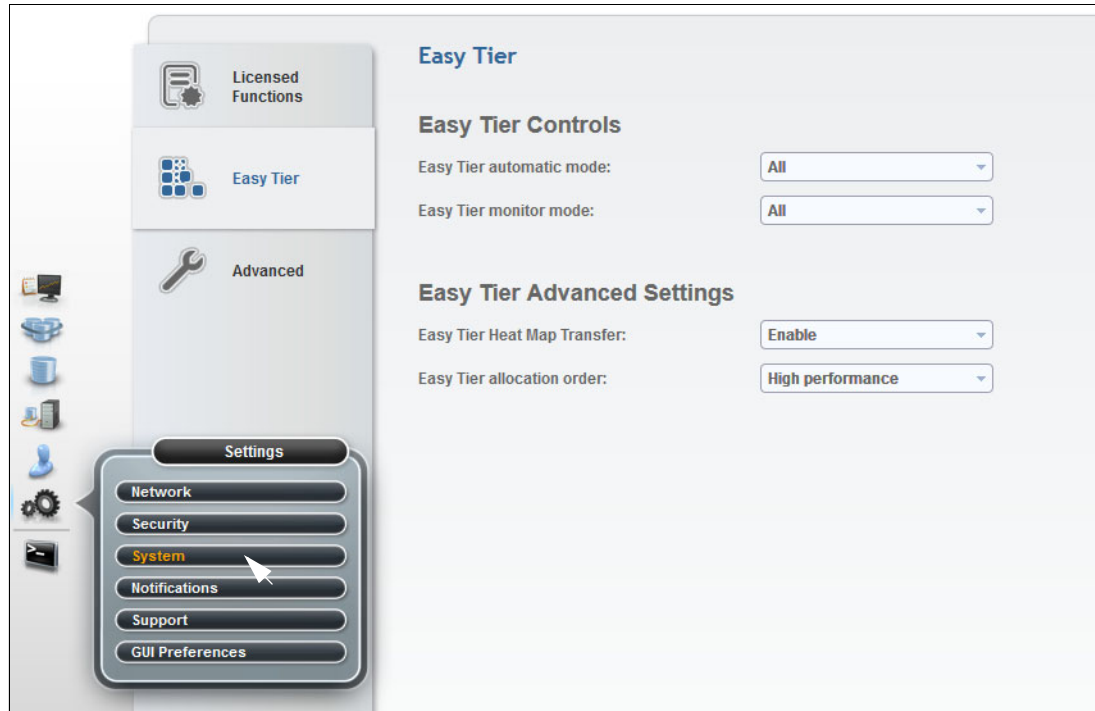
Figure 6-7   Easy Tier options to auto-balance all extent pools

Storage pool striping: Extent rotation

Storage pool striping is a technique for spreading data across several disk arrays. The I/O capability of many disk drives can be used in parallel to access data on the logical volume.

The easiest way to stripe is to create extent pools with more than one rank and use storage pool striping when you allocate a new volume. This striping method does not require any operating system support.

The number of random I/Os that can be run for a standard workload on a rank is described in 6.3, "Performance considerations for disk drives" on page 156. If a volume is on only one rank, the I/O capability of this rank also applies to the volume. However, if this volume is striped across several ranks, the I/O rate to this volume can be much higher.

The total number of I/Os that can be run on a set of ranks does not change with storage pool striping.

Important: Use storage pool striping and extent pools with a minimum of 4 - 8 ranks to avoid hot spots on the disk drives. Consider combining this configuration with Easy Tier auto-rebalancing.

Figure 6-8 shows an example of storage pool striping, considering large (1-GB) extents.

Rotate Extents

Volume 1 = 10GB
Volume 2 = 22GB
Volume 3 = 7GB
Volume 4 = 13GB
Volume 5 = 12GB
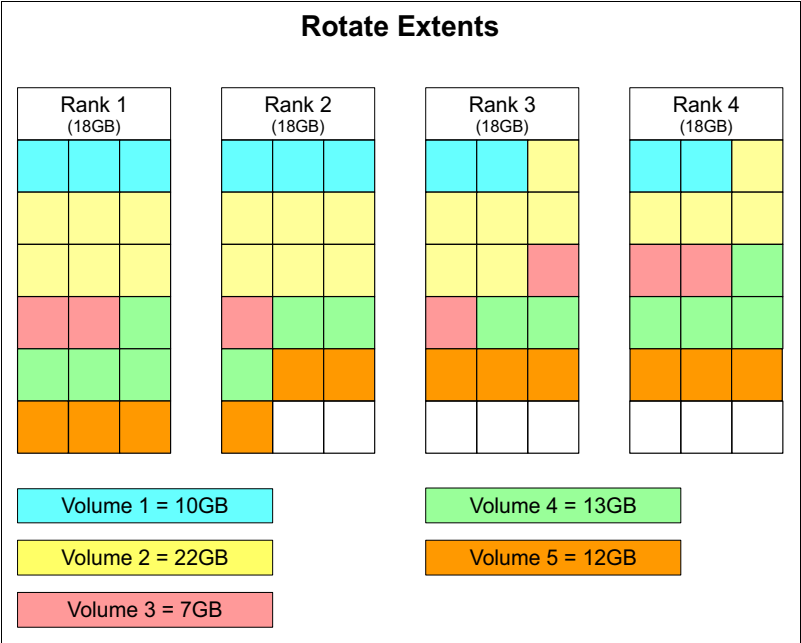
Figure 6-8   Storage pool striping

A sample configuration is shown in Figure 6-9. The ranks are attached to DS8880 server 0 and server 1 in a half-and-half configuration, which might be used with a mixed count key data (CKD) and fixed-block (FB) workload. Ranks on separate device adapters are combined in a multi-rank extent pool, and volumes are striped across all available ranks in each pool. In all-CKD or all-FB systems, consider the use of a single extent pool for each server.
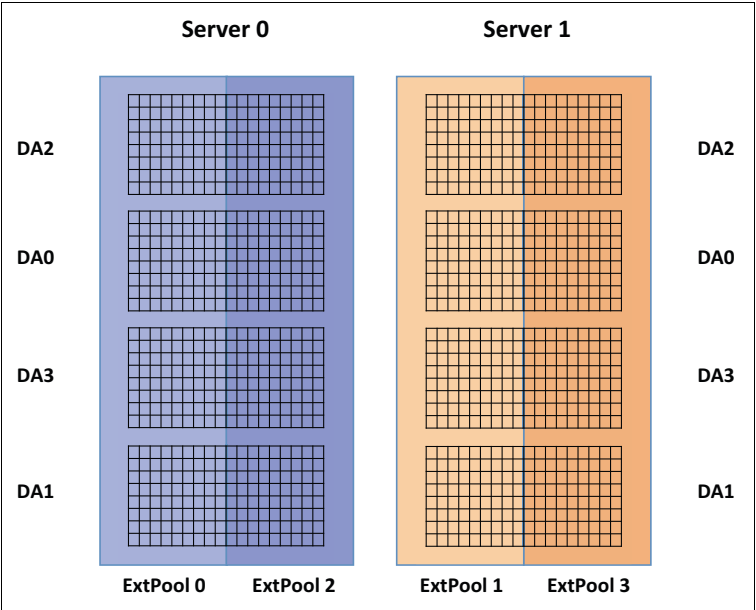
Figure 6-9   Balanced extent pool configuration

Tip: When using small extents (16 MB/21-cylinder), the load distribution across ranks becomes even more equal. Certain bursts of load that otherwise would go onto just a few ranks can be handled better by using small extents to distribute them further.

Striping at the host level

Many operating systems include the option to stripe data across several (logical) volumes. An example is the AIX Logical Volume Manager (LVM).

LVM striping is a technique for spreading the data in a logical volume across several disk drives so that the I/O capacity of the disk drives can be used in parallel to access data on the logical volume. The primary objective of striping is high-performance reading and writing of large sequential files, but benefits also exist for random access. Other examples for applications that stripe data across the volumes include the IBM SAN Volume Controller and IBM System Storage N series Gateway.

If you use a Logical Volume Manager (such as LVM on AIX) on your host, you can create a host logical volume from several DS8000 series logical volumes or logical unit numbers (LUNs). You can select LUNs from different DS8880 servers and device adapter pairs, as shown in Figure 6-10. By striping your host logical volume across the LUNs, the best performance for this LVM volume is realized.
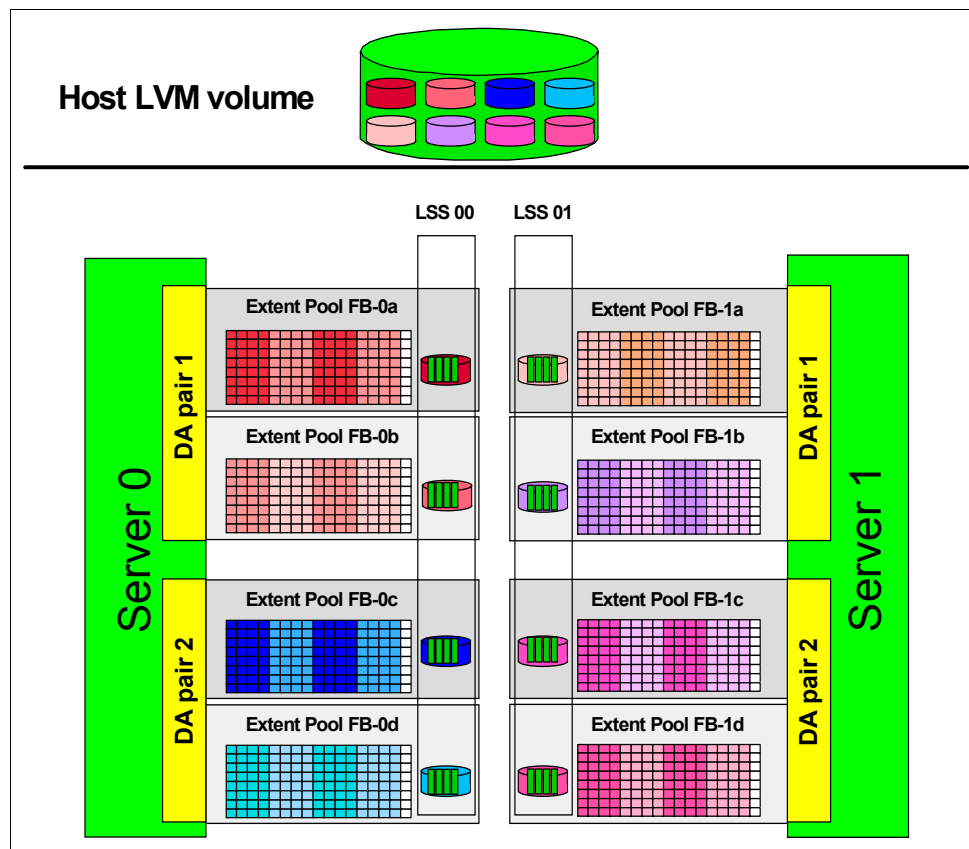
Figure 6-10   Optimal placement of data

Figure 6-10 shows an optimal distribution of eight logical volumes within a DS8880. You might have more extent pools and ranks, but when you want to distribute your data for optimal performance, ensure that you spread it across the two servers, across different device adapter pairs, and across several ranks.

The striping on the host level can also work with storage pool striping to help you spread the workloads on more ranks and disks. In addition, LUNs that were created on each extent pool offer another alternative balanced method to evenly spread data across the DS8880 without the use of storage pool striping, as shown on the left side of Figure 6-11.

If you use multi-rank extent pools and you do not use storage pool striping nor Easy Tier auto-rebalance, you must be careful where to put your data, or you can easily unbalance your system (as shown on the right side of Figure 6-11).
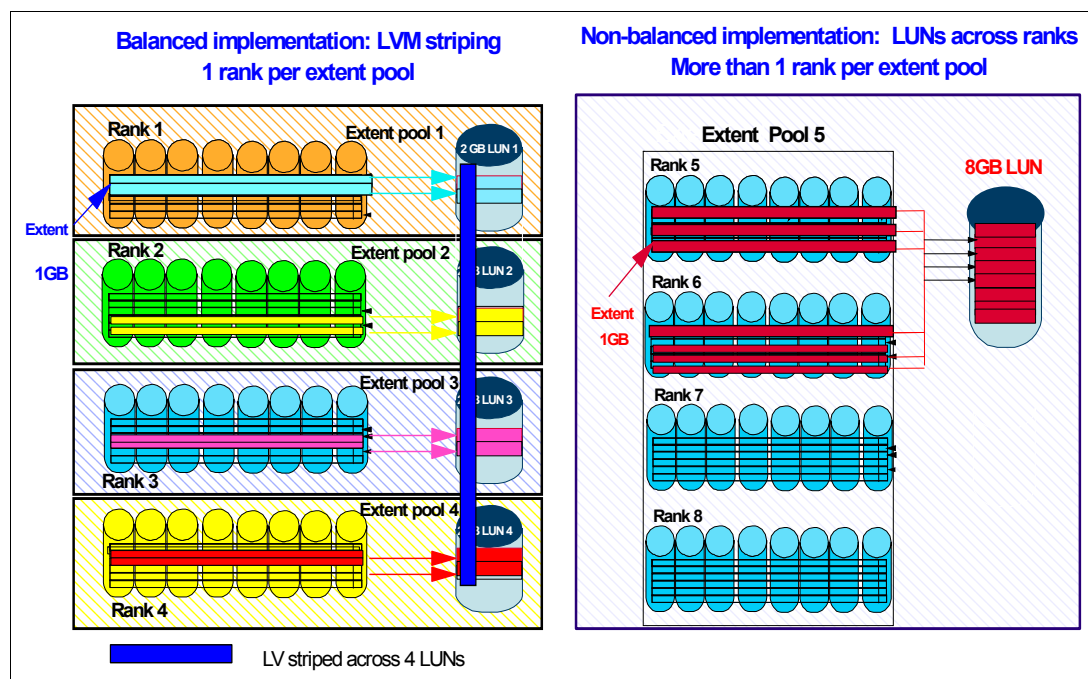
Figure 6-11   Spreading data across ranks (without the use of Easy Tier auto-rebalance)

Each striped logical volume that is created by the host's Logical Volume Manager has a stripe size that specifies the fixed amount of data that is stored on each DS8000 LUN.

The stripe size must be large enough to keep sequential data relatively close together, but not too large to keep the data on a single array.

Define stripe sizes by using your host's Logical Volume Manager in the range of 4 MB - 64 MB. Select a stripe size close to 4 MB if you have many applications that share the arrays, and a larger size when you have few servers or applications that share the arrays.

Combining extent pool striping and Logical Volume Manager striping

Striping by an LVM is performed on a stripe size in the MB range (about 64 MB). Extent pool striping is at a 1-GiB extent size for large-extent pools (or small-extent pools, 16 MiB). Both methods can be combined, but avoid working with two strip granularities on top of one another that are almost the same. LVM striping can stripe across extent pools and use volumes from extent pools that are attached to server 0 and server 1 of the DS8000 series. If you already use LVM physical partition (PP) wide striping, you might want to continue to use that striping.

Important: Striping at the host layer contributes to an equal distribution of I/Os to the disk drives to reduce hot spots. However, if you are using tiered extent pools with solid-state flash drives, IBM Easy Tier can work best if hot extents can be moved to flash drives. As an alternative, consider pools with small extents for a more even load distribution.

6.6 I/O Priority Manager

It is common practice to have large extent pools and stripe data across all disks. However, when production workload and, for example, test systems, share the physical disk drives, potentially the test system can negatively affect the production performance.

The DS8000 series I/O Priority Manager (IOPM) feature is included in the Base Function license group for the DS8880. It enables more effective storage consolidation and performance management, and the ability to align QoS levels to disparate workloads in the system. These workloads might compete for the same shared and possibly constrained storage resources.

I/O Priority Manager constantly monitors system resources to help applications meet their performance targets automatically, without operator intervention. The DS8880 storage hardware resources are monitored by the IOPM for possible contention within the RAID ranks and device adapters.

I/O Priority Manager uses QoS to assign priorities for different volumes, and applies network QoS principles to storage by using an algorithm called Token Bucket Throttling for traffic control. IOPM is designed to understand the load on the system and modify priorities by using dynamic workload control.

The I/O of a less important workload is slowed down to give a higher-priority workload a higher share of the resources.

Important: If you separated production and non-production data by using different extent pools and different device adapters, you do not need I/O Priority Manager.

Figure 6-12 shows a three-step example of how I/O Priority Manager uses dynamic workload control.
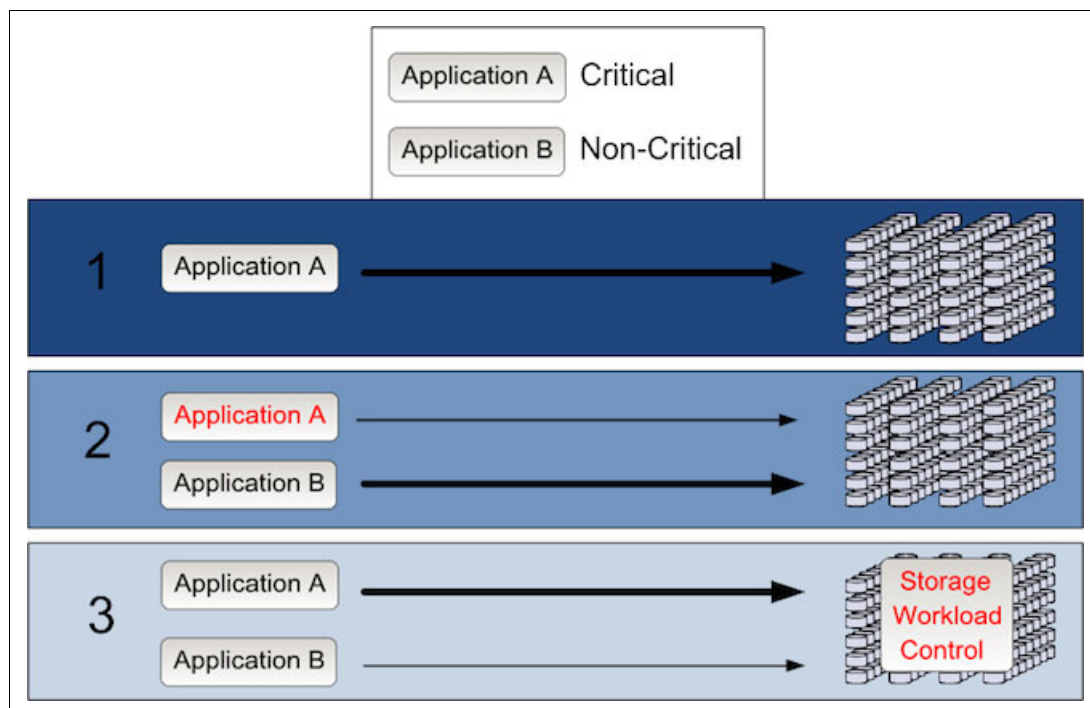
Figure 6-12   Automatic control of disruptive workload by I/O Priority Manager

In step 1, critical application A works normally. In step 2, a non-critical application B starts, causing performance degradation for application A. In step 3, I/O Priority Manager detects the QoS impact on critical application A automatically and dynamically restores the performance for application A.

6.6.1 Performance policies for open systems

When I/O Priority Manager is enabled, each volume is assigned to a performance group when the volume is created. Each performance group has a QoS target. This QoS target is used to determine whether a volume is experiencing appropriate response times.

A performance group associates the I/O operations of a logical volume with a performance policy that sets the priority of a volume relative to other volumes. All volumes fall into one of the performance policies.

For open systems, I/O Priority Manager includes four defined performance policies: Default (unmanaged), high priority, medium priority, and low priority. I/O Priority Manager includes 16 performance groups: Five performance groups each for the high, medium, and low performance policies, and one performance group for the default performance policy.

The following performance policies are available:

► Default performance policy

The default performance policy does not have a QoS target that is associated with it. I/Os to volumes that are assigned to the default performance policy are never delayed by I/O Priority Manager.

► High priority performance policy

The high priority performance policy has a QoS target of 70. I/Os from volumes that are associated with the high performance policy attempt to stay under approximately 1.5 times the optimal response of the rank. I/Os in the high performance policy are never delayed.

► Medium priority performance policy

The medium priority performance policy has a QoS target of 40. I/Os from volumes with the medium performance policy attempt to stay under 2.5 times the optimal response time of the rank.

► Low performance policy

Volumes with a low performance policy have no QoS target and have no goal for response times. If no bottleneck occurs for a shared resource, the low priority workload is not restricted. However, if a higher priority workload does not achieve its goal, the I/O of a low priority workload is slowed down first by delaying the response to the host. This delay is increased until the higher-priority I/O meets its goal. The maximum delay that is added is 200 ms.

6.6.2 Performance policies for z Systems

z Systems has 14 performance groups: Three performance groups for high performance policies, four performance groups for medium performance policies, six performance groups for low performance policies, and one performance group for the default performance policy.

With z Systems, two operation modes are available for I/O Priority Manager: Without software support, or with software support.

Important: Only z/OS operating systems use the I/O Priority Manager with software support.

I/O Priority Manager count key data support

In a z Systems environment, I/O Priority Manager includes the following characteristics:

► A user assigns a performance policy to each CKD volume that applies in the absence of software support.

► z/OS can optionally specify parameters that determine the priority of each I/O operation and allow multiple workloads on a single CKD volume to have different priorities.

► I/O Priority Manager is supported on z/OS V1.11, V1.12, V1.13, and later.

► Without z/OS software support, on ranks in saturation, the volume's I/O is managed according to the performance policy of the volume's performance group.

► With z/OS software support, the following actions occur:

– A user assigns application priorities by using IBM Enterprise Workload Manager™ (EWLM).

– z/OS assigns an importance value to each I/O based on eWLM inputs.

– z/OS assigns an achievement value to each I/O based on the previous history of I/O response times for I/O with the same importance and based on eWLM expectations for response time.

– The importance and achievement value on I/O associates this I/O with a performance policy (independently of the volume's performance group/performance policy).

– On ranks in saturation, I/O is managed according to the I/O's performance policy.

If no bottleneck exists for a shared resource, a low priority workload is not restricted. However, if a higher priority workload does not achieve its goal, the I/O of the low priority workload is slowed down first, by delaying the response to the host. This delay is increased until the higher-priority I/O meets its goal. The maximum delay that is added is 200 ms.

> **For more information:** For more information about I/O Priority Manager, see *DS8000 I/O Priority Manager*, REDP-4760.

6.7 IBM Easy Tier

IBM Easy Tier on the DS8880 can enhance performance and balance workloads through the following capabilities:

► Automated hot spot management and data relocation.

► Support for thin provisioned and fully provisioned extent space-efficient (ESE) volumes.

► Automatic inter-tier and intra-tier rebalancing.

  In systems that contain flash storage only, or systems with a single drive type, intra-tier rebalancing monitors for hot spots and relocates extents to balance the workload.

► Support for HPFEs, including intra-tier rebalancing for heterogeneous flash storage pools.

► Manual volume rebalancing and volume migration.

► Rank depopulation.

► Extent pool merging.

► Directive data placement from applications.

  Support is provided for z Systems applications, such as DB2, to communicate application hints to the Easy Tier Application API. An application hint that contains performance requirements for optimal data set placement instructs Easy Tier to move the data set to the correct tier.

► On FlashCopy activities, adequately assigning workloads to source and target volumes for the best production optimization.

► Heat map transfer from Peer-to-Peer Remote Copy (PPRC) source to target.

  The eighth generation of Easy Tier supports Easy Tier Heat Map Transfer (HMT) in 3-site and 4-site Metro Global Mirror (MGM) environments. Used with GDPS or IBM Tivoli Storage Productivity Center for Replication, this capability leads to performance-optimized disaster recovery.

► The administrator can pause and resume Easy Tier monitoring (learning) at the extent pool and volume level, and can reset Easy Tier learning for pools and volumes. The administrator can now pause and resume migration at the pool level. It is also possible to prevent volumes from being assigned to the nearline tier.

► Easy Tier Application is an application-aware storage interface to help deploy storage more efficiently through enabling applications and middleware to direct more optimal placement of data. Easy Tier Application enables administrators to assign distinct application volumes to a particular tier in the Easy Tier pool. This capability provides a flexible option for clients who want certain applications to remain on a particular tier to meet performance and cost requirements. Easy Tier Application is available at no additional cost, and no separate license is required.

Full Disk Encryption support: All drive types in the DS8880 support Full Disk Encryption. Encryption usage is optional. With or without encryption, disk performance or Easy Tier functions are the same.

For more information about IBM Easy Tier, see *IBM DS8000 EasyTier*, REDP-4667, *DS8870 Easy Tier Application*, REDP-5014, and *IBM DS8870 Easy Tier Heat Map Transfer*, REDP-5015.

6.8 Host performance and sizing considerations

This section describes performance and sizing considerations for open systems. For z Systems-specific performance and sizing considerations, see *IBM DS8870 and IBM z Systems Synergy*, REDP-5186.

6.8.1 Performance and sizing considerations for open systems

The following sections describe topics that relate to open systems.

Determining the number of paths to a LUN

When you configure a DS8000 series for an open systems host, a decision must be made about the number of paths to a particular LUN because the multipathing software allows (and manages) multiple paths to a LUN. The following opposing factors must be considered when you are deciding on the number of paths to a LUN:

► Increasing the number of paths increases availability of the data, which protects against outages.

► Increasing the number of paths increases the amount of CPU that is used because the multipathing software must choose among all available paths each time that an I/O is issued.

A good compromise is 2 - 4 paths for each LUN. Consider eight paths if a high data rate is required.

Dynamic I/O load-balancing: Subsystem Device Driver

Subsystem Device Driver (SDD) is an IBM provided pseudo-device driver that is designed to support the multipath configuration capabilities in the DS8000. SDD runs on each host system, in cooperation with the native disk device driver.

The dynamic I/O load-balancing option (default) of SDD is suggested to achieve the best performance by using these functions:

► SDD automatically adjusts data routing for optimum performance. The multipath load balancing of the data flow prevents a single path from becoming overloaded, causing I/O congestion when many I/O operations are directed to common devices along the same I/O path.

► The path to use for an I/O operation is chosen by estimating the load on each adapter to which each path is attached. The load is a function of the number of I/O operations currently in process. If multiple paths include the same load, a path is chosen at random from those paths.

IBM SDD is available for most operating environments. On certain operating systems, SDD offers an installable package to work with their native multipathing software, also. For example, IBM Subsystem Device Driver Path Control Module (SDDPCM) is available for AIX, and IBM Subsystem Device Driver Device Specific Module (SDDDSM) is available for Microsoft Windows.

For more information about the multipathing software that might be required for various operating systems, see the IBM System Storage Interoperation Center (SSIC):

https://www.ibm.com/systems/support/storage/ssic/

SDD is covered in more detail in the following IBM publications:

► *IBM System Storage DS8000: Host Attachment and Interoperability*, SG24-8887
► *IBM System Storage DS8000 Host Systems Attachment Guide*, SC27-4210

Automatic Port Queues

The DS8880 Fibre Channel host adapters and the server HBAs support I/O queuing. The length of this queue is called the *queue depth*. Because several servers can and usually do communicate with few DS8880 ports, the queue depth of a storage host adapter needs to be larger than the queue depth on the server side. This parameter is also true for the DS8880, which supports 2,048 FC commands that are queued on a port. However, sometimes the port queue in the DS8880 host adapter can be flooded.

When the number of commands that are sent to the DS8880 port exceeds the maximum number of commands that the port can queue, the port discards these additional commands. This operation is a *normal* error recovery operation in the Fibre Channel Protocol to allow overprovisioning on the SAN. The normal recovery is a 30-second timeout for the server. The server retries the command until the *command retry* value is exceeded, at which point the operation fails. Command Timeout entries can be seen in the server logs.

Automatic Port Queues is a mechanism that the DS8880 uses to self-adjust the queues based on workload. This mechanism allows higher port queue oversubscription while it maintains a fair share for the servers and the accessed LUNs. When an I/O port queue fills up, the port goes into SCSI Queue Full mode, and accepts no additional commands to slow down the I/Os. By avoiding error recovery and the 30-second blocking SCSI Queue Full recovery interval, the overall performance is better with Automatic Port Queues.

Determining where to attach the host

The DS8000 series host adapters have no server affinity, but the device adapters and the ranks have server affinity. When you determine where to attach multiple paths from a single host system to I/O ports in the storage system, the following considerations apply:

► Spread the connections across host adapters in all of the I/O enclosures
► Spread the connections across port pairs in the host adapters

Figure 6-13 shows a z13 system with four FICON paths, which are connected through two SAN384B-2 switches to eight DS8880 host adapters. The connections are spread evenly across the I/O enclosures. The diagram also shows the DS8880 internal PCIe pathing between the internal servers and the I/O enclosures and to a high performance flash enclosure, and the 8 Gb Fibre Channel connections from the device adapters to the standard drive enclosures.
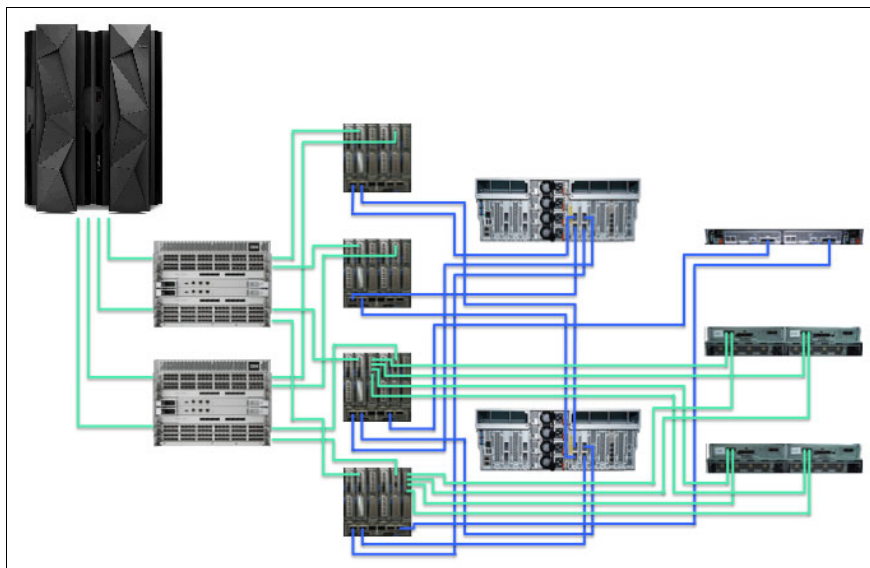
Figure 6-13   DS8880 multipathing

Options for four-port and eight-port 8 Gbps host adapters are available in the DS8880. Eight-port cards provide more connectivity, but not necessarily more total throughput because all the ports share a single PCIe connection to the I/O enclosure. Additionally, host ports are internally paired, driven by two-port Fibre Channel I/O controller modules. Four-port 16 Gbps adapters are also available.

For maximum throughput, consider the use of fewer ports for each host adapter and spread the workload across more host adapters. Where possible, avoid mixing FICON and FCP connections on a single adapter, or mixing host connections with PPRC connections.

Part 2

Planning and installation

This part of the book describes the installation planning process for the IBM DS8880.

This part contains the following chapters:

7

IBM DS8880 physical planning and installation

This chapter describes the various steps that are involved in the planning and installation of the IBM DS8880. It includes a reference listing of the information that is required for the setup and where to find detailed technical reference material.

This chapter covers the following topics:

- ► Considerations before the installation: Planning for growth
- ► Planning for the physical installation
- ► Network connectivity planning
- ► Remote Mirror and Copy connectivity
- ► Disk capacity considerations

For more information about the configuration and installation process, see the *IBM System Storage DS8880 Introduction and Planning Guide*, GC27-8525.

7.1 Considerations before the installation: Planning for growth

Start by developing and following a project plan to address the necessary topics for a successful implementation. Consider the following items for your installation plan checklist:

► Plan for growth to minimize disruption to operations.

> **Important:** Expansion frames can be placed only to the right (from the front) of the DS8880 base frame.

► Consider location suitability, floor loading, access constraints, elevators, and doorways.

► Analyze power requirements, such as redundancy and the use of an uninterruptible power supply (UPS).

► Examine environmental requirements, such as adequate cooling capacity.

► Full Disk Encryption (FDE) drives are a standard feature for the DS8880. If encryption activation is required, consider the location and connection needs for the external key servers, IBM Security Key Lifecycle Manager (SKLM), or Gemalto SafeNet KeySecure servers.

► Consider the integration of Lightweight Directory Access Protocol (LDAP) to allow a single user ID and password management. With Release 8.1, you can take advantage of the LDAP capability offered by the Copy Service Manager residing on the Hardware Management Console (HMC).

► Call home through an Internet Secure Sockets Layer (SSL) installation to provide a continued secure connection to the IBM Support center.

► Plan for disk types, such as flash cards, flash drives, and enterprise and nearline serial-attached Small Computer System Interface (SCSI) (SAS).

► Plan for logical configuration, Copy Services, and staff education. See Chapter 10, "Configuration flow" on page 245.

7.1.1 Client responsibilities for the installation

The DS8880 is specified as an IBM or IBM Business Partner installation and setup system. However, the following are several required planning and installation activities for which the client is responsible at a high level:

► Physical configuration planning. Your Storage Marketing Specialist can help you plan and select the DS8880 model physical configuration and features.

► Installation planning.

► Integration of LDAP. IBM can help in planning and implementation on client request.

► Installation of Assist On-site (AOS). IBM can help plan and implement on client request.

► Integration of IBM Spectrum™ Control and Simple Network Management Protocol (SNMP) into the client environment for monitoring of performance and configuration. IBM can provide services to set up and integrate these components.

► Configuration and integration of SKLM servers and DS8000 Encryption for extended data security. IBM provides services to set up and integrate these components. IBM Storage Appliance 2421 Model AP1 can be ordered either as a single isolated key server (Feature Code (FC) 1761) or as two isolated key servers (FC1761 and FC1762, ordered together). However, it is not the only option. A client might want to run two independent SKLMs instances that run on their own servers. Alternatively, clients who must conform to the Key

Management Interoperability Protocol (KMIP) can install the Gemalto SafeNet key servers.

► Logical configuration planning and application. *Logical configuration* refers to the creation of Redundant Array of Independent Disks (RAID) arrays and pools, and to the assignment of the configured capacity to servers. Application of the initial logical configuration and all subsequent modifications to the logical configuration also are client responsibilities. The logical configuration can be created, applied, and modified by using the data storage Management graphical user interface (DS GUI), DS command-line interface (DS CLI), or DS Open application programming interface (DS Open API).

IBM Services also can apply or modify your logical configuration, which is a fee-based service.

7.1.2 Participants

A project manager needs to coordinate the many tasks that are necessary for a successful installation. Installation requires close cooperation with the user community, IT support staff, and technical resources that are responsible for floor space, power, and cooling.

A storage administrator needs to also coordinate requirements from the user applications and systems to build a storage plan for the installation. This plan is needed to configure the storage after the initial hardware installation is complete.

The following people must be briefed and engaged in the planning process for the physical installation:

► Systems and storage administrators
► Installation planning engineer
► Building engineer for floor loading, air conditioning, and electrical considerations
► Security engineers for AOS, LDAP, key servers, and encryption
► Administrator and operator for monitoring and handling considerations
► IBM service support representative (SSR) or IBM Business Partner

7.1.3 Required information

A validation list to help the installation process must include the following items:

► Drawings that detail the DS8000 placement as specified and agreed upon with a building engineer, which ensures that the weight is within limits for the route to the final installation position.

► Approval to use elevators if the DS8880 weight and size are acceptable.

► Connectivity information, servers, storage area network (SAN), and mandatory local area network (LAN) connections.

► Agreement on the security structure of the installed DS8000 with all security engineers.

► Agreement on the detailed storage plan. Ensure that the configuration specialist has all of the information to configure all of the storage and set up the environment, as required.

► Activation codes for Base Functions, which are mandatory, and any optional feature activation codes.

7.2 Planning for the physical installation

This section describes the physical installation planning process and provides important tips and considerations.

7.2.1 Delivery and staging area

The shipping carrier is responsible for delivering and unloading the DS8880 as close to its final destination as possible. Inform the carrier of the weight and size of the packages to deliver. Also, inspect the site and the areas through which the packages will be moved (for example, hallways, floor protection, elevator size, and loading). The 46U tall racks are shipped as 40U, with the top-hat lowered for ease of shipping. The hardware positioned between 40U and 46U is shipped in a separate box and must be physically installed before formal installation is attempted.

Table 7-1 lists the final packaged dimensions and maximum packaged weight of the DS8880 storage unit ship group.

Table 7-1   Packaged dimensions and weight for DS8880 models

Shipping container	Packaged dimensions (in centimeters and inches)	Maximum packaged weight (in kilograms and pounds)
DS8884 Base Frame (Model 984)	Height 2.08 m (81.9 in.) Width 0.95 m (37.4 in.) Depth 1.50 m (59.1 in.)	1086 kg (2390 lb)
DS8886 Base Frame (Model 985 1-Phase)	Height 2.08 m (81.9 in.) Width 0.95 m (37.4 in.) Depth 1.50 m (59.1 in.)	1250.5 kg (2757 lb)
DS8886 Base Frame (Model 986 3-Phase)	Height 2.08 m (81.9 in.) Width 0.95 m (37.4 in.) Depth 1.50 m (59.1 in.)	1187.5 kg (2618 lb)
DS8888 base frame (Model 988)	Height 2.08 m (81.9 in.) Width 0.95 m (37.4 in.) Depth 1.50 m (59.1 in.)	1171 kg (2577 lb)
DS8884 Expansion frame (Model 84E)	Height 2.08 m (81.9 in.) Width 0.95 m (37.4 in.) Depth 1.50 m (59.1 in.)	First expansion frame: 1128 kg (2487 lb) Second expansion frame: 940 kg (2068 lb)
DS8886 Expansion Frame (Model 85E)	Height 2.08 m (81.9 in.) Width 0.95 m (37.4 in.) Depth 1.50 m (59.1 in.)	First expansion frame: 1205 kg (2657 lb) Second, and third expansion frames: 1021 kg (2246 lb) Fourth expansion frame: 902 kg (1984 lb)

Shipping container	Packaged dimensions (in centimeters and inches)	Maximum packaged weight (in kilograms and pounds)
DS8886 Expansion Frame (Model 86E)	Height 2.08 m (81.9 in.) Width 0.95 m (37.4 in.) Depth 1.50 m (59.1 in.)	First expansion frame: 1205 kg (2657 lb) Second through fourth expansion frames: 1026 kg (2256 lb)
DS8888 Expansion Frame (Model 88E)	Height 2.08 m (81.9 in.) Width 0.95 m (37.4 in.) Depth 1.50 m (59.1 in.)	999 kg (2198 lb)

Important: A fully configured model in the packaging can weigh over 1,200 kg (2,640 lbs). The use of fewer than three persons to move it can result in injury.

By using the shipping weight reduction option, you can receive delivery of a DS8880 model in multiple shipments that do not exceed 909 kg (2,000 lb) each.

For more information about the Shipping Weight Reduction option, see Chapter 9, "IBM DS8880 features and licensed functions" on page 219.

7.2.2 Floor type and loading

The DS8880 can be installed on a raised or nonraised floor. The total weight and space requirements of the storage unit depend on the configuration features that you ordered. You might consider calculating the weight of the unit and the expansion frame (if ordered) in their maximum capacity to allow for the addition of new features.

Table 7-2 lists the weights of the various DS8880 models.

Table 7-2 DS8880 weights

Model	Maximum weight
DS8880 Model 984	976 kg (2,147 lb)
DS8880 Model 985	1,144 kg (2,517 lb)
DS8880 Model 986	1,099 kg (2,418 lb)
DS8880 Model 988	1,080 kg (2376 lb)
DS8880 Model 84E (first expansion frame)	1038 kg (2,284 lb)
DS8880 Model 84E (second expansion frame)	849 kg (1,868 lb)
DS8880 Model 85E (first expansion frame)	1,115 kg (2,453 lb)
DS8880 Model 85E (second and third expansion frame)	931 kg (2,048 lb)
DS8880 Model 85E (fourth expansion frame)	812 kg (1,786 lb)

Model	Maximum weight
DS8880 Model 86E (first expansion frame)	952 kg (2,094 lb)
DS8880 Model 86E (second through fourth expansion frame)	934 kg (2,055 lb)
DS8880 Model 88E	908 kg (2,198 lb)

Important: You must check with the building engineer or other appropriate personnel to ensure that the floor loading is correctly considered.

Figure 7-1 for DS8880 shows the location of the cable cutouts. You can use the following measurements when you cut the floor tile:

- ► Width: 40.0 cm (15.7 in.)
- ► Depth: 14.5 cm (5.7 in.)
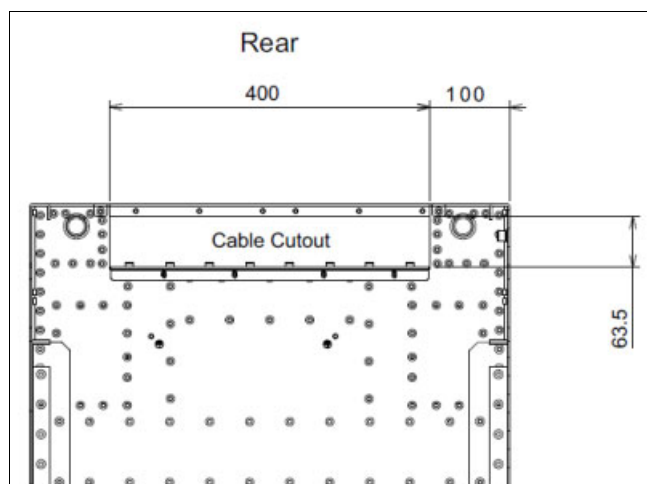
Figure 7-1   Floor tile cable cutout for DS8880

For more information about floor loading and weight distribution, see the *IBM System Storage DS8880 Introduction and Planning Guide*, GC27-8525.

7.2.3 Overhead cabling features

The overhead cabling (top exit) feature, as shown in Figure 7-2 on page 185, is available for DS8880 as an alternative to the standard rear cable exit. Verify whether you ordered the top exit feature before the tiles for a raised floor are cut.

This feature requires the following items:

- ► Feature Code (FC) 1400 Top exit bracket for overhead cabling
- ► FC1101 Safety-approved fiberglass ladder

For more information, see the *IBM System Storage DS8880 Introduction and Planning Guide*, GC27-8525.

Figure 7-2 shows the overhead cabling (top exit) feature.
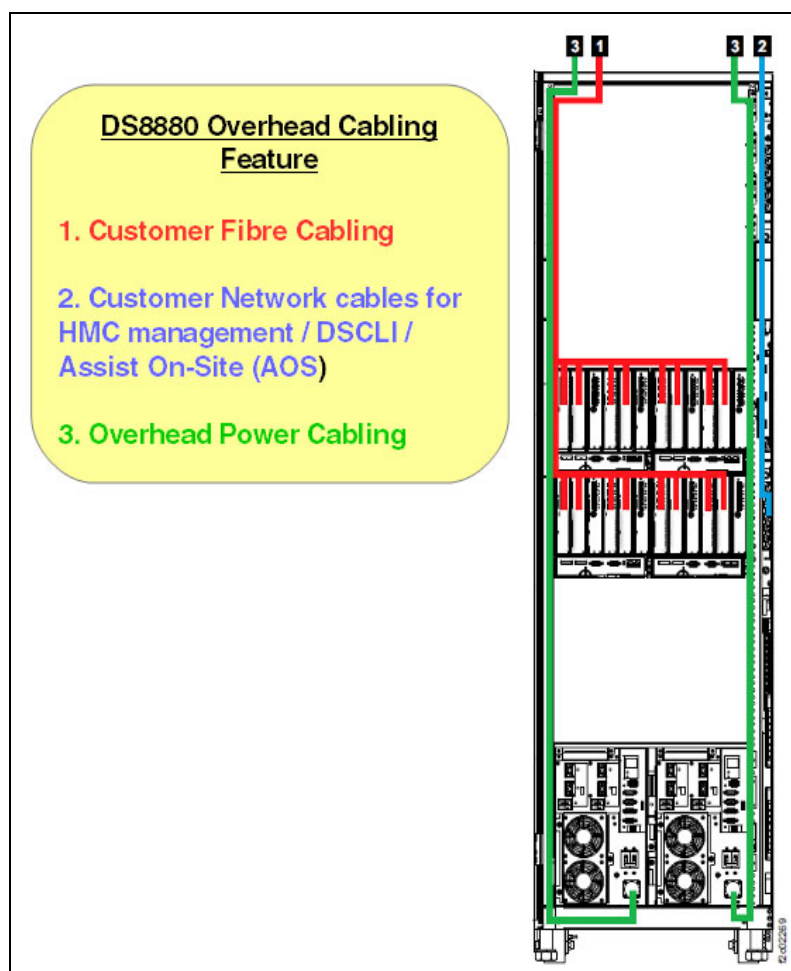
Figure 7-2   Overhead cabling for DS8880

7.2.4 Room space and service clearance

The total amount of space that is needed by the storage units can be calculated by using the dimensions that are shown in Table 7-3.

Table 7-3   DS8880 dimensions

Dimensions with casters and covers	Models 980/982/984	Models 981/985/986/988 without / with top extension
Height	193 cm (76 in.)	193 cm (76 in.) / 220 cm (86.5 in.)
Width	64 cm (25 in.)	64 cm (25 in.)
Depth	144 cm (56.5 in.)	144 cm (56.5 in.)

The storage unit location area also covers the service clearance that is needed by the IBM SSR when the front and rear of the storage unit are accessed. You can use the following minimum service clearances. Verify your configuration and the maximum configuration for your needs, keeping in mind that the DS8880 has a maximum of four expansion frames (for a total of five frames).

An example of the dimensions for a DS8880 with two expansion frames is shown in Figure 7-3. The following clearances are needed:

► For the front of the unit, allow a minimum of 121.9 cm (48 in.).
► For the rear of the unit, allow a minimum of 76.2 cm (30 in.).
► For the sides of the unit, allow a minimum of 12.7 cm (5 in.).

Important: When you install only the base frame, you also need a clearance of at least 61 cm (24.0 in.) on the right side (facing the system) to allow an operator to unfold and use the HMC keyboard and monitor.

Figure 7-3   DS8880 with three frames service clearance requirements

7.2.5 Power requirements and operating environment

Consider the following basic items when you plan for the DS8880 power requirements:

► Power connectors
► Input voltage
► Power consumption and environment
► Power control features
► Extended power line disturbance (ePLD) feature

Power connectors

Each DS8880 base and expansion frame features redundant power supply systems. Attach the two power cords to each frame to separate AC power distribution systems. For more information about power connectors and power cords, see the *IBM System Storage DS8880 Introduction and Planning Guide*, GC27-8525.
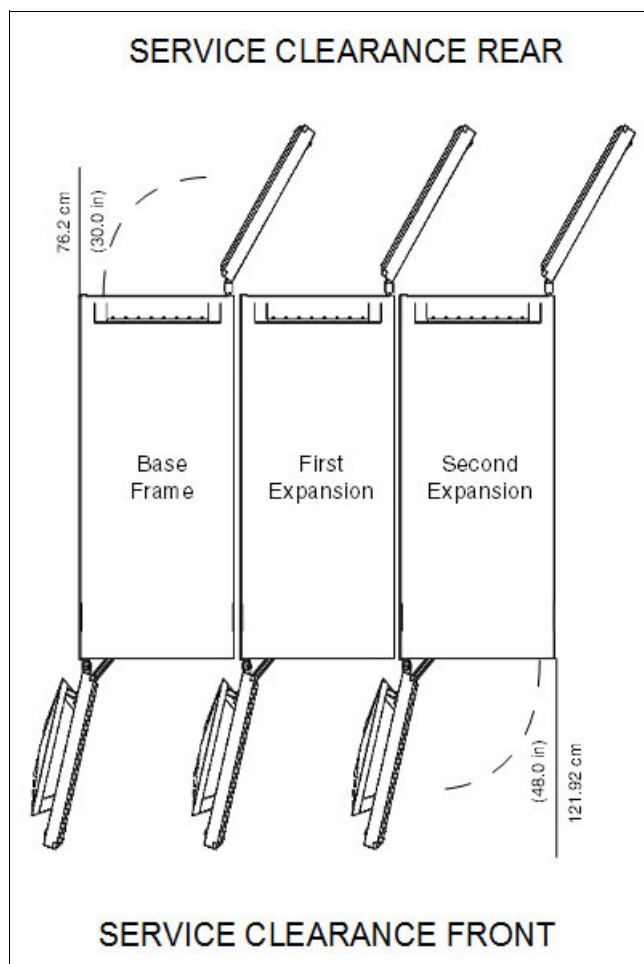
Input voltage

When you plan for the power requirements of the storage system, consider the input voltage requirements. Table 7-4 shows the DS8880 input voltages and frequencies.

Table 7-4   DS8880 input voltages and frequencies

Characteristic	Voltage (single-phase)
Nominal input voltage	200, 208, 220, or 240 RMS Vac
Minimum input voltage	180 RMS Vac
Maximum input voltage	256 RMS Vac
Customer wall breaker rating 1-phase	50 - 63 Amps[a]
Customer wall breaker rating 3-phase	30 - 60 Amps (delta), 20 - 32 Amps (wye)
Steady-state input frequency	50 ± 3 or 60 ± 3.0 Hz
PLD input frequencies (<10 seconds)	50 ± 3 or 60 ± 3.0 Hz

a. Can vary by region

Power consumption

Table 7-5 lists the power consumption specifications of the DS8880. The power estimates in this table are conservative and assume a high transaction rate workload.

Table 7-5   Power consumption and environmental information (fully equipped frames)

Measurement	Unit of measure	Base frame Model	Expansion frame Model
Peak electric power	kilovolt amperes (kVA)	984: 5.4 (single-phase) 985: 7.4 (single-phase) 986: 6.7 (three-phase) 982: 8.1 (three-phase)	84E: 5.3 (single-phase) 85E: 6.6 (single-phase) 86E: 6.4 (three-phase) 98F: 4.4 (three-phase)
Thermal load	British thermal units (BTU) per hour	984: 18425 (single-phase) 985: 25249 (single-phase) 986: 22860 (three-phase) 982: 27637 (three-phase)	84E: 18084 (single-phase) 85E: 22519 (single-phase) 86E: 21837 (three-phase) 98F: 15013 (three-phase)
Capacity of exhaust	cubic meters per min. (cubic feet per min., or CFM)	44.2 (1500)	51.8 (1800)

The values represent data that was obtained from the following configured systems:

► Standard base frames that contain 15 disk drive sets (16 drives for each disk drive set, 15 disk drive sets x 16 = 240 disk drives) and Fibre Channel adapters.

► All-flash configurations that contain eight sets of fully configured high-performance storage enclosures and 16 Fibre Channel adapters.

► Expansion models that contain 21 drive sets for each storage enclosure (21 drive sets x 16 = 336 drives) and Fibre Channel adapters.

DS8880 cooling

Air circulation for the DS8880 is provided by the various fans that are installed throughout the frame. All of the fans in the DS8880 direct air flow from the front of the frame to the rear of the frame. No air exhausts out of the top of the frame. The use of a directional air flow in this manner allows cool aisles to the front and hot aisles to the rear of the systems, as shown in Figure 7-4.
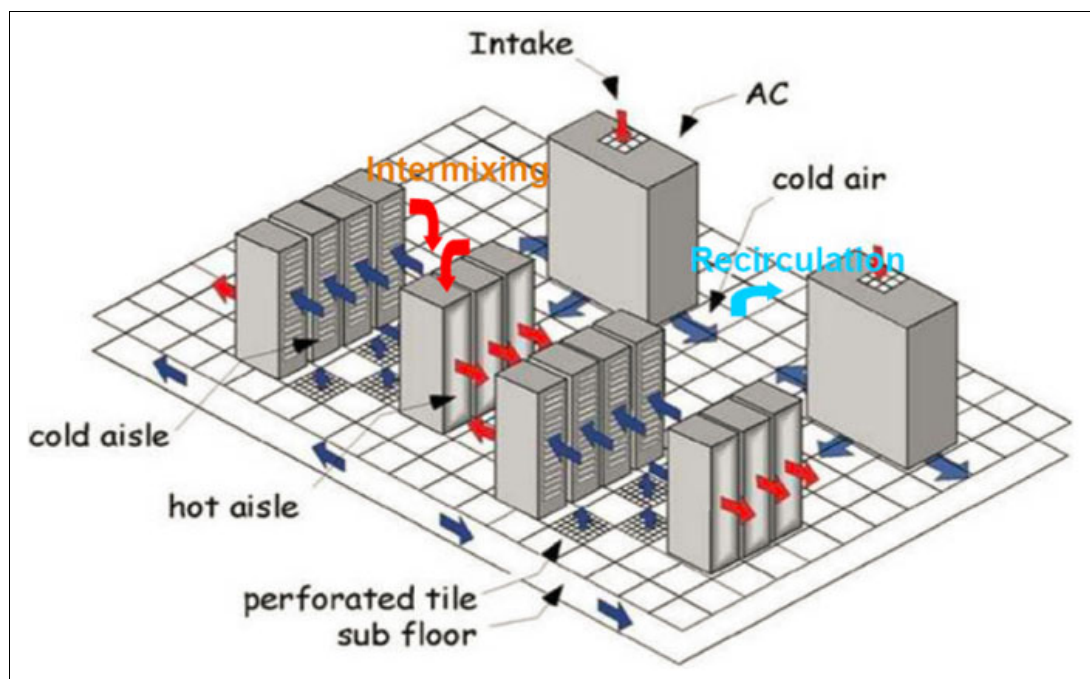
Figure 7-4   DS8880 air flow: Hot aisle/cold aisle approach

The operating temperature for the DS8880 is 16 - 32 °C (60 - 90 °F) at relative humidity limits of 20% - 80% and optimum at 45%.

> **Important:** The following factors must be considered when the DS8880 is installed:
>
> ► Ensure that the air circulation for the DS8880 base frame and expansion frames is maintained free from obstruction to keep the unit operating in the specified temperature range.
>
> ► For safety reasons, do not store anything on top of the DS8880.

Power control features

The DS8880 has remote power control features that are used to control the power of the storage system through the HMC. For more information about power control features, see the *IBM System Storage DS8880 Introduction and Planning Guide*, GC27-8525.

Extended power line disturbance feature

The ePLD feature extends the uptime duration to 40 seconds. The system operates on battery before it starts a controlled shutdown if power is lost to a DS8880 frame. Without the ePLD feature, the duration is 4 seconds of operating on battery before the controlled shutdown. No additional physical connection planning is needed for the client with or without the ePLD feature.

7.2.6 Host interface and cables

The DS8886 can contain a maximum of 16 host adapters (HAs) in the base frame, and a maximum of eight HAs can be configured in the DS8884. The DS8886 allows an additional 16 HAs to be installed in the first expansion frame. The DS8884 can contain an additional eight HAs. With sixteen 8-port HAs, the maximum number is 128 HA ports. With sixteen 4-port HAs, the maximum number is 64 HA ports. For a full breakdown of the available ports on a DS8886, see Table 2-1 on page 46 in Chapter 2, "IBM DS8880 hardware components and architecture" on page 23.

Table 7-6 shows the minimum and maximum numbers of host adapters that are supported by the DS8880.

Table 7-6   Minimum and maximum host adapters

Storage system type	Storage system configuration	Minimum number of host adapter features for the base frame	Maximum number of host adapter features for the storage system
6-core	Base frame plus two expansion frames	2	16
8-core	Base frame	2	16
16-core or 24-core	Base frame plus 1 - 4 expansion frames	2	32
48-core	Base frame plus one expansion frame	2	32

Fibre Channel and Fibre Channel connection

Each host adapter port supports Fibre Channel Protocol (FCP) or Fibre Channel connection (FICON). However, it cannot support both topologies simultaneously on the same port. Fabric components from various vendors, including IBM, QLogic, Brocade, and Cisco, are supported by both environments.

The Fibre Channel and FICON shortwave host adapter, when it is used with 50 micron multi-mode fiber cable, supports point-to-point distances. For information about cable limits, see Table 7-7.

Table 7-7   Cabling type and limits according to speed

Cable type	Distance limits		
	4 Gbps	8 Gbps	16 Gbps
OM2 (50 μm)	150 m (492 ft.)	N/A	35 m (114.8 ft.)
OM3 (50 μm)	380 m (1,246.7 ft.)	150 m (492 ft.)	100 m (328 ft.)
OM4 (50 μm)	400 m (1,312.3 ft.)	190 m (623.3 ft.)	125 m (410.1 ft.)

The Fibre Channel and FICON longwave host adapter, when it is used with 9-micron single-mode fiber cable, extends the point-to-point distance to 10 km (6.2 miles) on 8 Gbps link speed with four ports or eight ports.

A 31-meter (102 ft.) fiber optic cable or a 2-meter (6.5 ft.) jumper cable can be ordered for each Fibre Channel adapter port.

Table 7-8 lists the fiber optic cable features for the FCP/FICON adapters.

Table 7-8   FCP/FICON cable features

Feature code	Length	Characteristics	Compatible Fibre Channel host adapter features
1410	40 m (131 ft.)	50 micron OM3 or higher, multimode	► Shortwave Fibre Channel or FICON host adapters (Feature codes 3153 and 3157) ► **LC** connector
1411	31 m (102 ft.)	50 micron OM3 or higher, multimode	
1412	2 m (6.5 ft.)	50 micron OM3 or higher, multimode	
1420	31 m (102 ft.)	9 micron OM3 or higher, single mode	► Longwave Fibre Channel or FICON Host Adapters (Feature codes 3253 or 3257) ► **LC** connector
1421	31 m (102 ft.)	9 micron OM3 or higher, single mode	
1422	2 m (6.5 ft.)	9 micron OM3 or higher, single mode	

For more information about IBM supported attachments, see the *IBM System Storage DS8880 Introduction and Planning Guide,* GC27-8525.

For the latest information about host types, models, adapters, and operating systems that are supported by the DS8880, see the DS8000 System Storage Interoperation Center (SSIC) at this website:

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

7.2.7 Host adapter Fibre Channel specifics for open environments

Each storage system host adapter has four or eight ports, and each port has a unique worldwide port name (WWPN). Each port can be configured as SCSI-FCP or FICON topology by using the data storage Management GUI or the DS CLI. Host adapters can be shortwave or longwave. Additional host adapters can be installed up to two host adapters per I/O enclosure.

With host adapters that are configured as FC protocols, the DS8880 provides the following configuration capabilities:

► A maximum of 128 Fibre Channel ports.

► A maximum of 509 logins per Fibre Channel port, which includes host ports and Peer-to-Peer Remote Copy (PPRC) target and initiator ports.

► Access to 63750 logical unit numbers (LUNs) for each target (one target for each host adapter), depending on the host type.

► Either Fibre Channel Arbitrated Loop (FC-AL), switched-fabric (FC-SW), or point-to-point topologies.

► The 16 Gbps adapter does not support arbitrated loop topology at any speed.

7.2.8 FICON specifics on z/OS environment

For host adapters that are configured for FICON, the DS8880 provides the following configuration capabilities:

► Fabric or point-to-point topologies

► A maximum of 64 ports on the DS8884 storage systems and a maximum of 128 host adapter ports on the DS8886 storage system

► A maximum of 509 logins for each host adapter port

► A maximum of 8,192 logins for each storage unit

► A maximum of 1,280 logical paths on each host adapter port

► Access to all 255 control-unit images (65,280 count-key-data (CKD) devices) over each FICON port

► A maximum of 512 logical paths for each control-unit image

Note: The IBM z13 servers support 32,000 devices for each FICON host channel. The IBM zEnterprise EC12 and IBM zEnterprise BC12 servers support 24,000 devices for each FICON host channel. Earlier z Systems servers support 16,384 devices for each FICON host channel. To fully access 65,280 devices, it is necessary to connect multiple FICON host channels to the storage system. You can access the devices through a Fibre Channel switch or FICON director to a single storage system.

7.2.9 Preferred practice for host adapters

For optimum availability and performance, the following practices are preferred:

► To obtain the maximum ratio for availability and performance, install one HA card on each available I/O enclosure before you install the second HA card on the same I/O enclosure.

► The DS8880 supports 16 Gbps 4-port host adapters, and 8 Gbps 4-port and 8-port host adapters. Based on the configuration, these host adapters or an intermixture of them can be installed in the DS8880.

► Better performance can be obtained from Copy Services by using dedicated host adapters for remote copy links.

7.2.10 WWNN and WWPN determination

The incoming and outgoing data to the DS8880 is tracked by using worldwide node name (WWNN) and worldwide port name (WWPN). For the DS8000, each storage facility image (SFI) has its own unique WWNN. The storage unit itself also has a unique WWNN. Each host adapter port has a unique and persistent WWPN for attachment to a SAN. The WWNN and WWPN values can be determined by using the DS CLI or DS Storage Management GUI.

Determining a WWNN by using a DS CLI

The DS8880 WWNN has an address that is similar to the following strings:

```
50:05:07:63:0z:FF:Cx:xx
50:50:07:63:0z:FF:Dx:xx
```

The z and x:xx values are unique combinations for each system and each SFI that are based on a machine's serial number. Use the DS CLI command `lssi` to determine the SFI WWNN, as shown in Example 7-1.

Example 7-1   SFI WWNN determination

```
dscli> lssi
Date/Time: 09 December 2016 13:12:11 CET IBM DSCLI Version: 7.8.20.280 DS: -
Name ID                Storage Unit    Model WWNN              State ESSNet
===========================================================================
-    IBM.2107-75DMC11 IBM.2107-75DMC10 981   5005076306FFD33A Online Enabled
```

Do not use the `lssu` command because it determines the storage unit WWNN, which is not used. Attached hosts see only the SFI, as shown in Example 7-2.

Example 7-2   Machine WWNN

```
dscli> lssu
Date/Time: 09 December 2016 13:14:11 CET IBM DSCLI Version: 7.8.20.280 DS: -
Name ID               Model WWNN            pw state
====================================================
-    IBM.2107-75DMC10 981   5005076306FFEB3A On
```

Determining a WWPN by using a DS CLI

Similar to the WWNN, a WWPN in the DS8880 looks like the following address:

`50:05:07:63:0z:YY:Yx:xx`

However, the DS8880 WWPN is a child of the SFI WWNN, where the WWPN inserts the z and x:xx values from SFI WWNN. It also includes the YY:Y, from the logical port naming, which is derived from where the host adapter is physically installed. Use the DS CLI command `lsioport` to determine the SFI WWPN, as shown in Example 7-3.

Example 7-3   WWPN determination

```
dscli> lsioport
Date/Time: 09 December 2016 13:34:44 CET IBM DSCLI Version: 7.8.20.280 DS:
IBM.2107-75DMC11
ID    WWPN             State Type              topo     portgrp
================================================================
I0000 500507630600133A Online Fibre Channel-LW FICON    0
I0001 500507630600533A Online Fibre Channel-LW FICON    0
I0002 500507630600933A Online Fibre Channel-LW SCSI-FCP 0
I0003 500507630600D33A Online Fibre Channel-LW SCSI-FCP 0
I0010 500507630601133A Online Fibre Channel-SW SCSI-FCP 0
I0011 500507630601533A Online Fibre Channel-SW SCSI-FCP 0
I0012 500507630601933A Online Fibre Channel-SW SCSI-FCP 0
I0013 500507630601D33A Online Fibre Channel-SW SCSI-FCP 0
I0040 500507630604133A Online Fibre Channel-SW SCSI-FCP 0
I0041 500507630604533A Online Fibre Channel-SW SCSI-FCP 0
I0042 500507630604933A Online Fibre Channel-SW SCSI-FCP 0
I0043 500507630604D33A Online Fibre Channel-SW SCSI-FCP 0
I0100 500507630608133A Online Fibre Channel-LW FICON    0
I0101 500507630608533A Online Fibre Channel-LW FICON    0
I0102 500507630608933A Online Fibre Channel-LW SCSI-FCP 0
I0103 500507630608D33A Online Fibre Channel-LW SCSI-FCP 0
I0230 500507630613133A Online Fibre Channel-SW SCSI-FCP 0
```

```
I0231 500507630613533A Online Fibre Channel-SW SCSI-FCP 0
I0232 500507630613933A Online Fibre Channel-SW SCSI-FCP 0
I0233 500507630613D33A Online Fibre Channel-SW SCSI-FCP 0
I0330 50050763061B133A Online Fibre Channel-SW SCSI-FCP 0
I0331 50050763061B533A Online Fibre Channel-SW SCSI-FCP 0
I0332 50050763061B933A Online Fibre Channel-SW SCSI-FCP 0
I0333 50050763061BD33A Online Fibre Channel-SW SCSI-FCP 0
I0334 50050763065B133A Online Fibre Channel-SW SCSI-FCP 0
I0335 50050763065B533A Online Fibre Channel-SW SCSI-FCP 0
I0336 50050763065B933A Online Fibre Channel-SW SCSI-FCP 0
I0337 50050763065BD33A Online Fibre Channel-SW SCSI-FCP 0
```

Determining a WWNN by using a DS GUI

Use the following guidelines to determine the WWNN by using the DS8870 Storage Management GUI:

1. Connect by using a web browser to the HMC IP address:

   https://< *hmc IP address* >

2. Select **Actions**.

3. Select **Properties** to obtain the WWNN value, as shown in Figure 7-5.
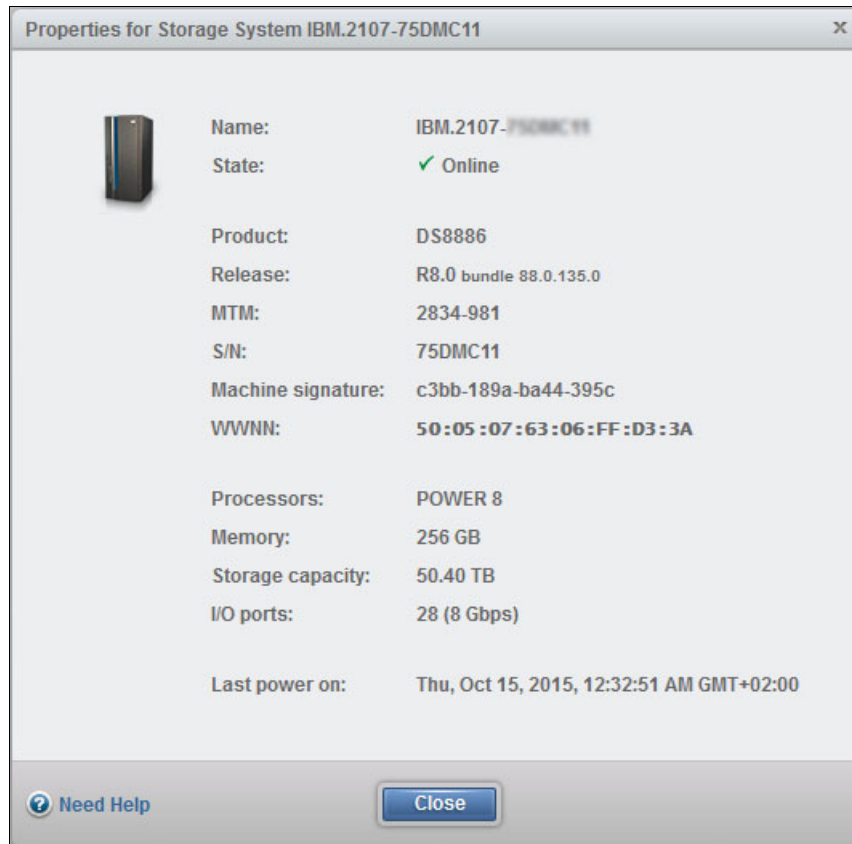
Figure 7-5   SFI WWNN value

Determine a WWPN by using a web GUI

Use the Storage Management GUI to determine the HA port WWPN:

1. Connect to the HMC IP address by using a web browser:

   `https://< hmc IP address >`

2. Select **Actions**.

3. Select **Modify I/O Port Protocols**.

   The default view shows protocols and state only. The view can be customized to display the port WWPN and the Frame.

4. Click **Actions**, and then select **Customize Columns** to include the WWPN and Frame in the view. You receive the full list of each installed I/O port with its WWPN and its physical location, as shown in Figure 7-6.
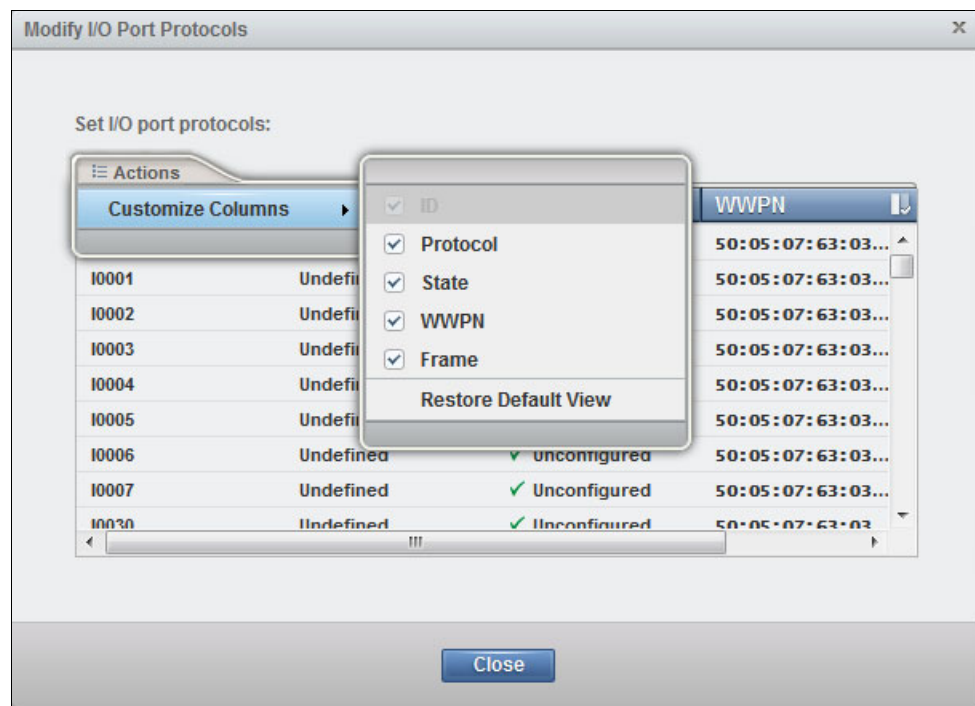
Figure 7-6   Determining the I/O port WWPN

7.3 Network connectivity planning

Implementing the DS8880 requires consideration of the physical network connectivity of the Hardware Management Console within your local area network.

Consider the following network and communications issues when you plan the location and interoperability of your storage systems:

► HMC network access (one IP per HMC)
► Remote support connection
► SAN connectivity
► IBM Security Key Lifecycle Manager connection if encryption is activated
► LDAP connection if LDAP is implemented

For more information about physical network connectivity, see the *IBM System Storage DS8880 Introduction and Planning Guide*, GC27-8525.

7.3.1 Hardware Management Console and network access

The HMCs are the focal point for the configuration of Advanced Function management, and maintenance for a DS8880 unit. The internal Hardware Management Console (HMC1) that is included with every base frame includes Ethernet adapters.

An optional, additional Hardware Management Console can be purchased. This secondary Hardware Management Console (HMC2) is mounted in the base of the primary rack, and it has connectivity to the storage system private networks. A dual Ethernet connection is also available for client access. The secondary HMC provides redundant management access to enable continuous availability access for encryption key servers and other advanced functions.

> **Important:** The secondary HMC is directly connected to the private DS8880 Ethernet switches. An Ethernet connection for the client network is also available.

The HMC can be connected to the client network for the following tasks:

▶ Remote management of your system by using the DS CLI

▶ Remote management by using the DS Storage Management GUI by opening a browser to the network address of the HMC:

```
https://< HMC IP address>
```

To access the Hardware Management Consoles (HMC1/HMC2) over the network, provide the following information:

▶ Hardware Management Console: For each Hardware Management Console, determine one TCP/IP address, host name, and domain name.

▶ Domain name server (DNS) settings: If a DNS is to be implemented, ensure that it is reachable to avoid contention or network timeout.

▶ Gateway routing information: Supply the necessary routing information.

For more information about HMC planning, see Chapter 8, "IBM DS8880 Management Console planning and setup" on page 203.

> **Important:** The DS8880 uses 172.16.y.z and 172.17.y.z private network addresses. If the client network uses the same addresses, the IBM SSR can reconfigure the private networks to use another address range option.

7.3.2 IBM Spectrum Control

IBM Spectrum Control™ is an integrated software solution that can help you improve and centralize the management of storage environments. With IBM Spectrum Control, you can manage and configure multiple DS8000 storage systems from a single point of control.

IBM Spectrum Control simplifies storage management by providing the following benefits:

- ► Centralizing the management of heterogeneous storage network resources with IBM storage management software
- ► Providing greater synergy between storage management software and IBM storage devices
- ► Reducing the number of servers that are required to manage your software infrastructure
- ► Migrating from basic device management to storage management applications that provide higher-level functions

7.3.3 Data storage command-line interface

The IBM DS CLI can be used to create, delete, modify, and view Copy Services functions and for the logical configuration of a storage unit. These tasks can be performed interactively, in batch processes (operating system shell scripts), or in DS CLI script files. A *DS CLI script file* is a text file that contains one or more DS CLI commands. It can be issued as a single command. DS CLI can also be used to manage other functions for a storage system, including managing security settings, querying point-in-time performance information or the status of physical resources, and exporting audit logs.

The DS CLI client can be installed on a workstation, and can support multiple operating systems. The DS CLI client can access the DS8880 over the client's network.

For more information about hardware and software requirements for the DS CLI, see the *IBM System Storage DS Command-Line Interface User's Guide for DS8000 series*, SC53-1127.

7.3.4 Remote support connection

Embedded AOS
The preferred remote support connectivity method for IBM is through Transport Layer Security (TLS) for the management console to IBM communication. DS8880 uses an embedded Assist On-Site server solution. Embedded AOS is a secure, fast, broadband form of remote access.

The previous remote solutions, *Internet Protocol Security* (IPSec), which is also called a *virtual private network* (VPN), and modem are no longer options in the DS8880 family.

For more information, see Chapter 8, "IBM DS8880 Management Console planning and setup" on page 203 and Chapter 15, "Remote support" on page 433.

A typical AOS remote support connection is shown in Figure 7-7.

AOS REMOTE SUPPORT FLOW

Secure SSL/TLS connection

Built in AOS Gateway

AOS Relay Servers

IBM Firewall

Customer Firewall

IBM Remote Support Infrastructure

[Step 1] - The AOS gateway establishes and maintains a secure connection between IBM and the customer.

[Step 2] - IBM Support starts a remote access session if allowed by the customer.

[Step 3] - The AOS server establishes a remote access session after verifying customer and IBM Support credentials.

[Step 4] - The AOS gateway server controls the application data flow between the DS8000 and the IBM support engineer.

Figure 7-7   DS8880 AOS remote support flow

Complete the following steps to prepare for attaching the DS8880 to the client's network:

1. Assign a TCP/IP address and host name to the HMC in the DS8880.

2. If email notification for service alerts is allowed, enable the support on the mail server for the TCP/IP addresses that is assigned to the DS8880.

3. Use the information that was entered on the configuration worksheets during your planning.

Remote Support Console

DS8880 Release 8.2 also supports a web-based remote support option called Remote Support Console (RSC). RSC provides additional isolation between IBM support and the DS8880 system by forcing all remote support actions to be performed by using a web-based console terminal interface.

7.3.5 Storage area network connection

The DS8880 can be attached to a storage area network (SAN) environment through its HA ports. The SAN provides the capability to interconnect open systems hosts, z Systems hosts, and other storage systems.

A SAN allows your host bus adapter (HBA) host ports to have physical access to multiple HA ports on the storage system. Zoning can be implemented to limit the access (and provide access security) of host ports to the storage system.

Shared access to a storage system HA port is possible from hosts that support a combination of host bus adapter types and operating systems.

> **Important:** A SAN administrator must verify periodically that the SAN is working correctly before any new devices are installed. SAN bandwidth must also be evaluated to ensure that it can handle the new workload.

7.3.6 Key manager servers for encryption

The DS8880 is delivered with FDE drives. When you activate encryption, isolated key managers are required.

With DS8880 Release 8.1 code, you have a choice between IBM SKLM and the Gemalto SafeNet KeySecure, if you must comply with KMIP. Note that you cannot mix SKLM and Safenet key servers.

For more information, see *IBM DS8880 Data-at-rest Encryption*, REDP-4500.

Encryption activation review planning

IBM Encryption offerings must be activated before they are used. This activation is part of the installation and configuration steps that are required to use the technology.

Using the IBM Security Key Lifecycle Manager

A Security Key Lifecycle Manager license is required for use with the SKLM software. Two isolated SKLM servers are required to enable encryption on the DS8880 system.

IBM Storage Appliance 2421 Model AP1 can be ordered in these configurations:

- ► A single isolated key server (Feature Code (FC) 1761)
- ► Two isolated key servers (FC1761 and FC1762, which are ordered together)

This order must include an indicator for IBM Security Key Lifecycle Manager (FC0952), which indicates that a DVD with IBM Security Key Lifecycle Manager software is provided with the Storage Appliance AP1. For more information, search for `IBM Storage Appliance 2421 Model AP1` at the IBM Publications Center website:

http://www.ibm.com/shop/publications/order

> **Important:** Regardless of the ordering method, clients must acquire an SKLM license for the use of the SKLM software, which is ordered separately from the stand-alone server hardware.
>
> **Note:** The licensing for SKLM includes both an installation license for the SKLM management software and licensing for the encrypting drives.

The DS8000 series supports IBM Security Key Lifecycle Manager v2.5 or later.

For encryption and National Institute of Standards and Technology (NIST) SP-8131, a compliant connection between the HMC and the key server, IBM Security Key Lifecycle Manager v2.5 or later, is required. Starting from v2.5, the name of Security Key Lifecycle Manager was changed to IBM Security Key Lifecycle Manager.

Isolated key servers that are ordered with MT 2421-AP1 and Feature Code 1762 include a Linux operating system and IBM Security Key Lifecycle Manager or IBM Security Key

Lifecycle Manager software that is preinstalled. You are advised to upgrade to the latest version of the IBM Security Key Lifecycle Manager.

Security Key Lifecycle Manager connectivity and routing information

To connect the IBM SKLM to your network, provide the following settings to your IBM SSR:

► IBM Security Key Lifecycle Manager server network IDs, host names, and domain name
► DNS settings (if you plan to use DNS to resolve network names)

Two network ports must be opened on a firewall to allow DS8880 connection and to obtain an administration management interface to the IBM Security Key Lifecycle Manager server. These ports are defined by the IBM Security Key Lifecycle Manager administrator.

For more information, see the following IBM publications for SKLM:

► *IBM Security Key Lifecycle Manager Quick Start Guide*, GI13-2316
► *IBM Security Key Lifecycle Manager Installation and Configuration Guide*, SC27-5335

7.3.7 Lightweight Directory Access Protocol server for single sign-on

The DS8000 system supports both unified sign-on or single sign-on (SSO) functions, which are available through the DS8000 Storage Management GUI, and the ability to specify an existing LDAP server. The LDAP server can use existing users and user groups for authentication on the DS8000 system.

With Release 8.1 and later, LDAP can be enabled through Copy Services Manager, which is packaged with the HMC code. Alternatively, you can use the LDAP access going through any Copy Services Manager server. For more information, see *IBM DS8880 Integrated Copy Services Manager and LDAP Client on the HMC*, REDP-5356.

7.4 Remote Mirror and Copy connectivity

The DS8880 uses the high-speed FCP for Remote Mirror and Copy connectivity. Ensure that you assigned sufficient FCP paths for the remote mirroring between the source and target sites to address performance and redundancy issues. When you plan Metro Mirror and Global Copy modes between a pair of storage systems, use separate logical and physical paths for the Metro Mirror. Use another set of logical and physical paths for the Global Copy.

Plan the distance between the primary and auxiliary storage systems carefully to correctly acquire the necessary length of fiber optic cables that are required. If necessary, the Copy Services solution can include hardware, such as channel extenders or dense wavelength division multiplexing (DWDM).

For more information, see *IBM DS8870 Copy Services for Open Systems*, SG24-6788, and *IBM DS8870 Copy Services for IBM z Systems*, SG24-6787.

7.5 Disk capacity considerations

The effective capacity of the DS8880 is determined by the following factors:

► The spare configuration
► The capacity of the installed drives
► The selected RAID configuration: RAID 5, RAID 6, or RAID 10
► The storage type: Fixed block (FB) or CKD

7.5.1 Disk sparing

RAID arrays automatically attempt to recover from a drive failure by rebuilding the data for the failed drive to a spare disk drive module (DDM). For sparing to occur, a drive with a capacity equal to or greater than the failed drive must be available on the same device adapter pair. After the sparing is initiated, the spare and the failing drives are swapped between their respective array sites so that the spare drive becomes part of the array site that is associated with the array at the failed drive. The failing drive becomes a failed spare drive in the array site from which the spare came.

Standard drive enclosure

The DS8880 assigns spare disks automatically. The following list shows the requirements for spares:

► Four spares are needed for each DA pair of the same capacity and speed.

► If the speed is the same, the spares are the higher capacity.

► If only two arrays are on the DA pair, only two spares are assigned. Two more spares are assigned when two more arrays are added to the DA pair.

The *enhanced sparing* feature allows deferred service of the failed drives until sufficient drives have failed to warrant a service call. A service call will always occur if any DA pair has only one remaining good spare.

High-Performance Flash Enclosure Gen2

High-Performance Flash Enclosures (HPFEs) are installed in pairs, with 16, 32, or 48 Flash Cards per enclosure pair. Two spare flash cards are assigned for each HPFE Gen2 pair. If a flash card fails, and a spare is taken, the system will call for service, because only one good spare remains in the HPFE Gen2 pair (DA pair).

For more information about the DS8000 sparing concepts, see 3.5.9, "Spare creation" on page 89.

7.5.2 Disk capacity

The DS8880 operates in a RAID 5, RAID 6, or RAID 10 configuration. The following RAID configurations are possible:

► 6+P RAID 5 configuration: The array consists of six data drives and one parity drive. The remaining drive on the array site is used as a spare.

► 7+P RAID 5 configuration: The array consists of seven data drives and one parity drive.

► 5+P+Q RAID 6 configuration: The array consists of five data drives and two parity drives. The remaining drive on the array site is used as a spare.

► 6+P+Q RAID 6 configuration: The array consists of six data drives and two parity drives.

► 3+3 RAID 10 configuration: The array consists of three data drives that are mirrored to three copy drives. Two drives on the array site are used as spares.

► 4+4 RAID 10 configuration: The array consists of four data drives that are mirrored to four copy drives.

> **Note:** DS8880 hardware release 8.2 uses RAID 6 by default. RAID 5 is supported for disk sizes smaller than 1 TB, but is not recommended. RAID 5 is not supported for disk sizes greater than 1 TB.

Refer to the *IBM DS8880 Intoduction and Planning Guide*, GC27-8525 for the effective capacity of one rank in the various possible configurations.

A drive set contains 16 drives, which form two array sites. The drive capacity is added in increments of one disk drive set. Flash cards are added in increments of drive sets. Sixteen cards are included with HPFE, and an optional second drive set of 14 flash cards can be added. The capacities in the table are expressed in decimal gigabytes and as the number of extents.

> **Important:** Because of the larger metadata area, the net capacity of the ranks is lower than in previous DS8000 models.

> **Important:** When you review teh effective capacity, keep in mind the following points:
>
> ► Effective capacities are in decimal gigabytes (GB). One GB is 1,000,000,000 bytes.
>
> ► Although drive sets contain 16 drives, arrays use only eight drives. The effective capacity assumes that you have two arrays for each disk drive set.

An updated version of Capacity Magic can help you determine the raw and net storage capacities and the numbers for the required extents for each available type of RAID.

7.5.3 DS8000 flash drives: Considerations

Flash drives (solid-state drives (SSDs)) are a higher performance option when compared to hard disk drives (HDDs). For the DS8880, flash drives are available in 400 GB, 800 GB, and 1.6 TB capacities.

All drives that are installed in a standard drive enclosure pair must be of the same capacity and speed.

Flash drives are ordered in drive sets of 16 for all capacities.

Limitations
Drives of different capacities cannot be intermixed in a standard drive enclosure pair.

Placement
The following rules apply to the placement of flash drives:

► Flash drive sets are installed and configured by IBM manufacturing based on the flash drive enclosure feature that is listed on the client's order. These features are an indicator that determines how manufacturing distributes flash drive sets over the available flash drive enclosure pairs.

For example, if the client order specifies two flash drive enclosure pairs (FC1245) and two drive sets of 400 GB flash drives, the system ships with two flash drive enclosure pairs with 16 flash drives for each enclosure pair.

► The sequence in which drive types are installed is based on capacity, from the highest to the lowest. Therefore, 1.6 TB drives are installed first, followed by 800 GB, and then 400 GB flash drives.

7.5.4 DS8000 flash cards: Considerations

Flash cards are a higher performance option when compared to flash drives and other hard disk drives. Flash cards are installed in the HPFE Gen2 Pairs. 2.5-inch Flash cards are available with 400 GB, 800 GB, 1.6 TB, or 3.2 TB capacity.

When an HPFE is ordered, also order a set of 16 flash cards. Additionally, the optional second flash card set of 14 can be ordered to fully configure the HPFE to 30 flash cards.

Flash cards in HPFE Gen2 pairs support RAID 6 by default. RAID 10 is also supported, as is RAID 5 for flash card sizes smaller than 1 TB.

The system always places volume metadata on the highest performance tier of an extent pool. To ensure best performance for your system, have at least ten percent of flash capacity in each pool. This percentage of flash capacity should be enough to ensure that all volume metadata resides on flash.

8

IBM DS8880 Management Console planning and setup

This chapter describes the planning tasks that are involved in the setup of the required DS8880 Management Console, which is also known as the Hardware Management Console (HMC).

This chapter covers the following topics:

► DS8880 Management Console overview
► Management Console software
► Management Console activities
► Management Console and IPv6
► Management Console user management
► Optional Secondary Management Console

8.1 DS8880 Management Console overview

The Management Console (MC) is a multi-purpose piece of equipment that provides the services that the client needs to configure and manage the storage and manage several of the operational aspects of the storage system. It also provides the interface where service personnel perform diagnostic and repair tasks. The MC does not process any of the data from hosts. It is not even in the path that the data takes from a host to the storage. The MC is a configuration and management station for the DS8880.

The MC, which is the focal point for DS8880 management, includes the following functions:

► DS8880 power control
► Storage provisioning
► Storage Subsystem health monitoring
► Storage Subsystem performance monitoring
► Advanced Copy Services management
► Embedded Copy Services Manager
► Interface for onsite service personnel
► Collection of diagnostic data and call home
► Problem management and alerting
► Remote support access
► Storage management through the data storage graphical user interface (DS GUI)
► Connection to IBM Security Key Lifecycle Manager (SKLM) for encryption management functions, if required
► Interface for Licensed Internal Code (LIC) and other firmware updates

Every DS8880 installation includes an MC that is in the base rack. A second optional MC can be installed in the DS8880 base rack to provide redundancy.

8.1.1 Management Console hardware

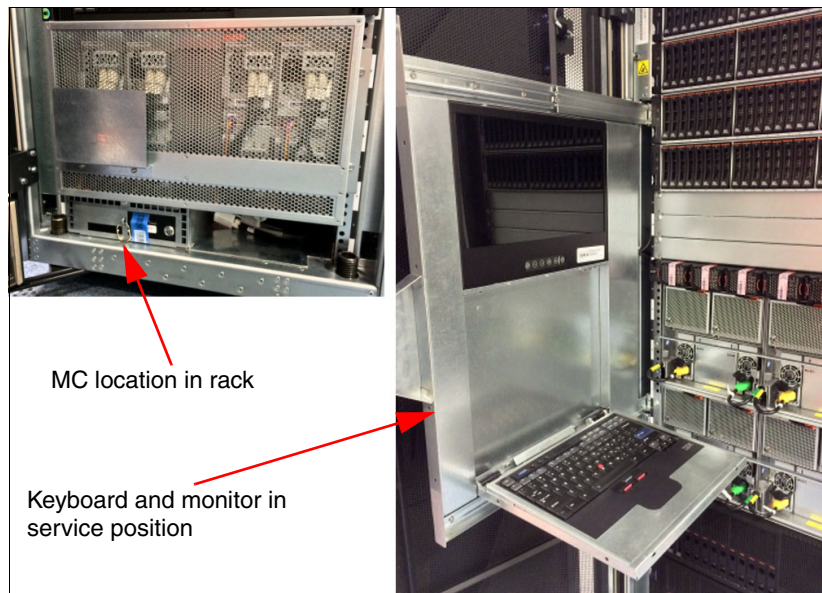The MC consists of a small form factor (SFF) personal computer, as shown in Figure 8-1.

Figure 8-1   MC location in DS8880

The use of an SFF personal computer makes the MC efficient in many ways, including power consumption. The MC is mounted on the bottom of the rack, underneath the direct current uninterruptible power supply (DC-UPS). Because of the small width, both primary and secondary MCs are mounted in the base rack. The keyboard and display slide out from the left side of the rack for service. The MCs are equipped with Ethernet connections for the client's network.

A second, redundant, MC workstation is orderable and needs to be used in environments that use encryption management or Advanced Copy Services functions. The second MC is installed in the DS8880 rack. For more information about adding a secondary MC, see 8.6, "Optional Secondary Management Console" on page 217.

8.1.2 Private Ethernet networks

The MC communicates with the storage facility through a pair of redundant Ethernet networks, which are designated as the Black Network and the Gray Network. Two switches are included in the rear of the DS8880 base rack. Each MC and each central processor complex (CPC) is connected to both switches. Figure 8-2 on page 206 shows how each port is used on the pair of DS8880 Ethernet switches. Do not connect the client network (or any other equipment) to these switches because they are for the DS8880 internal use only.

In most DS8880 configurations, two or three ports are unused on each switch.

Important: The internal Ethernet switches that are shown in Figure 8-2 on page 206 are for the DS8880 private network only. Do not connect directly to these internal switches from the client network.
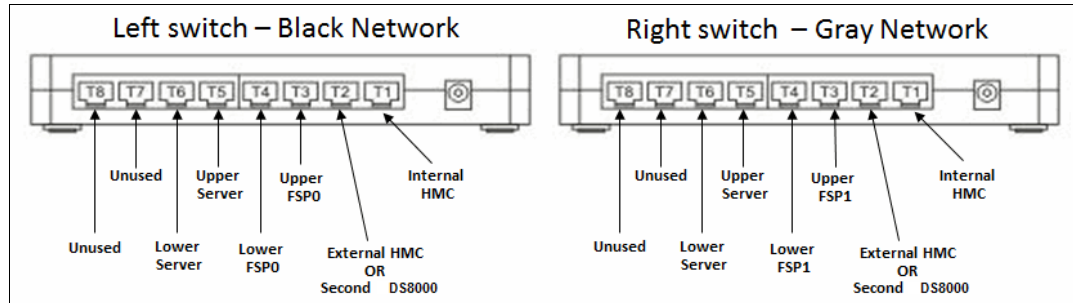

Figure 8-2   DS8880 internal Ethernet switches

8.2 Management Console software

The MC, which is based on Linux, includes the following application servers:

► DS Management GUI

   The DS GUI server is used to perform configuration and management tasks.

► IBM enterprise storage server network interface (ESSNI) server

   ESSNI is the logical server that communicates with the DS GUI server and interacts with the two processor nodes of the DS8880. It is also referred to as the DS Network Interface (DSNI).

► RESTful application programming interface (API) services

   The DS8880 provides industry-standard Representational State Transfer (RESTful) API services for management applications and RESTful clients. RESTful services are running on the MC and can be upgraded without requiring a DS8880 LIC update.

The DS8880 MC provides the following management interfaces:

► DS Management graphical user interface (GUI)
► Data storage command-line interface (DS CLI)
► DS Open application programming interface (DS Open API)
► RESTful application programming interface (RESTful API)
► Web-based user interface (Web UI)
► Copy Services Manager interface (CSM)

The GUI and the DS CLI are comprehensive, easy-to-use interfaces for a storage administrator to perform DS8880 management tasks to provision the storage arrays, manage application users, and change MC options. The interfaces can be used interchangeably, depending on the particular task.

The DS Open API provides an interface for external storage management programs, such as Spectrum Control, to communicate with the DS8880. It channels traffic through the IBM System Storage Common Information Model (CIM) agent, which is a middleware application that provides a CIM-compliant interface.

Similar to DS Open API, the RESTful API services provide an interface for external management programs and applications to interact with the DS8880. Clients can develop and tailor their specific DS8880 management applications based on the standard RESTful APIs.

Copy Services Manager, which replaced the Tivoli Productivity Center for Replication, is now embedded within the DS8880 code from release 8.1. This simple and effective replication management and automation tool is now part of the MC code, removing the requirement for an external server to host the software.

8.2.1 DS Management GUI

Although the DS Management GUI runs on the MC, it is not possible to access it when you are logged in to the MC console. It can be accessed remotely either by using a web browser on a workstation that is attached to the client's network to which the MC is attached, or through Spectrum Control. For more information, see 11.2.1, "Accessing the DS8000 Storage Management GUI" on page 253.

8.2.2 Data storage command-line interface

The data storage command-line interface (DS CLI), which must be run in the command environment of an external workstation, is a second option to communicate with the MC. The DS CLI is a good choice for configuration tasks when many updates are needed. This option avoids the web page load time for each window in the DS GUI when you perform Copy Services tasks.

For more information about DS CLI use and configuration, see Chapter 12, "Configuration with the data storage command-line interface" on page 347. For a complete list of DS CLI commands, see *IBM Systems Storage DS8000 Series: Command-Line Interface User's Guide*, GC27-4212.

8.2.3 DS Open application programming interface

Calling DS Open application programming interfaces (DS Open APIs) from within a program is a third option to implement communication with the MC. The DS CLI and DS Open API communicate directly with the ESSNI server software that is running on the MC.

For the DS8000, the CIM agent is preinstalled with the MC code and is started when the MC starts. An active CIM agent allows access only to the DS8000s that are managed by the MC on which it is running. Configuration of the CIM agent must be performed by an IBM SSR by using the DS CIM command-line interface (DSCIMCLI). For more information about the CIM agent, see this website:

http://www.snia.org/forums/smi/tech_programs/lab_program

8.2.4 RESTful application programming interface

DS8880 RESTful API services provides a similar but easier to use application programming interface (RESTful API) to manage DS8880 through communication with the MC. The RESTful API communicates with RESTful services that run on the MC. The RESTful services in turn interact with the ESSNI server software that runs on the MC to pass requests and receive replies. For more information about the RESTful API, see *Deploying Flex System in a BladeCenter Environment*, REDP-5122.

8.2.5 Copy Services Manager interface

Embedded CSM offers a preinstalled GUI and CLI on your DS8880 HMC to manage and automate replication and disaster recovery for up to four DS8000 storage systems. Although this feature is only available in release 8.1 of the DS8880 microcode, it can also manage previous release systems based on CSM interoperability. Embedded CSM functions in active/standby coordination with CSM installed and running an external server.

This new feature removes the requirement for an external server to host CSM, providing savings on infrastructure costs and OS licensing. Administration costs are also reduced, as the embedded CSM instance is upgraded through the DS8880 code maintenance schedule, which is performed by IBM support personnel.

For more information about embedded CSM, see *IBM DS8880 Integrated Copy Services Manager and LDAP Client on the HMC*, REDP-5356.

> **Important:** Although possible, IBM does *not* recommend configuring the primary HMC and the secondary HMC of the same storage system as the active and standby CSM servers of an environment.

8.2.6 Web-based user interface

The web-based user interface (Web UI) is a browser-based interface that is used for remote access to system utilities.

Complete the following steps to log in to the MC by using the Web UI:

1. Start the MC Web GUI as shown in Figure 8-3. Click the **Service** icon to access the Service Management Console.
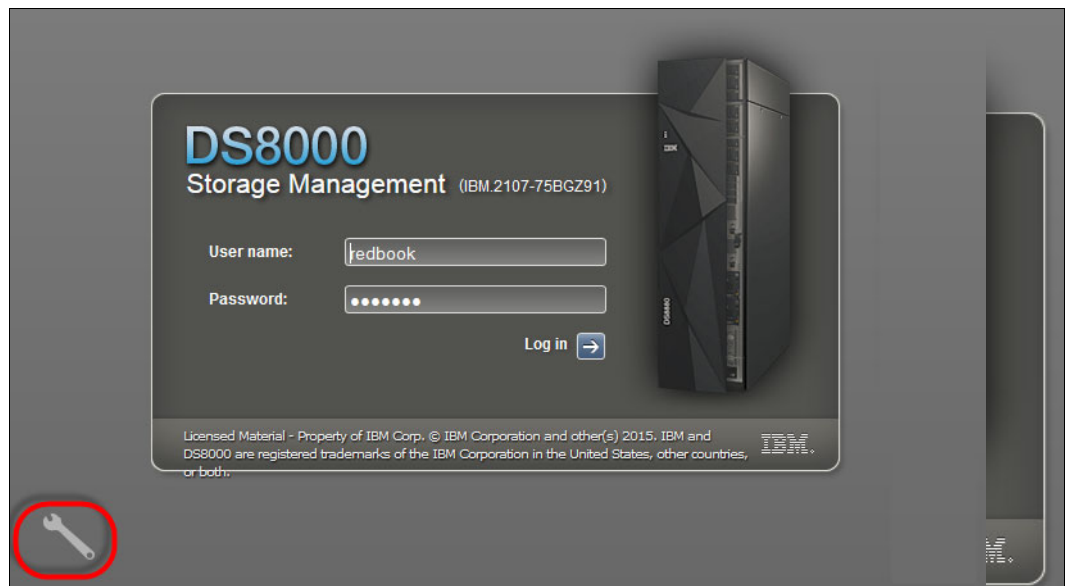
Figure 8-3   DS Management GUI logon panel

2. Click **Log on and launch the Hardware Management Console web application** to open the login window as shown in Figure 8-4 and log in. The default user ID is `customer` and the default password is `cust0mer`.

Figure 8-4   Service Management Console application

3. If you are successfully logged in, you see the MC window, in which you can select **Status Overview** to see the status of the DS8880. Other areas of interest are shown in Figure 8-5.

Figure 8-5   Web UI main window

Because the MC Web UI is mainly a services interface, it is not covered here. For more information, see the **Help** menu.

8.3 Management Console activities

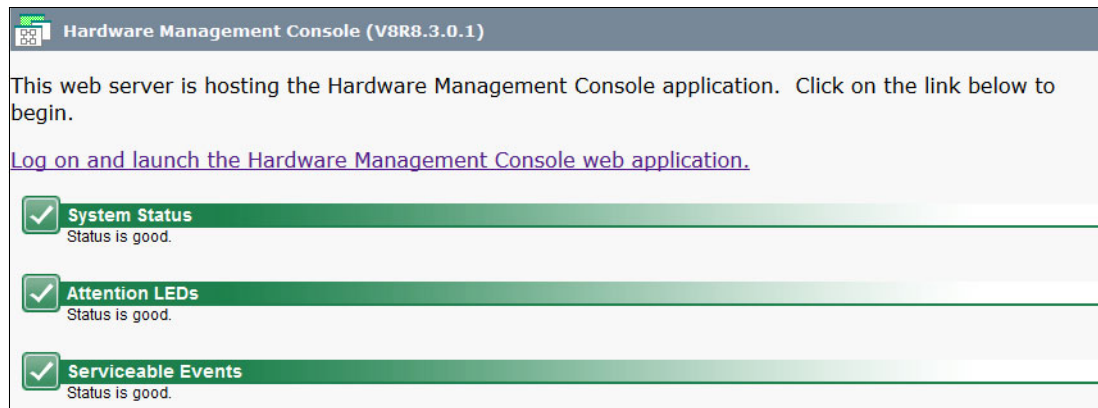This section covers planning and maintenance tasks for the DS8880 MC. For more information about overall planning, see Chapter 7, "IBM DS8880 physical planning and installation" on page 179. If a second external MC was ordered for the DS8880, information about planning that installation is included. If a second, external MC was not ordered, the information can be safely ignored.

8.3.1 Management Console planning tasks

The following tasks are needed to plan the installation or configuration:

► A connection to the client network is needed at the base rack for the primary MC. Another connection is also needed at the location of the secondary MC. The connections must be standard CAT5/6 Ethernet cabling with RJ45 connectors.

► IP addresses for the primary and secondary MCs are needed. The DS8880 can work with IPv4 and IPv6 networks. For more information about procedures to configure the DS8880 MC for IPv6, see 8.4, "Management Console and IPv6" on page 213.

► Most users access the DS GUI remotely through a browser. You can also use Spectrum Control in your environment to access the DS GUI.

► The web browser to be used on any administration workstation must be supported, as described in the *IBM System Storage DS8880 Introduction and Planning Guide*, GC27-8525.

► The IP addresses of Simple Network Management Protocol (SNMP) recipients must be identified if the client wants the DS8880 MC to send SNMP traps to a monitoring station.

► Email accounts must be identified if the client wants the DS8880 MC to send email messages for problem conditions.

► The IP addresses of Network Time Protocol (NTP) servers must be identified if the client wants the DS8880 MC to use NTP for time synchronization.

► When a DS8880 is ordered, the license and certain optional features must be activated as part of the customization of the DS8880. For more information, see Chapter 9, "IBM DS8880 features and licensed functions" on page 219.

► The installation tasks for the optional external MC must be identified as part of the overall project plan and agreed upon with the responsible IBM personnel.

> **Important:** Applying increased feature activation codes is a concurrent action.

8.3.2 Planning for Licensed Internal Code upgrades

The following tasks must be considered regarding the LIC upgrades on the DS8880:

► LIC changes

 IBM might release changes to the DS8880 series Licensed Machine Code.

► LIC installation

 An IBM SSR can install the changes. Check whether the new LIC requires new levels of DS CLI and DS Open API. Plan on upgrading them on the relevant workstations, if necessary.

► Code prerequisites

When you are planning for initial installation or for LIC updates, ensure that all prerequisites for the environment are identified correctly. These include, but are not limited to, host operating system versions, fixes, and host bus adapter (HBA) levels, interconnect/fabric types, and operating system versions.

DS8880 interoperability information is available at the IBM System Storage Interoperation Center (SSIC) at this website:

http://www.ibm.com/systems/support/storage/config/ssic

To prepare for downloading the drivers, see Interoperability Search Details in the SSIC report. This report provides an end-to-end support matrix from the host to the DS8880, and covers all versions of operating system, multipathing software, and firmware. This check is necessary to ensure that the DS8880 storage subsystem is in a supported environment.

> **Important:** The SSIC includes information about the latest supported code levels. This availability does not necessarily mean that former levels of HBA firmware or drivers are no longer supported. Some host type interoperability, such as Netapp Ontap might need to be confirmed in the vendor's support matrix. If you are in doubt about any supported levels, contact your IBM representative.
>
> *Never* proceed with a LIC update without adhering to all prerequisites.

► Maintenance windows

The LIC update of the DS8880 is a designed to be a nondisruptive action. A concurrent maintenance window, with added time for contingency, is still desirable. This contingency period gives you time to confirm all environment pre-requisites are met before the upgrade begins.

For more information about LIC upgrades, see Chapter 13, "Licensed machine code" on page 409.

8.3.3 Time synchronization

With the DS8880, the MC can use the NTP service. Clients can specify NTP servers on their internal or external network to provide the time to the MC. It is a client responsibility to ensure that the NTP servers are working, stable, and accurate. An IBM SSR enables the MC to use NTP servers (ideally at the time of the initial DS8880 installation). Changes can be made by the client by using the **Change Date and Time** action under MC Management on the MC.

> **Important:** For correct error analysis, it is important that the date and time information is synchronized as much as possible on all components in the DS8880 environment. The components include the DS8880 MC, the attached hosts, Spectrum Control, and DS CLI workstations.

8.3.4 Monitoring DS8880 with the Management Console

A client can receive notifications from the MC through SNMP traps and email messages. Notifications contain information about your storage complex, such as open serviceable events. You can choose one or both of the following notification methods:

► SNMP traps

   For monitoring purposes, the DS8880 uses SNMP traps. An SNMP trap can be sent to a server in the client's environment, perhaps with System Management Software, which handles the trap that is based on the Management Information Base (MIB) that was delivered with the DS8880 software. A MIB that contains all of the traps can be used for integration purposes into System Management Software. The supported traps are described in the documentation that comes with the LIC on the CDs that are provided by the IBM SSR. The IP address to which the traps need to be sent must be configured during initial installation of the DS8880. For more information about the DS8880 and SNMP, see Chapter 14, "Monitoring with Simple Network Management Protocol" on page 419.

► Email

   When you enable email notifications, email messages are sent to all the addresses that are defined on the MC whenever the storage complex encounters a serviceable event or must alert individuals to other information.

   During the planning process, create a list of the individuals who need to be notified.

Additionally, when the DS8880 is attached to a z Systems server, a service information message (SIM) notification occurs automatically and requires no setup. A SIM message is displayed on the operating system console if a serviceable event occurs. These messages are not sent from the MC, but from the DS8880 through the channel connections that run between the server and the DS8880.

SNMP and email notification options for the DS8880 require setup on the MC.

8.3.5 Event Notification through Syslog

To meet ever increasing security requirements, beginning with DS8000 release 8.1, the DS8880 supports security and logging events to be forwarded to a syslog server. Previously only available on the MC, events that are contained in the audit log are forwarded to configured syslog receivers. Up to eight external syslog servers can be configured, with varying ports if required. Events that are forwarded include user login/logout, all commands issued by using the GUI or DSCLI while the user is logged in, and remote access events. Events are sent from Facility 19, and are logged as level 6.

8.3.6 Call home and remote support

The MC uses outbound (call home) and inbound (remote service) support.

Call home is the capability of the MC to contact the IBM Support center to report a serviceable event. Remote support is the capability of IBM support representatives to connect to the MC to perform service tasks remotely. If the IBM Support center is allowed to connect to the MC to perform service tasks remotely based on the setup of the client's environment, an IBM SSR can connect to the MC to perform detailed problem analysis. The IBM SSR can view error logs and problem logs and start trace or memory dump retrievals.

Remote support can be configured for Remote Support Center (RSC), IBM Tivoli Assist On-site (AOS), or embedded AOS. Setup of the remote support environment is performed by the IBM SSR during the initial installation. For more information, see Chapter 15, "Remote support" on page 433.

8.4 Management Console and IPv6

The DS8880 MC can be configured for an IPv6 network. IPv4 also is still supported.

8.4.1 Configuring the Management Console in an IPv6 environment

Usually, the IBM SSR configures the MC during the DS8880 initial installation. Complete the following steps to configure the DS8880 MC client network port for IPv6:

1. Start and log in to the Web UI. For more information, see 8.2.6, "Web-based user interface" on page 208. The MC Welcome window opens, as shown in Figure 8-6.
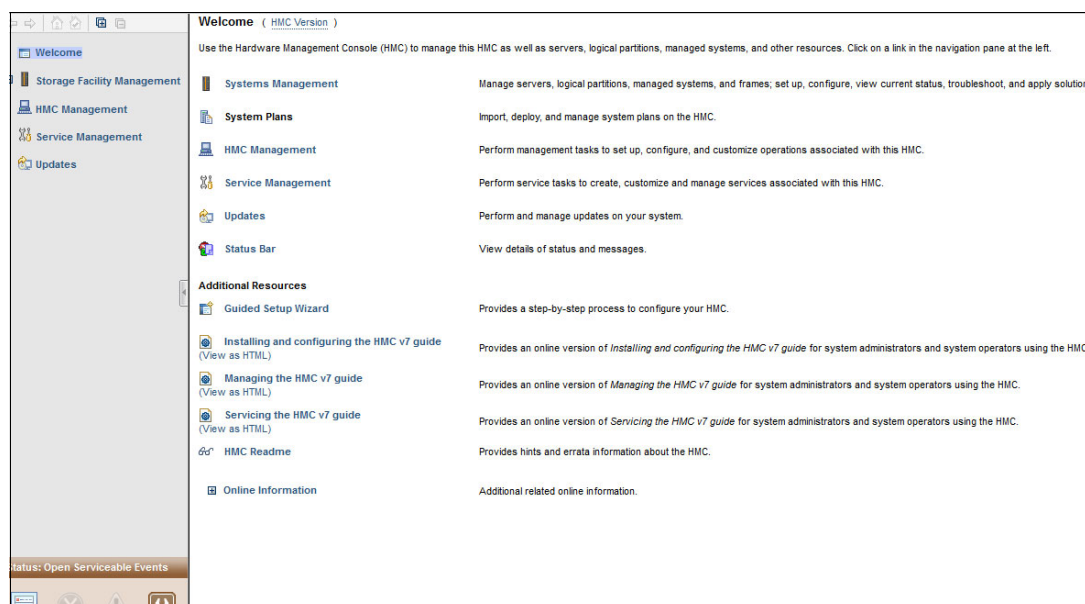

Figure 8-6   Web UI Welcome window

2. In the HMC Management window, select **Change Network Settings** as shown in Figure 8-7.
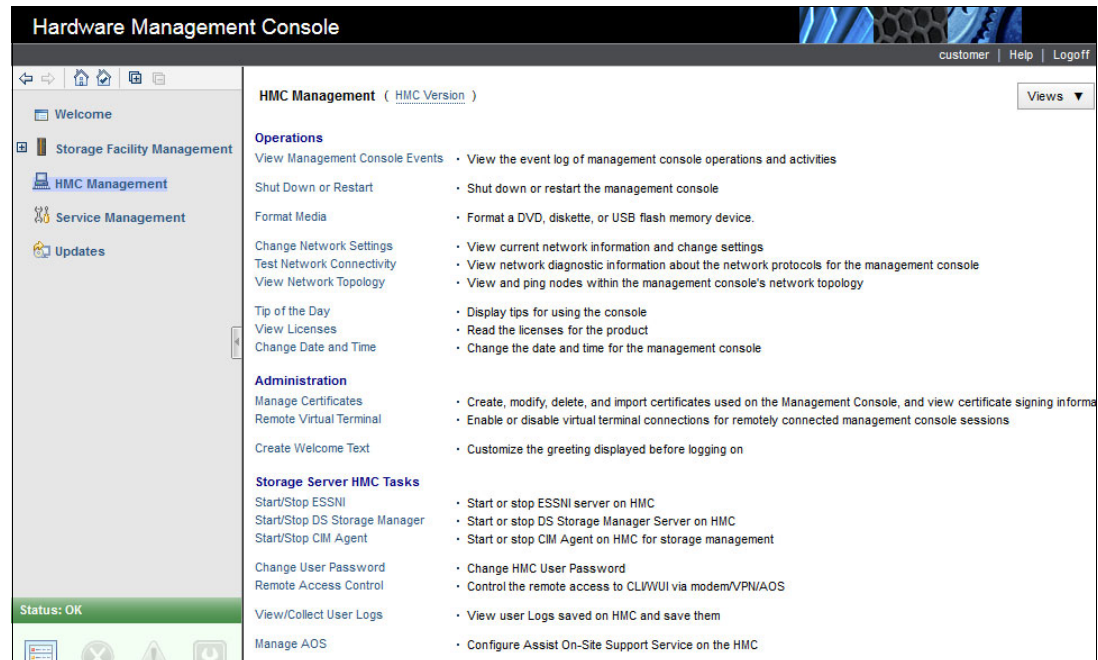
Figure 8-7   Web UI HMC Management window

3. Click the LAN Adapters tab.

4. Only eth2 is shown. The private network ports are not editable. Click **Details**.

5. Click the IPv6 Settings tab.

6. Click **Add** to add a static IP address to this adapter. Figure 8-8 shows the LAN Adapter Details window where you can configure the IPv6 values.

Figure 8-8   Web UI IPv6 settings window

8.5 Management Console user management

User management is performed by using the DS CLI or the DS GUI. An administrator user ID is preconfigured during the installation of the DS8880 and this user ID uses the following defaults:

► User ID: `admin`
► Password: `admin`

The password of the `admin` user ID must be changed before it can be used. The GUI forces you to change the password when you first log in. By using the 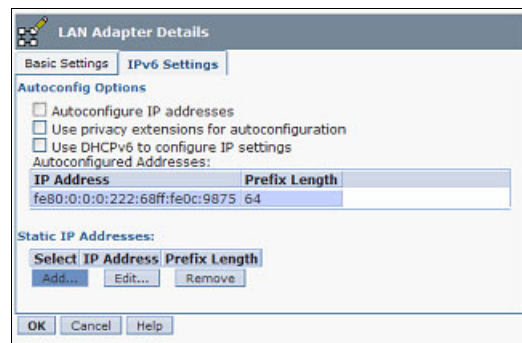DS CLI, you log in but you cannot run any other commands until you change the password. For example, to change the admin user's password to `passw0rd`, use the following DS CLI command:

```
chuser-pw passw0rd admin
```

After you issue that command, you can run other commands.

8.5.1 Password policies

DS8880 supports different role-based users. For more information about user and role management, see 10.2, "User and role management" on page 246. When the administrator adds a user, the administrator enters a password. During the user's first login, this password must be changed. Password settings include the time period (in days) after which passwords expire and a number that identifies how many failed logins are allowed. The user ID is deactivated if an invalid password is entered more times than the limit. Only a user with administrator rights can then reset the user ID with a new initial password.

> **General rule:** Do not set the value of the `chpass` command to 0 because this setting indicates that passwords never expire and unlimited login attempts are allowed.

If access is denied for the administrator because of the number of invalid login attempts, the administrator can use the security recovery utility tool on the Management Console to reset the password to the default value. The detailed procedure is described by selecting **Help Contents** and can be accessed from the DS Management GUI.

> **Important:** Upgrading an existing storage system to the latest code release does not change old default user-acquired rules. Existing default values are retained to prevent disruption. The user might opt to use the new defaults with the `chpass -reset` command. The command resets all default values to the new defaults immediately.

The password for each user account is forced to adhere to the following rules:

► Passwords must contain one character from at least two groups and must be 8 - 16 characters. In addition, the following changes were made:
  – Groups now include alphabetic, numeric, and punctuation.
  – The old rules required at least five alphabetic and one numeric character.
  – The old rules required the first and last characters to be alphabetic.

► Passwords cannot contain the user's ID.

► Initial passwords on new user accounts are expired.

► Passwords that are reset by an administrator are expired.

► Users must change expired passwords at the next logon.

The following password security implementations are included:

► Password rules are checked when passwords are changed.

► The valid character set, embedded user ID, age, length, and history are also checked.

► Passwords that are invalidated by a change remain usable until the next password change.

► Users with invalidated passwords are not automatically disconnected from the DS8880.

► The following password rules are checked when a user logs on:

– Password expiration, locked-out user, and failed attempts are checked.

– Users with passwords that expire or that are locked out by the administrator while they are logged on are not automatically disconnected from the DS8880.

> **Important:** User names and passwords are case-sensitive. For example, if you create a user name that is called `Anthony`, you cannot log in by using the user name `anthony`.

8.5.2 Remote Authentication

Starting with DS8000 version 8.1 you can enable and configure remote authentication through either IBM Spectrum Control (formerly IBM Tivoli Productivity Center) or IBM Copy Services Manager to connect to a Lightweight Directory Access Protocol (LDAP) repository.

With Copy Services Manager now preinstalled on the Management Console, an external proxy is now optional. When remote authentication is enabled, the installation is guided though the Remote Authentication Wizard. Figure 8-9 shows the Welcome page. After you complete all the steps of the wizard, the DS8000 is enabled and configured for remote authentication.
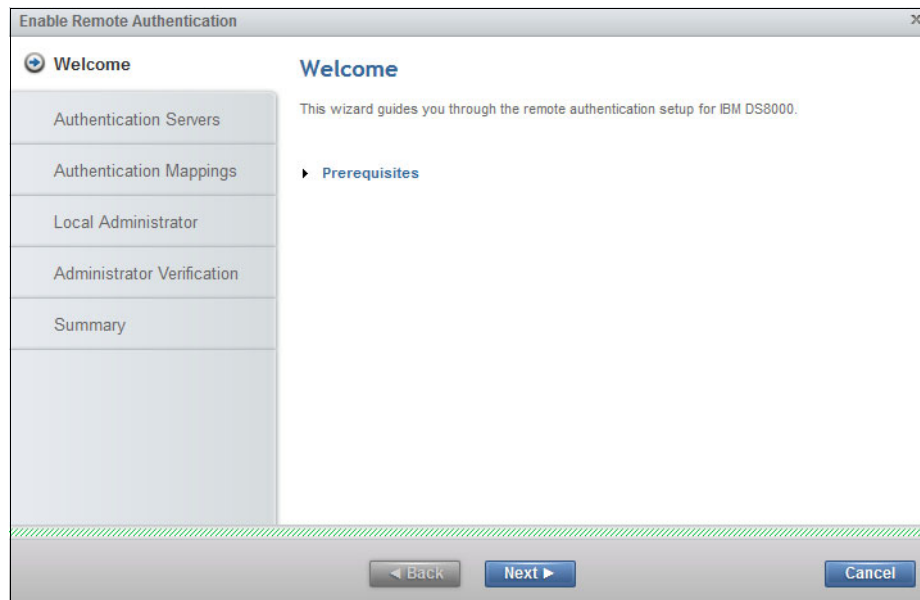
Figure 8-9   Remote Authentication Wizard Welcome page

The following prerequisites are required to complete the Remote Authentication Wizard:

► Access to create users and groups on your remote authentication server.
► A primary LDAP repository URI is required.
► A truststore file with password is required.
► An IBM WebSphere® user name with password is required.
► A secondary LDAP repository URI is optional.

For more information about LDAP-based authentication and configuration, see *IBM DS8880 Integrated Copy Services Manager and LDAP Client on the HMC*, REDP-5356.

8.6 Optional Secondary Management Console

An optional secondary MC (for redundancy) can be ordered for the DS8880. The secondary MC is an optional purchase, but it is highly useful. The primary MC is referred to as $MC1$, and the secondary MC is referred to as $MC2$. The two MCs run in a dual-active configuration, so either MC can be used at any time. Each MC is assigned a role of either primary (normally MC1) or secondary (normally MC2). Certain service functions can be performed only on the primary MC. For this book, the distinction between the primary and secondary MC is only for the purposes of clarity and explanation because they are identical in function.

The DS8880 can run all storage duties while the MC is down or offline, but configuration, error reporting, and maintenance capabilities become severely restricted. Any organization with high availability requirements should consider deploying an MC redundant configuration.

> **Important:** The primary and secondary MCs are not available to be used as general-purpose computing resources.

8.6.1 Management Console redundancy benefits

MC redundancy provides the following advantages:

► Enhanced maintenance capability

Because the MC is the only interface that is available for service personnel, an alternative MC provides maintenance operational capabilities if the internal MC fails.

► Greater availability for power management

The use of the MC is the only way to safely power on or power off the DS8880. A secondary MC is necessary to shut down the DS8880 if the primary MC fails.

► Greater availability of encryption deadlock recovery

If the DS8880 is configured for Full Disk Encryption and an encryption deadlock situation occurs, the use of the MC is the only way to input a Recovery Key to allow the DS8880 to become operational.

► Greater availability for Advanced Copy Services

Because all Copy Services functions are driven by the MC, any environment that uses Advanced Copy Services needs to include dual MCs for operational continuity.

► Greater availability for configuration operations

All configuration commands must go through the MC. This requirement is true regardless of whether access is through Spectrum Control, DS CLI, DS Management GUI, or DS Open API with another management program. A secondary MC allows these operations to continue if the primary MC fails.

When a configuration or Copy Services command is run, the DS CLI or DS Management GUI sends the command to the first MC. If the first MC is unavailable, it automatically sends the command to the second MC instead. Typically, you do not need to reissue the command.

Any changes that are made by using one MC are instantly reflected in the other MC. No host data is cached within the MC, so no cache coherency issues occur.

9

IBM DS8880 features and licensed functions

This chapter describes licensed functions and their activation for the DS8880.

This chapter covers the following topics:

- ► IBM DS8880 licensed functions
- ► Activating licensed functions
- ► Licensed scope considerations

9.1 IBM DS8880 licensed functions

The licensed functions are now bundled into three groups as listed in Table 9-1.

Table 9-1   DS8000 licensed functions

Licensed function for DS8000 with Enterprise Choice warranty	License scope	IBM 283*y*-LF8 indicator feature code numbers
Base Function (BF)	ALL	8151 - 8160
Copy Services (CS)	ALL, FB, or CKD	8250 - 8260
z-synergy Services (zsS)	CKD	8350 - 8360

The grouping of licensed functions facilitates ordering, which differs from earlier DS8000 models for which licensed functions were more granular and were ordered specifically.

The three license bundles contain the following functions:

► Base Function license (BF):

  – Operating Environment License (OEL)
  – Logical Configuration support for Fixed Block (FB) (open systems)
  – Thin Provisioning
  – Easy Tier
  – I/O Priority Manager

► Copy Services license (CS):

  – FlashCopy
  – Metro Mirror
  – Global Mirror
  – Metro/Global Mirror
  – z/Global Mirror
  – z/Global Mirror Resync
  – Multi-Target Peer-to-Peer Remote Copy (PPRC)

► z-synergy Services license (zsS):

  – Fibre Channel connection (FICON) attachment
  – Parallel access volume (PAV)
  – HyperPAV
  – SuperPAV
  – High Performance FICON for z Systems (zHPF)
  – IBM z/OS Distributed Data Backup

The licensed functions are enabled through a 283*x* (*x*=1 - 4) licensed function indicator feature, plus a 283*y*-LF8 (*y*=6 - 9) licensed function authorization feature number:

► The DS8000 provides Enterprise Choice warranty options that are associated with a specific machine type. The *x* in 283*x* designates the machine type according to its warranty period, where *x* can be 1, 2, 3, or 4. For all-flash machines, the machine type is 533*x*, where *x* can be 1 - 4.

  For example, a 2834-985 machine type designates a hybrid DS8886 storage system with a four-year warranty period.

- The *y* in 283*y* can be 6, 7, 8, or 9, according to the associated 2831/2832/2833/2834 base unit model. The 2836 function authorizations apply to 2831 base units, 2837 - 2832, and so on.

  The all-flash systems come with advanced function authorizations 904*y*, *y*=6 - 9.

  For example, 2839-LF8 designates a DS8000 Licensed Function Authorization for a 2834 machine with a four-year warranty period.

- The licensed function indicator feature numbers enable the technical activation of the function, subject to a feature activation code that is made available by IBM and applied by the client. The 283*y*-LF8 (*y*=6 - 9) licensed function authorization feature numbers establish the extent of authorization for that function on the 283*x*-98*z* (*x*=1 - 4, *z*=0 - 8) machine for which it was acquired.

- Licensed functions are activated and enforced with a defined license scope. *License scope* refers to the type of storage and the type of servers, that the function can be used with. For instance, the zsS licenses are only available with the count key data (CKD) (z/FICON) scope.

  The base functions are mandatory. The base functions must always be configured for both mainframe and open systems, which is a scope of `ALL`. Also, to configure CKD volumes, the activation of Feature Code 8300 is required.

  With CS, if these services are only used for either mainframe **or** open systems, the restriction to either FB or CKD is possible. However, most clients will want to configure CS for scope `ALL`.

For each group of licensed functions, specific feature code numbers indicate the licensed capacity, as shown in Table 9-2. Also, see "Ordering granularity" on page 228.

Table 9-2   License feature codes

Feature Code			Feature code for licensed function indicator
BF	CS	zsS	
8151	8251	8351	10 TB (up to 100 TB capacity)
8152	8252	8352	15 TB (from 100.1 TB to 250 TB capacity)
8153	8253	8353	25 TB (from 250.1 TB to 500 TB capacity)
8154	8254	8354	75 TB (from 500.1 to 1,250 TB capacity)
8155	8255	8355	175 TB (from 1,250.1 TB to 3,000 TB capacity)
8156	8256	8356	300 TB (from 3,000.1 TB to 6,000 TB capacity)
8160	8260	8360	500 TB (from 6,000.1 TB to 10,000 TB capacity)

9.1.1 Feature Codes for Copy Services Manager on HMC

With DS8880 and DS8000 microcodes starting with release 8.1, you can activate the Copy Services Manager code integrated on the HMC. See Table 9-3 on page 222.

For Copy Services Manager on HMC entitlement, an order of IBM Copy Services Manager V6 (5725-Z54) is required before you can configure CSM on the HMC. The CSM software is licensed with resource value units (RVUs) according the replicated source capacity managed in total by CSM. Ensure the number of years for software support and maintenance that is purchased is tied to the length of the warranty on the DS8880.

Table 9-3   CSM on HMC

Feature Code	CSM on HMC licensed capacity
8450	Inactive
8451	1 to 100 TB
8452	101 to 250 TB
8453	251 to 500 TB
8454	501 to 1250 TB
8455	1251 to 3000 TB
8456	3001 to 6000 TB
8457	6001 to 10000 TB

9.1.2 Licensing general introduction

Several of the orderable feature codes must be activated through the installation of a corresponding license key bundle. These feature bundle codes are listed in Table 9-1 on page 220. Certain features can be configured directly for the client by the IBM marketing representative during the ordering process.

Features that work with a license key

Important: All Copy Services are now bundled.

The following features are available after the license bundle is activated:

► Metro Mirror (MM) is a synchronous way to perform remote replication. Global Mirror (GM) enables asynchronous replication, which is useful for longer distances and lower bandwidth.

► Metro/Global Mirror (MGM) enables cascaded 3-site replications, which combine synchronous mirroring to an intermediate site with asynchronous mirroring from that intermediate site to a third site at a long distance. Combinations with other Copy Services features are possible, and sometimes even needed. Usually, the 3-site MGM installation also requires an MM sublicense on site A with the MGM license (and even a GM sublicense, if after a site B breakdown, you want to resynchronize between site A and site C). At site B, on top of the MGM, you also need the MM and GM licenses. At site C, you then need sublicenses for MGM, GM, and FlashCopy.

► Multiple-target PPRC (MT-PPRC) enhances the disaster recovery solutions by allowing data at a single primary site to be mirrored to two remote sites simultaneously. The function builds and extends Metro Mirror and Global Mirror capabilities and is supported on DS8880, on newer DS8870 firmware, and also on Copy Services Manager (CSM) or IBM z Systems software, such as IBM Geographically Dispersed Parallel Sysplex (IBM GDPS)/Multiple-target Metro Mirror (MTMM). Various interfaces and operating systems support the function. With DS8870, this feature was a separate chargeable feature code for each machine. For the DS8880 family, this feature is integrated with the CS license bundle.

► Two possibilities exist for FlashCopy (point in time copy (PTC)): Use it with thick (standard) volumes or thin-provisioned extent space-efficient (ESE) volumes.

The ESE thin volumes can also be used in remote-mirroring relationships. ESE volumes offer the same good performance as standard (thick) volumes, and can be managed by IBM Easy Tier.

► The z/OS Global Mirror (zGM) license, which is also known as Extended Remote Copy (XRC), enables z/OS clients to copy data by using System Data Mover (SDM). This is an asynchronous copy. For more information, see 5.3.7, "z/OS Global Mirror" on page 142.

► Copy Services Manager on HMC offers full CSM functionality and must be enabled through DFSA activation key. CSM enablement files are activated on the HMC when the key is applied.

► For z Systems clients, PAVs allow multiple concurrent I/O streams to the same CKD volume.

HyperPAV reassigns the alias addresses dynamically to the base addresses of the volumes based on the needs of a dynamically changing workload. Both features result in such large performance gains that for many years they were configured as an effective standard for mainframe clients, similar to FICON, which is required for z/OS.

SuperPAV is an extension to HyperPAV support and allows aliases to be borrowed from eligible peer logical control units.

► zHPF is a feature that uses a protocol extension for FICON and allows data for multiple commands to be grouped in a single data transfer. This grouping increases the channel throughput for many workload profiles. It works on all newer IBM zEnterprise Systems and it is preferred for these systems because of the performance gains that it offers.

► z/OS Distributed Data Backup (zDDB) is a feature for clients with a mixture of mainframe and distributed workloads to use their powerful z Systems host facilities to back up and restore open systems data. For more information, see *IBM System Storage DS8000: z/OS Distributed Data Backup*, REDP-4701.

► Easy Tier is available in the following modes:

 – Automatic mode works on the subvolume level (extent level) and allows auto-tiering in hybrid extent pools. The most-accessed volume parts go to the upper tiers. In single-tier pools, it allows auto-rebalancing if it is turned on.

 – Manual dynamic volume relocation mode works on the level of full volumes and allows volumes to be relocated or restriped to other places in the DS8000 online. It also allows ranks to be moved out of pools. For more information, see *IBM DS8000 EasyTier*, REDP-4667.

As part of Easy Tier, the following functions can be implemented:

 – Easy Tier Application provides an extra application programming interface (API) for software developers to use so that applications direct Easy Tier data placement on the DS8000. This API also enables clients to assign ("pin") volumes to a particular tier within an Easy Tier pool to meet performance and cost requirements. For more information, see *DS8870 Easy Tier Application*, REDP-5014.

 – Easy Tier Heat Map Transfer automatically replicates a heat map to remote systems to ensure that they are also optimized for performance and cost after a planned or unplanned outage. For more information, see *IBM DS8870 Easy Tier Heat Map Transfer*, REDP-5015.

► I/O Priority Manager is the quality of service (QoS) feature for the System Storage DS8000 series. When larger extent pools are used that include many servers that are competing for the same rank and device adapter resources, clients can define Performance Groups of higher-priority and lower-priority servers and volumes. In overload conditions, the I/O Priority Manager throttles the lower-priority Performance Groups to maintain service on the higher-priority groups. With DS8870 and earlier models, the I/O Priority Manager function was an extra charge by TB. The I/O Priority Manager function is now included in the Base Function so that it is available to all DS8880 clients. For more information, see *DS8000 I/O Priority Manager*, REDP-4760.

► IBM Database Protection provides the highest level of protection for Oracle databases by using more end-to-end checks for the detection of data corruption as data passes through different storage area network (SAN) and storage hardware layers. This feature complies with the Oracle Hardware Assisted Resilient® Data (HARD) initiative. For more information about this feature, see the *IBM Database Protection User's Guide*, GC27-2133-02.

Ordering feature options without needing a license key

The following ordering options of the DS8880 do not require the client to install a license key:

► Earthquake Resistance Kit, Feature Code (FC) 1906: The Earthquake Resistance Kit is an optional seismic kit for stabilizing the storage unit racks so that the racks comply with IBM earthquake resistance standards. It includes cross braces on the front and rear of the racks, and the racks are secured to the floor. These stabilizing features limit potential damage to critical DS8000 machine components and help to prevent human injury.

► Overhead cabling: For more information about FC1400 (top-exit bracket) and FC1101 (ladder), see 7.2.3, "Overhead cabling features" on page 184. One ladder for each site is sufficient. A smaller ladder (FC1102) is offered when you use the DS8886 and DS8888 option to expand the rack height to full length.

► Shipping Weight Reduction (FC0200): If your site features delivery weight constraints, IBM offers this option that limits the maximum shipping weight of the individually packed components to 909 kg (2,000 lb). Because this feature increases installation time, order it only when required.

► Extended Powerline Disturbance Feature (FC1055): This feature extends the available uptime in case both power cords lose external power, as described in "Extended power line disturbance feature" on page 188.

► Storage Appliance indicator (FC0952): The feature is for DS8880 models that are meant to be actively used with Full Disk Encryption (FDE). It allows a Storage Appliance (2421-AP1) pair (FC1762) or single Storage Appliance (FC1761) with Security Key Lifecycle Manager (SKLM) software to be ordered and preinstalled on a small Linux server. This server pair can then manage the encryption key with SKLM. For more information about disk encryption, see *IBM DS8880 Data-at-rest Encryption*, REDP-4500.

► Epic (FC0964), VMware VAAI (FC0965): For clients who want to use the Epic healthcare software or VMware vStorage APIs for Array Integration (VAAI), these features must be selected by the IBM marketing representative. For the VAAI XCOPY/Clone primitive, the PTC (FlashCopy) license (in the CS bundle) is also needed.

► IBM ProtecTIER® indicator (FC0960), IBM SAN Volume Controller (FC0963), IBM Storwize® V7000 virtualization (FC0961), N series Gateway (FC0962) indicator: If the DS8880 is used or virtualized behind any of these data deduplication or virtualization devices or an NAS gateway, select the correct feature to indicate this device.

For more information about these features, see the *IBM DS8880 Introduction and Planning Guide*, GC27-8525.

9.1.3 Licensing cost structure

For the three possible license bundles, the following general guidelines apply:

► Base Function (BF) license

The license needs to be equal to or greater than the total physical capacity of the storage system. Select the full raw capacity and order the number of TBs to support the total physical capacity of your storage system.

► Copy Services (CS) license

The license can be equal to, less than, or more than the total physical capacity of the storage system. You need to license the TBs that you actually use in copy relationships.

► z-synergy Services (zsS) license

The license can be equal to or less than the total physical capacity of the storage system, but order less only if you have a mixed machine FB/Open and CKD/mainframe. You need to license the full CKD/z raw capacity for zsS, that is, the physical capacity of all ranks that will be formatted as CKD ranks.

Copy Services license specifics

You can order the CS license to support the *total usable capacity* of all volumes that are involved in one or more copy services functions. However, this sort of subcapacity licensing (less than the total physical raw capacity) requires capacity monitoring and a steady remote connection on the client side.

> **Important:** With the CS license bundle, only order subcapacity licensing, which is less than the total physical raw capacity, when a steady remote connection for the DS8000 is available.

By using a remote connection for call home, the CS license can be based on the usable capacity of the volumes that will potentially be in CS relationships. This amount typically is less than the total physical capacity.

> **Note:** The Copy Services license goes by the usable capacity of all volumes that are involved in at least one Copy Services relationship. If overprovisioning is used on the DS8880 with a significant amount of Copy Services functionality, the Copy Services license needs to be equal to the total rank usable capacity only. This is true even if the logical volume capacity of volumes in Copy Services is greater.

For example, with overprovisioning, if the total rank usable capacity of a DS8880 is 100 TB but 200 TB of thin-provisioning volumes are in Metro Mirror, only a 100 TB of Copy Services license is needed.

For FlashCopy volumes, you need to count the source plus target volumes as usable capacity. Several examples are shown.

Pricing examples for Copy Services

The following examples are provided to illustrate your CS licensing requirements in an FB environment:

► Scenario 1: For FlashCopy of a 10 TB source, the purchase of 20 TBs of capacity of CS license is required.

► Scenario 2: To use Metro Mirror on a source of 10 TB and then a single FlashCopy on the target of 10 TB, the purchase of 10 TBs of CS license on the source and the purchase of 20 TBs of CS on the target DS8000 is required.

► Scenario 3: To use Global Mirror on a source of 10 TB and then a single FlashCopy on the target DS8000 of 10 TB, the purchase of 10 TBs of CS license on the source and the purchase of 20 TBs of CS on the target DS8000 is required.

► Scenario 4: To use MGM on a source of 10 TB and then a single FlashCopy on the target of 10 TB, the purchase of 10 TB of CS license on the source and secondary, and the purchase of 20 TB of CS on the target is required.

  However, consider that with MGM, certain scenarios can require more FlashCopy targets on the local machines, and so larger CS TB scopes are necessary.

► Scenario 5: A client wants to perform Global Mirror for 10 TB and perform a FlashCopy on the target for practicing disaster recovery, but not affect the normal GM. This situation requires a GM secondary, GM Journal, and FlashCopy volume on the secondary system. The source DS8880 requires 10 TB and the target DS8880 requires a 30 TB Copy Services license.

► Scenario 6: For zGM/XRC 10 TB, the purchase of 10 TB and 10 TB of capacity for the CS license on both the source and the target is required.

► Scenario 7: To perform 4-site replications, the purchase of the correct capacity license requirement for each storage system is required.

z-synergy Services (zsS) licensing

A z-synergy Services license is usually required for only the total physical capacity that is logically configured as CKD ranks for use with z Systems host systems.

> **Note:** If z/OS Distributed Data Backup is used on a system with no CKD ranks, a 10 TB z-synergy Services license must be ordered to enable the FICON attachment functionality.

Drive features

The BF is based on the raw (decimal TB) capacity of the drives. The pricing is based on the disk drive performance, capacity, speed, and other characteristics that provide more flexible and optimal price and performance configurations. Table 9-4, Table 9-5, and Table 9-6 on page 227 provide an overview of the available hard disk drives (HDDs) and flash drives.

To calculate the raw (gross) physical capacity, multiply for each drive set the number of drives with their individual capacities. Therefore, for example, a drive set of sixteen 1.2 TB drives has a 19.2 TB raw capacity.

Table 9-4 shows the feature codes for disk drive sets.

Table 9-4   Feature codes for disk drive sets

Drive set feature number	Drive size	Drive type	Drive speed	Encryption capable[a]	Drives per set
5308	300 GB	2.5-inch disk drives	15K RPM	Yes	16
5618	600 GB	2.5-inch disk drives	15K RPM	Yes	16
5708	600 GB	2.5-inch disk drives	10K RPM	Yes	16
5768	1.2 TB	2.5-inch disk drives	10K RPM	Yes	16
5778	1.8 TB	2.5-inch disk drives	10K RPM	Yes	16
5868	4 TB	3.5-inch disk drives	7.2K RPM	Yes	Half: 8
5878	6 TB	3.5-inch disk drives	7.2K RPM	Yes	Half: 8

a. Drives are FDE self-encrypting drive (SED)-capable.

Table 9-5 shows the feature codes for flash drive sets.

Table 9-5   Feature codes for flash drive sets

Drive set feature number	Drive size	Drive type	Drive speed	Encryption capable	Drives per set
6158	400 GB	2.5-inch flash drives	N/A	Yes	16
6258	800 GB	2.5-inch flash drives	N/A	Yes	16
6358	1.6 TB	2.5-inch flash drives	N/A	Yes	16

Table 9-6 shows the feature codes for flash card sets.

Table 9-6   Feature codes for flash card sets

Feature code	Drive capacity	Drive type	Drive speed	Encryption capable	Drives per set
1610	400 GB	2.5-inch flash card	N/A	Yes	16
1611	800 GB	2.5-inch flash card	N/A	Yes	16
1612	1.6 TB	2.5-inch flash card	N/A	Yes	16
1613	3.2 TB	2.5-inch flash card	N/A	Yes	16

Important: Check with an IBM representative or consult the IBM website for an up-to-date list of available drive types.

Related information: New storage system expansions for DS8880 are delivered only with FDE drives. Existing DS8870 disks are supported only in a DS8870-to-DS8886 model conversion.

IBM Easy Tier is a license feature that is available at no charge. Therefore, it is configured, by default, with the BF license bundle.

The Database Protection (for Open/FB) feature, the Thin Provisioning feature, and the I/O Priority Manager feature come with the BF license bundle.

The IBM z/OS Distributed Data Backup feature comes with the zsS license bundle.

Ordering granularity

You order the license bundles by TB, but not by single TB. The granularity is slightly larger. For instance, below 100 TB total physical capacity, the granularity increment for an upgrade is 10 TB. With larger total capacities, the granularity is larger. See Table 9-7.

Table 9-7   Ordering granularity

Tier	Minimum (TB)	Maximum (TB)	Tier TB granularity	Feature quantity (maximum)	Range delta (TB)	Feature number
1	1	100	10	10	100	8x51
2	101	250	15	10	150	8x52
3	251	500	25	10	250	8x53
4	501	1250	75	10	750	8x54
5	1251	3000	175	10	1750	8x55
6	3001	6000	300	10	3000	8x56
7	6001	10000	500	8	4000	8x60

Tip: For more information about the features and considerations when you order DS8880 licensed functions, see the following announcement letters:

► *IBM DS8880 (IBM 283x)*
► *IBM DS8880 (M/T 283x) high-performance flagship - Function Authorizations*

IBM announcement letters are available at this website:

http://www.ibm.com/common/ssi/index.wss

Use the DS8880 keyword as a search criterion in the **Contents** field.

9.2 Activating licensed functions

You can activate the license keys of the DS8000 after the IBM SSR completes the storage complex installation. If you plan to use the Storage Management GUI to configure your new storage, after the initial login as admin, the setup wizard will guide you to download your keys from the IBM Disk storage feature activation (DSFA) website and activate them. However, if you plan to use the data storage command-line interface (DS CLI) to configure your new storage, you must first obtain the necessary keys from the IBM DSFA website at this location:

http://www.ibm.com/storage/dsfa

Before you connect to the IBM DSFA website to obtain your feature activation codes, ensure that you have the following items:

► The IBM License Function Authorization documents. If you are activating codes for a new storage unit, these documents are included in the shipment of the storage unit. If you are activating codes for an existing storage unit, IBM sends the documents to you in an envelope.

► A USB memory device can be used for downloading your activation codes if you cannot access the DS Storage Manager from the system that you are using to access the DSFA website. Instead of downloading the activation codes in softcopy format, you can print the activation codes and manually enter them by using the DS Storage Manager GUI or the DS CLI. However, this process is slow and error-prone because the activation keys are 32-character strings.

9.2.1 Obtaining DS8000 machine information and activating license keys

To obtain the license activation keys from the DSFA website, you need to know the serial number and machine signature of your DS8000 unit.

You can obtain the required information by using the DS Storage Management GUI or the DS CLI. If you use the Storage Management GUI, you can obtain and apply your activation keys at the same time. These options are described next.

DS Storage Management GUI

Complete these steps to obtain the required information by using the Management GUI:

1. Open a browser and enter `https://< IP address of HMC >`.

2. Log in by using a user ID with administrator access. If you are accessing the system for the first time, contact your IBM service support representative (SSR) for the user ID and password. After a successful login, the system monitor window opens.

3. If this machine is new, a System Setup wizard window opens automatically that guides you through the initial setup and configuration tasks, as shown in Figure 9-1.
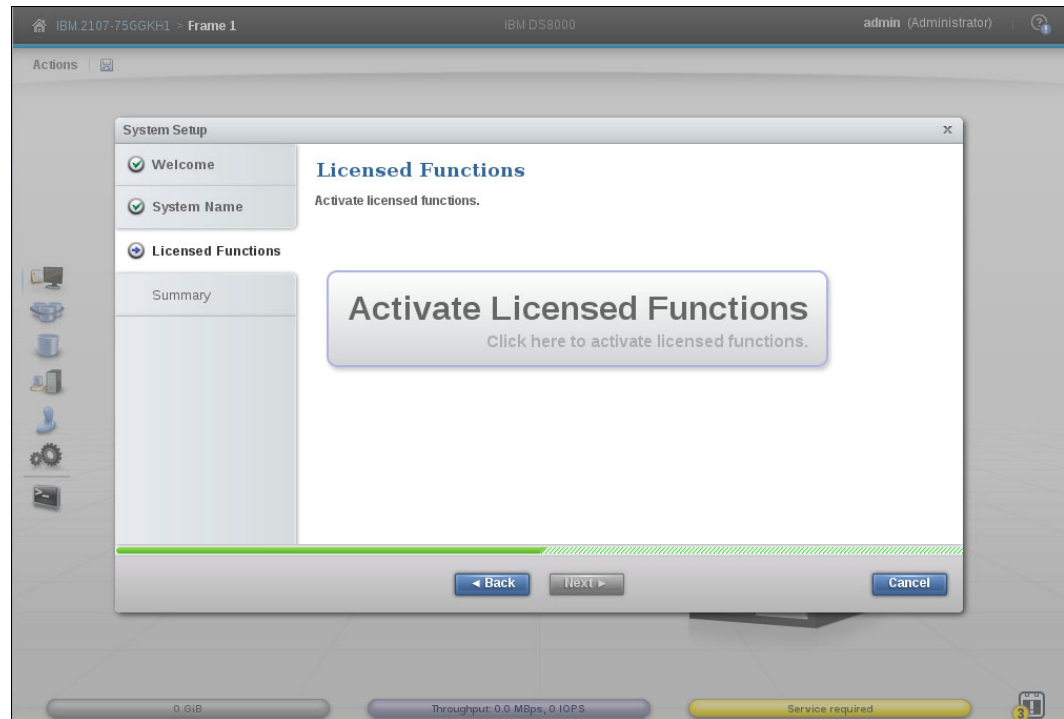

Figure 9-1   Activate Licensed Functions

Note: Before you begin this task, resolve any current DS8000 problems that exist. You can contact IBM Support to help you resolve these problems.

4. Click **Activate Licensed Functions** (on Figure 9-1) to begin the guided procedure to acquire and activate your feature activation keys, as shown in Figure 9-2.
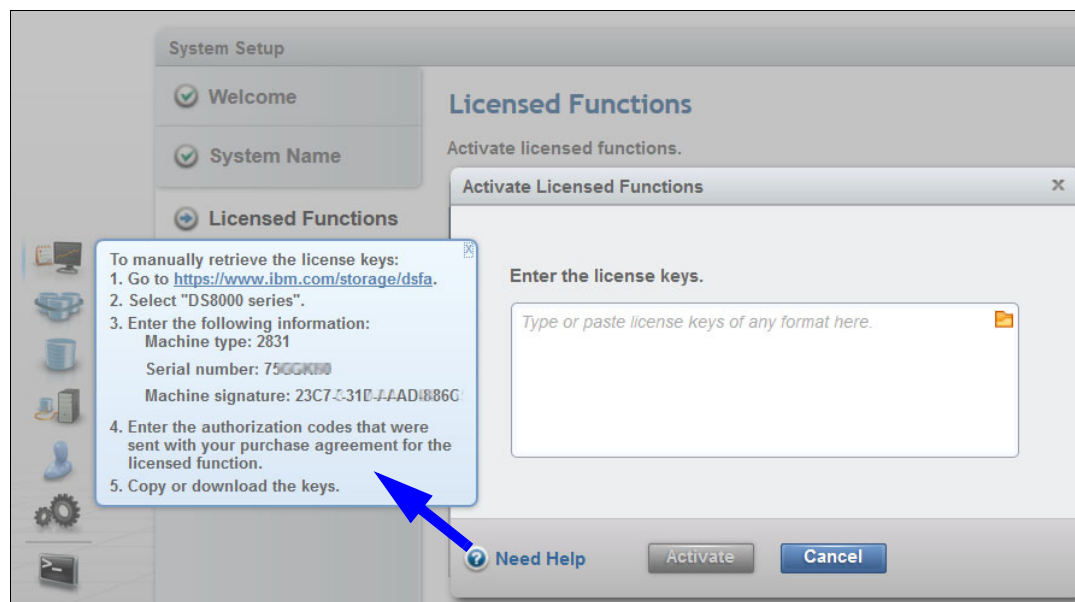
Figure 9-2   Enter the license keys with help

Note: You can download the keys and save the XML file to the folder that is shown here, or you can copy the license keys from the IBM DSFA website.

5. After you enter all of your license keys, click **Activate** to start the activation process, as shown in Figure 9-3.
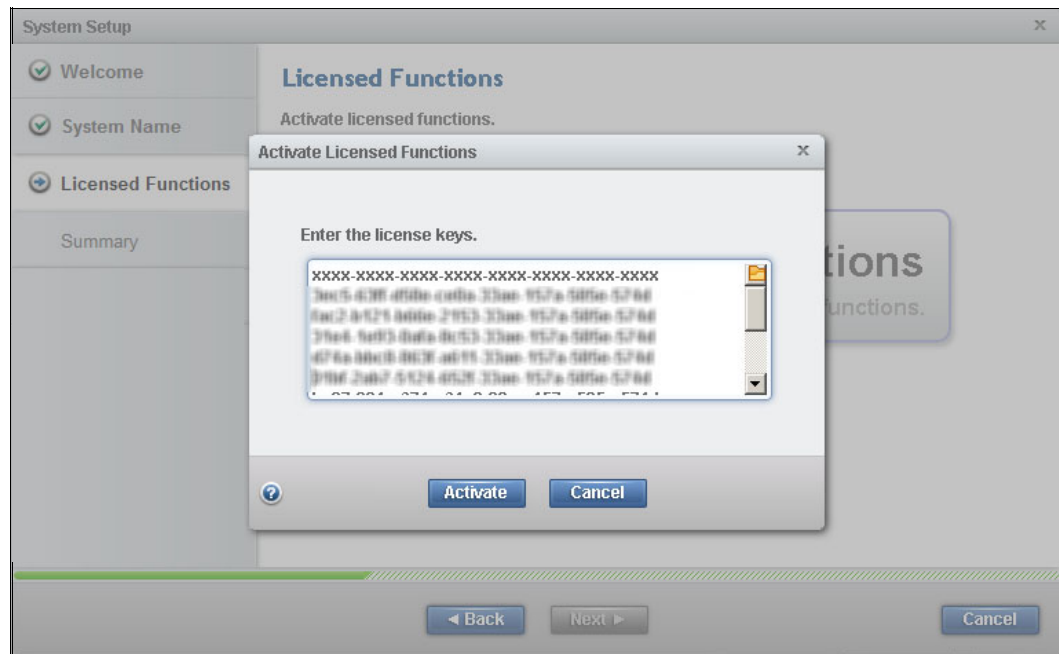

Figure 9-3   Add the license keys

6. Click **Summary** in the System Setup wizard to view the list of licensed functions or feature keys that are installed on your DS8000, as shown in Figure 9-4.
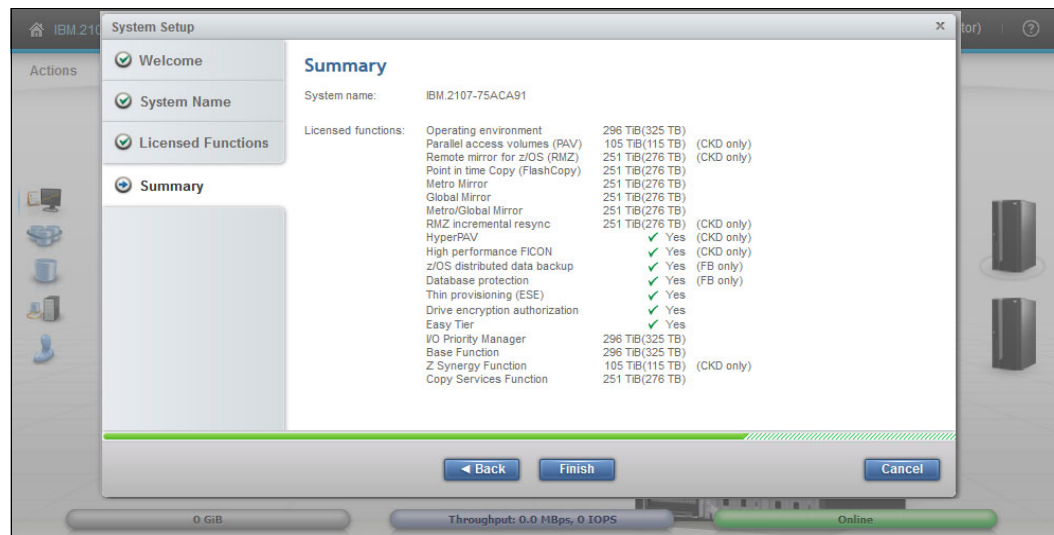

Figure 9-4   Summary of licensed functions

7. If you must activate additional feature keys after the initial installation, click the **System** icon in the Settings monitor window and then select **Licensed Functions**, as shown in Figure 9-5.

Figure 9-5   Licensed Functions

8. To obtain the machine signature and machine type and model (MTM) after the installation, click **Actions** → **Properties**, as shown in Figure 9-6.
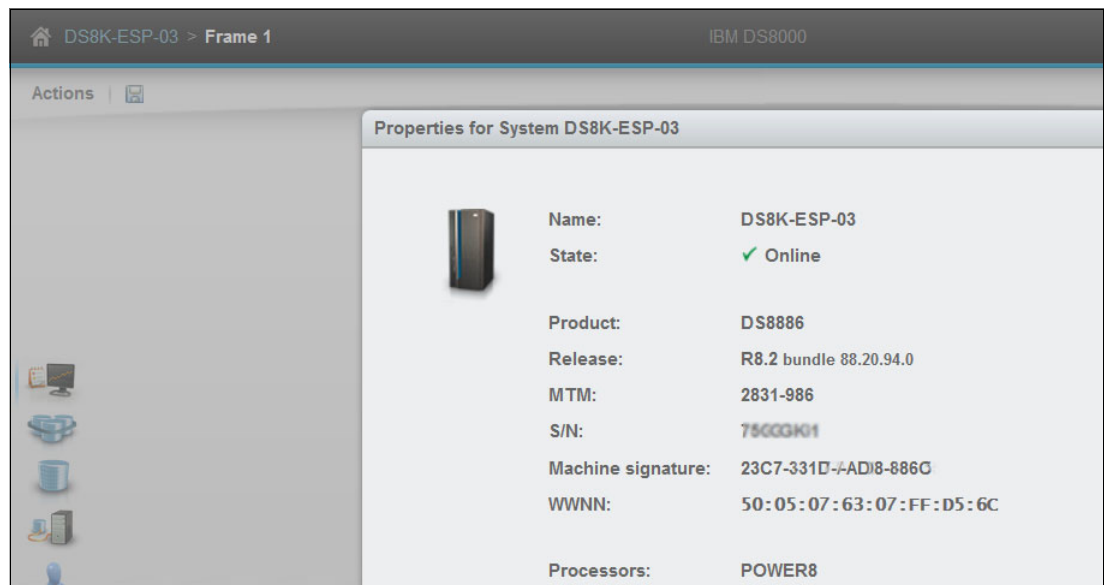

Figure 9-6   Properties window showing the machine signature and MTM

Important: The initial enablement of any optional DS8000 licensed function is a concurrent activity (assuming that the correct level of Licensed Internal Code (LIC) is installed on the system for the function).

The following activation activities are disruptive and require an initial machine load (IML) or restart of the affected image:

► Removal of a DS8000 licensed function to deactivate the function.

► A lateral change or reduction in the license scope. A *lateral change* is defined as changing the license scope from FB to CKD or from CKD to FB. A *reduction* is defined as changing the license scope from all physical capacity (ALL) to only FB or only CKD capacity.

Note: Before you begin this task, you must resolve any current DS8000 problems that exist. You can contact IBM Support for help to resolve these problems.

9. Click **Activate** to enter and activate your licensed keys, as shown in Figure 9-2 on page 230.

10. Wait for the activation process to complete and select **Licensed Functions** to display the list of activated features, as illustrated in Figure 9-7.

Name	Enabled	Capacity	Storage Type
Operating environment	✔ Yes	91 TiB(100 TB)	All
Parallel access volumes (PAV)	✔ Yes	69 TiB(76 TB)	CKD
Remote mirror for z/OS (RMZ)	✔ Yes	91 TiB(100 TB)	CKD
Point in time Copy (FlashCopy)	✔ Yes	91 TiB(100 TB)	All
Metro Mirror	✔ Yes	91 TiB(100 TB)	All
Global Mirror	✔ Yes	91 TiB(100 TB)	All
Metro/Global Mirror	✔ Yes	91 TiB(100 TB)	All
RMZ incremental resync	✔ Yes	91 TiB(100 TB)	CKD
HyperPAV	✔ Yes		CKD
High performance FICON	✔ Yes		CKD
z/OS distributed data backup	✔ Yes		FB
Database protection	✔ Yes		FB
Thin provisioning (ESE)	✔ Yes		All
Drive encryption authorization	✔ Yes		All
Easy Tier	✔ Yes		All
I/O Priority Manager	✔ Yes	91 TiB(100 TB)	All

Figure 9-7   List of licensed functions

DS command-line interface

To obtain the required information by using the DS CLI, log on to the DS CLI and run the `lssi` and `showsi` commands as shown in Figure 9-8.

```
dscli> lssi
Name             ID               Storage Unit    Model WWNN            State  ESSNet
==============================================================================================
IBM.2107-75XXXX1 IBM.2107-75XXXX1 IBM.2107-75XXXX0 984    50050763YYYYYYYY Online Enabled

dscli> showsi
Name            IBM.2107-75XXXX1
desc            Caribbean Shark BC
ID              IBM.2107-75XXXX1
Storage Unit    IBM.2107-75XXXX0
Model           984
WWNN            5005076303FFD13E
Signature       1ef8-90ab-cdef-0123
State           Online
ESSNet          Enabled
Volume Group    V0
os400Serial     13E
NVS Memory      16.0 GB
Cache Memory    217.2 GB
Processor Memory 251.6 GB
MTS             IBM.2831-75XXXX0
numegsupported  1
ETAutoMode      all
ETMonitor       all
IOPMmode        Managed
ETCCMode        -
ETHMTMode       Enabled
ETSRMode        Enabled
ETTierOrder     High performance
```

Figure 9-8   Obtain DS8000 information by using the DS CLI

Note: The `showsi` command can take the storage facility image (SFI) serial number as a possible argument. The SFI serial number is identical to the storage unit serial number, except that the SFI serial number ends in 1 instead of 0 (zero).

Gather the following information about your storage unit:

► The Machine Type - Serial Number (MTS), which is a string that contains the machine type and the serial number. The machine type is $283x$ (or $242x$, for earlier DS8000 models) and the last seven characters of the string are the machine's serial number (XYABCDE), which always ends with 0 (zero).

► The model, which, for example, is 986 for a DS8886.

► The machine signature, which is found in the Machine signature field and uses the following format: ABCD-EF12-3456-7890.

Use Table 9-8 to document this information, which is entered in the IBM DSFA website to retrieve the activation codes.

Table 9-8   DS8000 machine information

Property	Your storage unit's information
Machine type	
Machine's serial number	
Machine signature	

9.2.2 Obtaining the activation codes

If you plan to use the DS CLI to configure your system, you need to obtain your activation keys before you configure your machine.

> **Note:** A DS8800 is shown in the following figures. However, the steps are identical for all models of the DS8000 family.

Complete the following steps to obtain the activation codes:

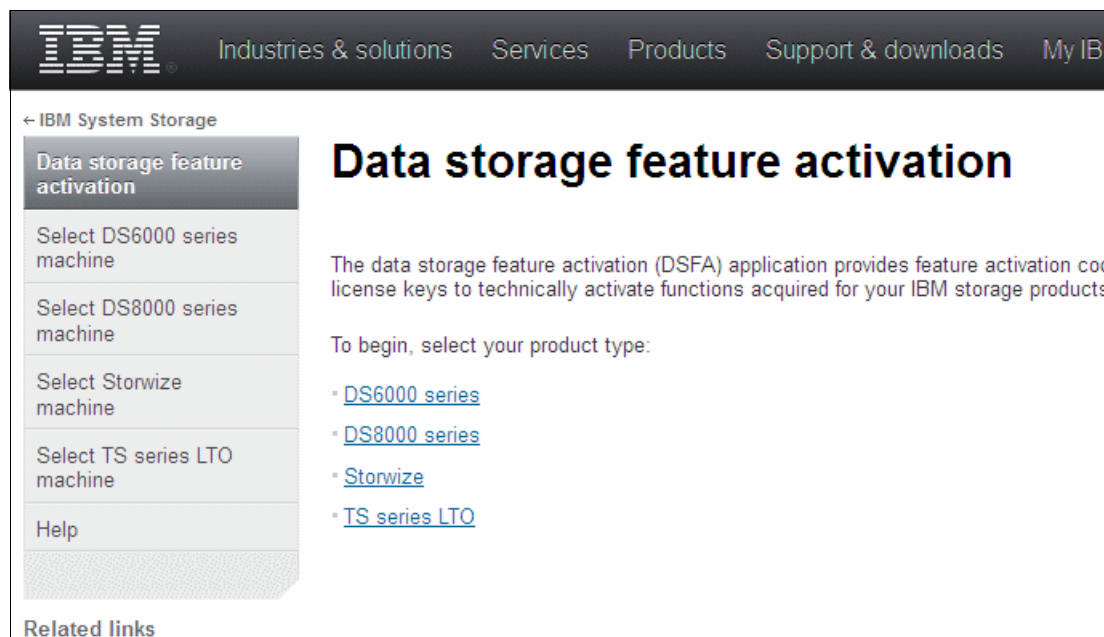1. As shown in Figure 9-9, connect to the DSFA website at the following address:

   http://www.ibm.com/storage/dsfa

Figure 9-9   IBM DSFA website

2. Click **DS8000 series**. The Select DS8000 series machine window opens, as shown in Figure 9-10. Select the appropriate 283*x* (or 242*x*, 533*x*, 958*x*) machine type.

Figure 9-10   DS8000 DSFA machine information entry window

3. Enter the machine information that was collected in Table 9-8 on page 235, and click **Submit**. The View machine summary window opens, as shown in Figure 9-11.

Figure 9-11   DSFA View machine summary window

Figure 9-11 shows the View machine summary window, which shows the total purchased licenses and the number of licenses that are currently assigned. When you assign licenses for the first time, the Assigned field shows `0.0 TB`.

4. On the left, click **Manage activations**. The Manage activations window opens, as shown in Figure 9-12. For each license type and storage image, enter the following information that is assigned to the storage image:

   – Select the license scope from the list box:

      • FB (fixed-block data)
      • CKD (count key data)
      • All

   – Type the capacity value (in TB) to assign to the storage image

   The capacity values are expressed in decimal terabytes. The sum of the storage image capacity values for a license cannot exceed the total license value.

Figure 9-12   DSFA Manage activations window

5. After the values are entered, click **Submit**.

6. On the left on Figure 9-12, select **Retrieve activation codes**.

7. The Retrieve activation codes window opens, which shows the license activation codes for the storage image, as shown in Figure 9-13. Print the activation codes or click **Download now** to save the activation codes in an XML file that you can import into the DS8000.
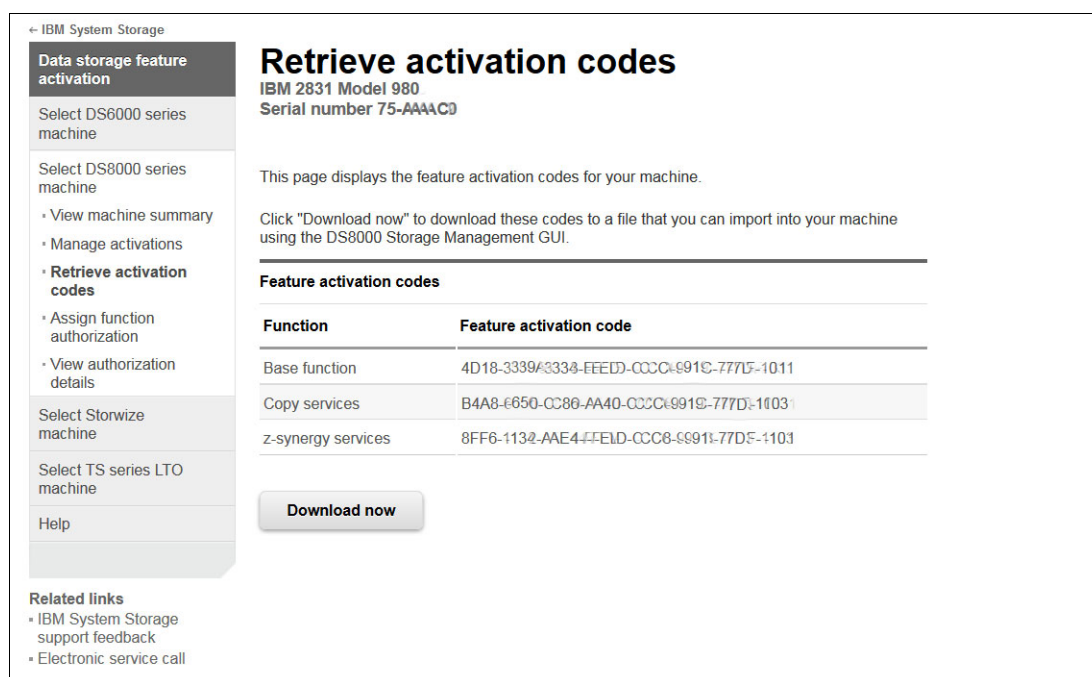
Figure 9-13   DSFA Retrieve activation codes window

8. Click **Settings** → **System** → **Licensed Functions** → **Activate** in the GUI to enter the activation codes after the initial installation. Click the yellow symbol to import the XML file, or alternatively, you can enter individual license keys, as shown in Figure 9-14.
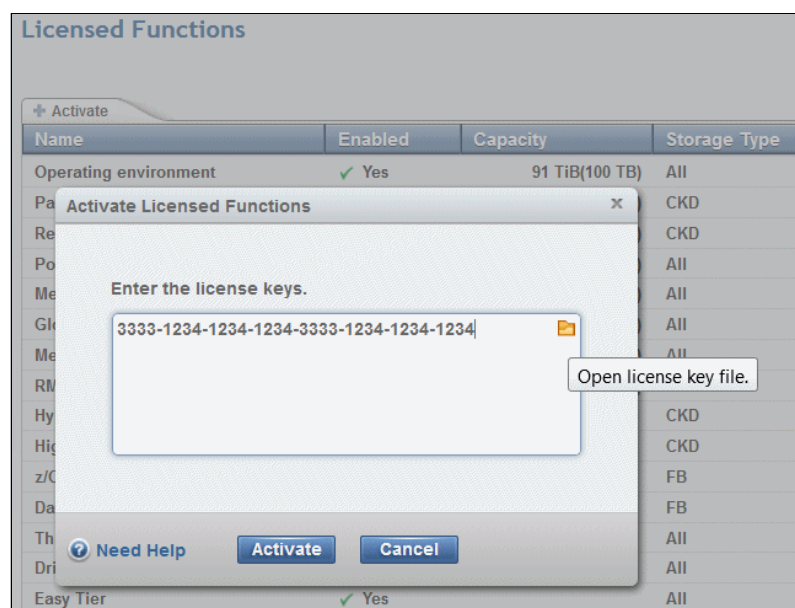
Figure 9-14   Activating additional licenses in the GUI

> **Important:** In most situations, the DSFA application can locate your 283*y* (*y*=6 - 9) licensed function authorization record when you enter the DS8880 283*x* (*x*=1 - 4) serial number and signature. However, if the 283*y* licensed function authorization record is not attached to the 283*x* record, you must assign it to the 283*x* record by using the **Assign function authorization** link on the DSFA application. In this case, you need the 283*y* serial number (which you can find on the License Function Authorization document).

9.2.3 Applying activation codes by using the DS CLI

The license keys also can be activated by using the DS CLI. This option is available only if the machine OEL was activated and you have installed a compatible DS CLI program on your console.

Complete the following steps to apply activation codes by using the DS CLI:

1. Run the `showsi` command to display the DS8000 machine signature, as shown in Figure 9-15.

```
dscli> showsi
Name            DS8K-R8-01
desc            Caribbean Shark EC
ID              IBM.2107-75XXXX1
Storage Unit    IBM.2107-75XXXX0
Model           986
WWNN            50050763YYYYYYYY
Signature       bcde-f012-3456-789a
State           Online
ESSNet          Enabled
Volume Group    V0
os400Serial     693
NVS Memory      32.0 GB
Cache Memory    451.0 GB
Processor Memory 506.4 GB
MTS             IBM.2831-75XXXX0
numegsupported  1
ETAutoMode      tiered
ETMonitor       all
IOPMmode        Monitored
ETCCMode        -
ETHMTMode       Enabled
ETSRMode        Enabled
ETTierOrder     High utilization
```

Figure 9-15   DS CLI showsi command

2. Obtain your license activation codes from the IBM DSFA website, as described in 9.2.2, "Obtaining the activation codes" on page 235.

3. Enter an `applykey` command at the following `dscli` command prompt. The `-file` parameter specifies the key file. The second parameter specifies the storage image.

   ```
   dscli> applykey -file c:\2831_75XXXX0.xml IBM.2107-75XXXX1
   ```

Or you can apply individual keys:

```
dscli> applykey -key f190-1234-1234-1234-1234-5678-1234-5678 IBM.2107-75ZA571
CMUC00199I applykey: License Machine Code successfully applied to storage image
IBM.2107-75ZA571.
```

4. Verify that the keys were activated for your storage unit by running the **lskey** command, as shown in Figure 9-16.

```
dscli> lskey
Activation Key                                 Authorization Level (TB) Scope
============================================================================
Base function                                  325                      All
CSM on HMC                                      325                      All
Copy services                                   275                      All
Encryption Authorization                        on                       All
Global mirror (GM)                              275                      All
High Performance FICON for System z (zHPF)     on                       CKD
I/O Priority Manager                            325                      All
IBM HyperPAV                                     on                       CKD
IBM System Storage DS8000 Thin Provisioning     on                       All
IBM System Storage Easy Tier                    on                       All
IBM database protection                         on                       FB
IBM z/OS Distributed Data Backup                on                       FB
Metro/Global mirror (MGM)                       275                      All
Metro mirror (MM)                               275                      All
Operating environment (OEL)                     325                      All
Parallel access volumes (PAV)                   115                      CKD
Point in time copy (PTC)                        275                      All
RMZ Resync                                      275                      CKD
Remote mirror for z/OS (RMZ)                     275                      CKD
z-synergy services                              115                      CKD
```

Figure 9-16   Using the lskey command to list the installed licenses

For more information about the DS CLI, see the *IBM DS8000 Series Command-Line Interface User's Guide*, SC27-8526.

The BF license bundle needs to be installed before ranks can be formatted for FB (open systems). The zsS license bundle needs to be installed before ranks can be formatted for CKD (mainframe).

9.3 Licensed scope considerations

An increase in license capacity is concurrent. A deactivation (or decrease in license capacity) takes place only after one additional machine IML. Similar considerations apply to the license scopes:

► The BF license and CS license are available with several license scopes, such as ALL, FB, and CKD.

► An increase in license scope, for example, changing FB or CKD to ALL, is a concurrent activity.

► A lateral change, such as changing `FB` to `CKD` or changing `CKD` to `FB`, or a reduction of the license scope, such as changing `ALL` to `FB` or `CKD`, is a disruptive activity because it requires an IML to activate the change.

Tip: Because the BF license must be ordered for the full physical capacity anyway, and because the CS license can be ordered for only those volumes that are in Copy Services relationships, consider the following tip: For BF and CS, configure these bundles with scope "`ALL`" from the beginning.

Part 3

Storage configuration

This part of the book describes the storage configuration tasks that are required on an IBM DS8880.

This part contains the following chapters:

- ► Chapter 10, "Configuration flow" on page 245
- ► Chapter 11, "DS8880 Storage Management GUI" on page 251
- ► Chapter 12, "Configuration with the data storage command-line interface" on page 347

10

Configuration flow

This chapter provides a brief overview of the sequence of tasks that are required for the configuration of an IBM DS8880.

This chapter covers the following topics:

- ► Configuration worksheets
- ► User and role management
- ► Disk encryption
- ► Network security
- ► Configuration flow
- ► General storage configuration guidelines

10.1 Configuration worksheets

Before a new DS8880 is delivered, the client is given the DS8880 customization worksheets. The configuration worksheets can be found in Appendix E of the *IBM DS8880 Introduction and Planning Guide*, GC27-8525. The guide provides all of the information that is required to plan for a successful installation.

The purpose of the configuration worksheets is to provide the required information to the IBM service support representative (IBM SSR) to customize the DS8880. It is best to present the completed worksheets to the IBM SSR before the delivery of the DS8880.

The completed customization worksheets specify the initial setup for the following items:

► Company information: Provides important contact information.

► Management console network settings: Supplies the IP address and local area network (LAN) settings for connectivity to the management console.

► Remote support (which includes call home and remote service settings): Specifies the inbound and outbound remote support settings.

► Notifications: Lists Simple Network Management Protocol (SNMP) trap and email notification settings.

► Power control: Selects and controls the various power modes for the storage complex.

► Control switch settings: Specifies certain DS8880 settings that affect host connectivity. This information is required by the IBM SSR so that the SSR can enter these settings during the DS8880 installation.

10.2 User and role management

During the planning phase (when you use the customization worksheet), list all users who need access to the data storage graphical user interface (DS GUI) or data storage command-line interface (DS CLI). This action helps you manage secure authorization, which specifies the resource and access for different role-based users. A user can be assigned to more than one group. Assign at least one user to each of the following roles:

► The Administrator (`admin`) has access to all Hardware Management Console (HMC) or Management Console (MC) service methods and all storage image resources, except for encryption functions. This user authorizes the actions of the Security Administrator during the encryption deadlock prevention and resolution process.

► The Security Administrator (`secadmin`) has access to all encryption functions. This role requires an Administrator user to confirm the actions that are taken during the encryption deadlock prevention and resolution process.

► The Physical operator (`op_storage`) has access to physical configuration service methods and resources, such as managing the storage complex, storage image, rank, array, and extent pool objects.

► The Logical operator (`op_volume`) has access to all service methods and resources that relate to logical volumes, hosts, host ports, logical subsystems, and volume groups, excluding security methods.

► The Monitor group has access to all read-only, nonsecurity MC service methods, such as the **list** and **show** commands.

► The Service group has access to all MC service methods and resources, such as running code loads and retrieving problem logs. This group also has the privileges of the Monitor group, excluding security methods.

► The Copy Services operator has access to all Copy Services methods and resources, and the privileges of the Monitor group, excluding security methods.

> **Important:** Available resource groups offer an enhanced security capability that supports the hosting of multiple customers with Copy Services requirements. It also supports the single client with requirements to isolate the data of multiple operating systems' environments. For more information, see *IBM System Storage DS8000 Copy Services Scope Management and Resource Groups*, REDP-4758.

► No access prevents access to any service method or storage image resources. This group is used by an administrator to deactivate a user ID temporarily. By default, this user group is assigned to any user account in the security repository that is not associated with any other user group.

10.3 Disk encryption

Additional planning is required if you intend to activate encryption for the DS8880. It is important to plan and configure encryption before you perform the logical configuration.

The DS8880 provides disk-based encryption for data that is within the storage system, for increased data security. This disk-based encryption is combined with an enterprise-scale key management infrastructure.

Although all DS8880 systems have certificates installed, encryption is optional. It is activated when licensed Feature Code 1750 is ordered. Activation must be completed before you perform any logical configuration. For more information about encryption license considerations, see "Encryption activation review planning" on page 198.

The current DS8880 encryption solution requires the use of the IBM Security Key Lifecycle Manager, IBM Security Key Lifecycle Manager for z/OS, or the third-party solution SafeNet. These key lifecycle managers assist with generating, protecting, storing, and maintaining encryption keys that are used to encrypt information that is written to and decrypt information that is read from devices.

For more information, including current considerations and preferred practices for DS8880 encryption, see 7.3.6, "Key manager servers for encryption" on page 198 and *IBM DS8880 Data-at-rest Encryption*, REDP-4500.

10.4 Network security

The security of the network that is used to communicate to and manage the DS8880 (specifically the HMC) is important, depending on the client requirements. The DS8880 supports compliance to the National Institute of Standards and Technology (NIST) SP800-131a standards, which are also known as *Gen-2 security*.

Two components are required to provide full network protection:

► The first component is Internet Protocol Security (IPSec), and for Gen-2 security, IPsec-v3 is required. IPSec protects the network communication at the Internet layer, or the packets that are sent over the network. This configuration ensures that a valid workstation or server communicates with the HMC and that the communication between them cannot be intercepted.

► The second component is Transport Layer Security (TLS) 1.2, which provides protection at the application layer to ensure that valid software (external to the HMC or client) is communicating with the software (server) in the HMC.

> **Note:** The details for implementing and managing Gen-2 security requirements are provided in *IBM DS8870 and NIST SP 800-131a Compliance*, REDP-5069.

10.5 Configuration flow

This section shows the list of tasks to perform when storage is configured in the DS8880. Depending on the environment and requirements, not all tasks might be necessary.

Logical configuration can be performed by using the DS GUI, DS CLI, or both. Depending on the client's preference and experience, one method might be more efficient than the other. The DS8880 R8.1 GUI provides a powerful, yet simple process for logical configuration. If you use the DS Storage Management GUI, not all of the steps that are listed in this book are explicitly performed by the user. For more information about the DS Storage Management GUI, see Chapter 11, "DS8880 Storage Management GUI" on page 251.

If you perform logical configuration by using the DS CLI, the following steps provide a high-level overview of the configuration flow. For more detailed information about using and performing logical configuration with the DS CLI, see Chapter 12, "Configuration with the data storage command-line interface" on page 347.

The following is the general configuration flow:

1. Install license keys: Activate the license keys for the DS8880 storage system. For more information about activating licensed functions, see Chapter 9, "IBM DS8880 features and licensed functions" on page 219.

   > **Important:** If encryption will be activated, the encryption configuration must be performed before the logical configuration that is described in the next steps.

2. Create arrays: Configure the installed disk drives as RAID 6, RAID 10, or RAID 5 arrays.

3. Create ranks: Assign each array to be a fixed-block (FB) rank or a count key data (CKD) rank.

4. Create extent pools: Define extent pools, associate each one with Server 0 or Server 1, and assign at least one rank to each extent pool. To take advantage of storage pool striping, you must assign multiple ranks to an extent pool. For more information about storage pool striping, see "Storage pool striping: Extent rotation" on page 166, and "Storage pool striping" on page 375.

> **Important:** If you plan to use IBM Easy Tier (in particular, in automatic mode), select the **All pools** option to receive all of the benefits of Easy Tier data management. For more information, see 6.7, "IBM Easy Tier" on page 173.

5. Create a repository for space-efficient volumes. For more information, see *IBM DS8880 Thin Provisioning*, REDP-5343.

6. Configure the I/O ports: Define the topology of the I/O ports. The port type can be Switched Fabric or Fibre Channel Protocol (FCP), Fibre Channel Arbitrated Loop (FC-AL), or Fibre Channel connection (FICON).

7. Create the volume groups for open systems: Create volume groups where FB volumes are assigned.

8. Create the host connections for open systems: Define open systems hosts and their Fibre Channel (FC) host bus adapter (HBA) worldwide port names (WWPNs). Assign volume groups to the host connections.

9. Create the open systems volumes: Create striped open systems FB volumes and assign them to one or more volume groups.

10. Create the z Systems logical control units (LCUs): Define their type and other attributes, such as subsystem identifiers (SSIDs).

11. Create the striped z Systems volumes: Create z Systems CKD base volumes and parallel access volumes (PAV) aliases for them.

10.6 General storage configuration guidelines

Observe the following general guidelines when storage is configured in the DS8000:

► To achieve a well-balanced load distribution, use at least two extent pools (also known as a *pool pair*), each assigned to one of the internal servers (extent pool 0 and extent pool 1). If CKD and FB volumes are required on the same storage system, configure at least four extent pools: Two for FB and two for CKD.

► The volume type for the first volume that is created in an address group is either FB or CKD. That volume type determines the type for all other volumes (FB or CKD) for the entire address group. A volume is one of 256 in a logical subsystem (LSS) or LCU. An LSS is one of 16 in an address group (except address group F, which has only 15 LSSs). For more information about logical subsystems and address groups, see 4.4.5, "Logical subsystems" on page 122.

► Volumes of one LCU/LSS can be allocated on multiple extent pools in the same rank group.

► Assign multiple ranks to extent pools to take advantage of storage pool striping. Additionally, assign ranks from multiple device adapter (DA) pairs to an extent pool to spread the workload and increase performance. See 6.5.2, "Data placement in the DS8000" on page 164.

► The following options are available for fixed-block (FB) pools:

  – Create a volume group for each server unless logical unit number (LUN) sharing is required.

  – Assign the volume group for one server to all of its host connections.

– If LUN sharing is required, the following options are available (Figure 10-1):

- Create one volume group for each server. Assign the shared volumes in each volume group. Assign the individual volume groups to the corresponding server's host connections. The advantage of this option is that you can assign private and shared volumes to each host. This configuration can be used in an environment such as application sharing.

- Create one common volume group for all servers. Place the shared volumes in the volume group and assign the volume group to the host connections. This configuration can be used in an environment such as clustering.
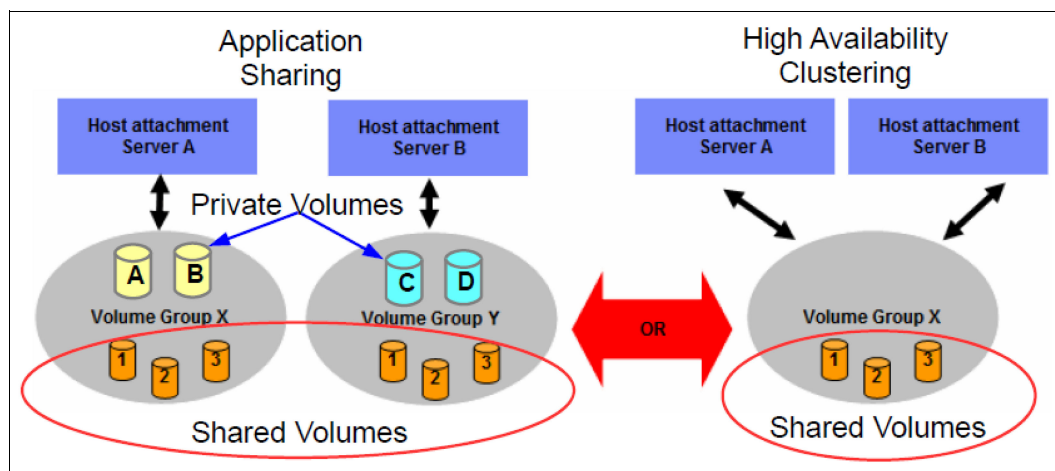
Figure 10-1   LUN configuration for shared access

► The following options are available for I/O ports:

– Configure a port to be FICON, FCP, or FC-AL.

– Distribute host connections of each type (FICON, FCP, and FC-AL) evenly across the I/O enclosures.

– Ensure that each host is connected to at least two different host adapters (HAs) in two different I/O enclosures for redundancy and availability.

– Use access `any` typically for I/O ports with access to ports that are controlled by storage area network (SAN) zoning.

Note: Avoid intermixing host I/O with Copy Services I/O on the same ports.

11

DS8880 Storage Management GUI

This chapter describes the DS8000 Storage Management GUI that is available with the DS8880 (R8.2). It allows you to perform these functions on the storage system:

► Initial system setup for a new installation
► Activation of licensed functions
► Simplified logical configuration for open systems and IBM z Systems
► Advanced logical configuration for open systems and IBM z Systems
► Graphical view of system resource availability
► System status and events viewer
► Access to advanced help and the IBM Knowledge Center
► Performance monitoring and statistic offload

This chapter covers the following topics:

► Introduction
► DS8000 Storage Management GUI: Getting started
► System configuration overview
► Logical configuration overview
► Logical configuration for fixed-block volumes
► Logical configuration for count key data volumes
► Monitoring system health
► Performance monitoring
► Fibre Channel error rate statistics
► Accessing the previous DS GUI

11.1 Introduction

The Storage Management GUI (DS8000 GUI) is packed with usability features that offer an easy to use way of administering a DS8000. An enhanced GUI that was introduced with DS8870 Release 7.5.

The DS8000 GUI was designed and developed with three major objectives:

► Speed: A responsive GUI is a requirement.
► Simplicity: A simplified and intuitive design can drastically reduce the total cost of ownership (TCO).
► Commonality: Common graphics, widgets, terminology, and metaphors facilitate the management of multiple IBM storage products and software products.

With these deliverables in mind, a system administrator can use the DS8000 GUI to configure the storage system logically and prepare the storage system for I/O quickly after completing the setup. Then the system administrator can manage the system with minimal knowledge. The GUI provides a performance monitoring tool for the main storage parameters.

Logical storage configuration is streamlined in the DS8000 GUI. Creating and configuring pools is easy. The conceptual approach was changed to combine array site, array, and rank into a single resource, which is referred to as an *array*. The storage system can automatically manage device adapter (DA) pairs and balance arrays and spares over the two processor nodes.

With the enhanced GUI, the process for creating volumes is simplified. The system can automatically balance volume capacity over a pool pair. If storage needs require that the user balances arrays, spares, and volume capacity manually, the DS8000 GUI also provides custom options for configuring pools and volumes.

Configuring connections to hosts is also easier. Host ports are updated automatically and host mapping is allowed at volume creation. From the data storage command-line interface (DS CLI) perspective, commands are available to make, change, or remove a host connection. See Chapter 12, "Configuration with the data storage command-line interface" on page 347.

The underlying architecture and virtualization are not changed, but the GUI navigation and its wizards make them more intuitive and self-explanatory. For more information about DS8880 virtualization, see Chapter 4, "Virtualization concepts" on page 97.

The storage system status can be viewed at any time in the pods that are displayed on the bottom of each page. The status and attributes of all hardware elements are displayed on the system page. All changes to the storage system, whether they were initiated by a user or the system, can be viewed on the Events page. Alerts are displayed at the bottom of each page and they link to the Events page.

Additionally, functions that include user access, licensed function activation, setup of Encryption and Remote Authentication, and modifying power or I/O port protocols are available to the system administrator.

The DS8000 GUI provides performance metrics that are grouped into four sections:

► Storage System (all DS8880 performance metrics)
► Array (array level performance metrics)
► Pool (pool-dependent performance metrics)
► I/O Port (I/O port-dependent performance metrics)

The DS8000 GUI provides links to the Storage Management IBM Knowledge Center, User Learning tutorials, and context help.

For certain special functions, access to the former version of the DS GUI might be needed. This possibility still exists with Release 8.2 of DS8880 microcode. Directions are provided to access the previous DS GUI for functions that are not available with the current GUI. For more information, see 11.10, "Accessing the previous DS GUI" on page 345.

If required, the DS command-line interface (DS CLI) can be used to manage functions that are not supported in this release of the DS8000 Storage Management GUI.

11.2 DS8000 Storage Management GUI: Getting started

This section describes how to accomplish the following tasks:

► Accessing the DS8000 Storage Management GUI
► Using the Storage Management Setup Wizard
► Managing and monitoring the system
► Performance monitoring the system
► Using the Help functions

11.2.1 Accessing the DS8000 Storage Management GUI

The Storage Management GUI can be accessed in the following ways:

► With a browser that is connected to the DS8880 Hardware Management Console (HMC)
► By using IBM Spectrum Control that has connectivity to the HMC

Accessing the Storage Management GUI with a browser

Web browsers including Mozilla Firefox, Microsoft Internet Explorer, and Google Chrome are supported for the DS8000 GUI.

The DS8880 supports the ability to use a single point of authentication function for the Storage Management GUI through a centralized Lightweight Directory Access Protocol (LDAP) server. This capability is supported locally by the internal Copy Services Manager (CSM) or through external proxy by IBM Spectrum Control. For more information about remote authentication and LDAP for the DS8880, see *IBM DS8880 Integrated Copy Services Manager and LDAP Client on the HMC*, REDP-5356.

The DS8000 GUI has a new URL. The DS8000 GUI can be accessed from a browser by using the http address that is shown in Example 11-1. In this address, *HMC_IP* is the IP address or host name of the Management Console (MC). You are not required to specify a port ID as in previous releases of the DS8000 GUI.

Example 11-1 DS8880 Storage Management GUI URL

```
https://<HMC_IP>
```

On a new storage system, the user must log on as the administrator. The password will expire and the user is forced to change the password. Figure 11-1 shows the following information:

1. Storage Management GUI URL: `https://<HMC_IP>`

   This Storage Management GUI URL is configured by the IBM service support representative (IBM SSR) during the installation.

2. Initial login on the new machine

   For a new machine, the following logon credentials are required:

   – User = `admin`.

   – Password = `admin`.

3. The default password is expired, and the user is forced to change the password for the `admin` user ID.

4. Service Management icon (the 'Wrench')

   Clicking the Service Management icon opens a new window to the Management Console Web User Interface (WUI), which can be accessed to perform the following tasks:

   – View open problems.

   – Manage remote access.

   – Restart the enterprise storage server network interface (ESSNI).

   – Perform various other tasks.
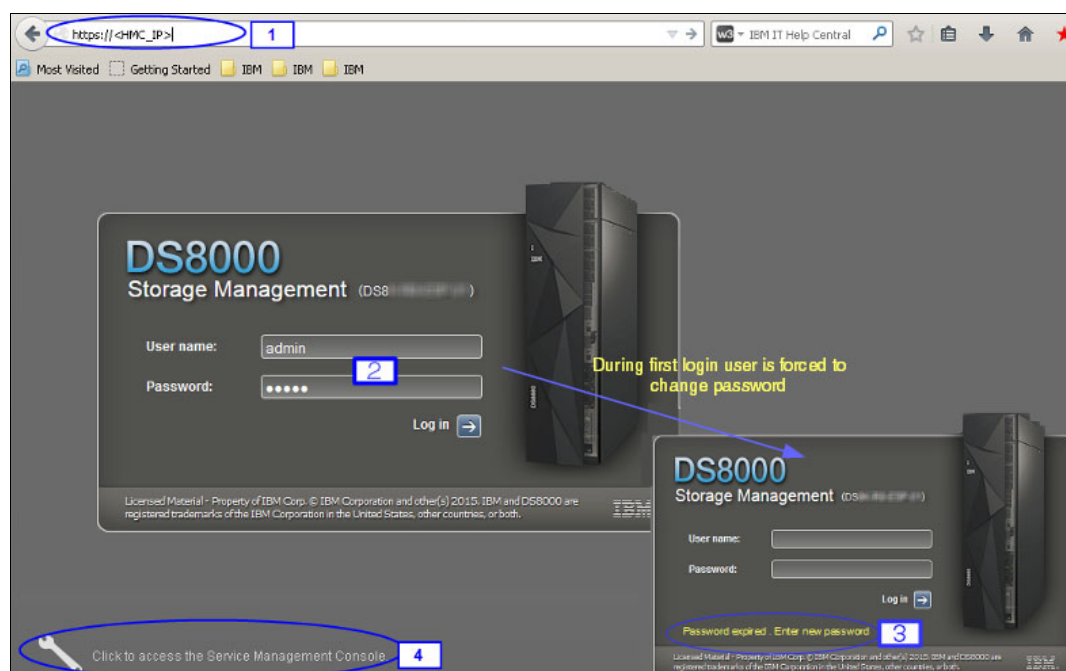
Figure 11-1   Storage Management GUI initial login

11.2.2 Storage Management GUI System Setup wizard

For a new storage system installation, the DS8000 GUI starts the *System Setup wizard*. The setup wizard starts automatically after the `admin` password is changed and a user with the Administrator role and authority logs in.

The setup wizard guides the `admin` user through the following tasks:

1. Set the system name.
2. Activate the licensed functions.
3. Provide a summary of actions.

The System Setup window opens the Welcome window, as shown in Figure 11-2.
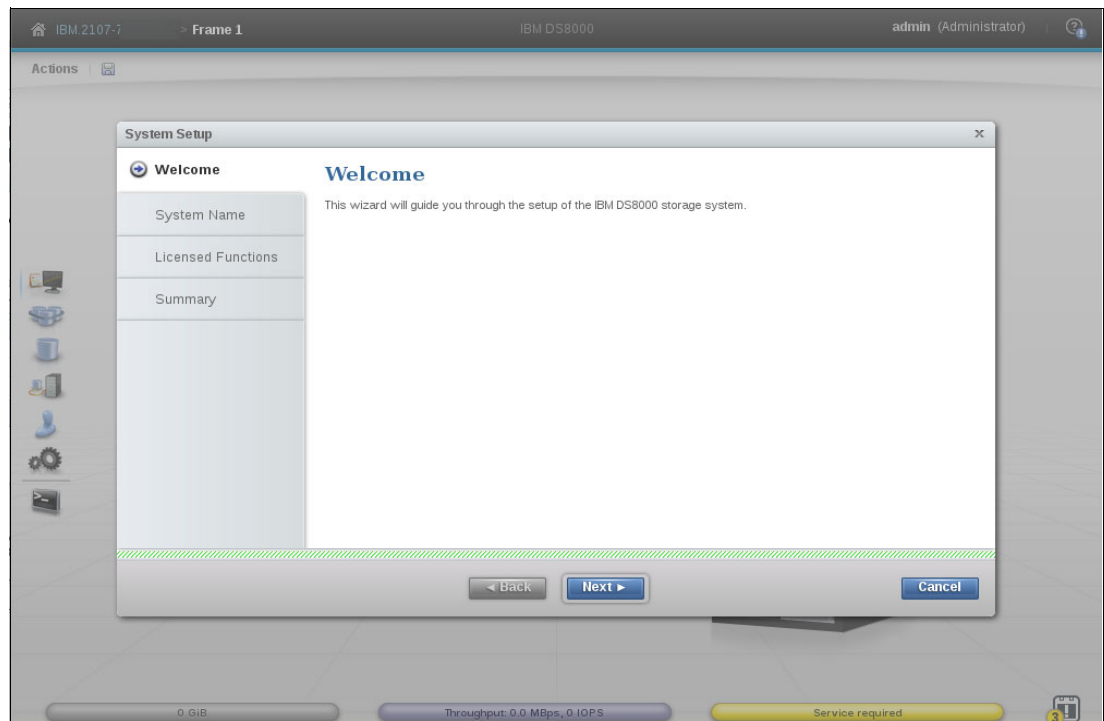
Figure 11-2   System Setup: Welcome window

Starting with the Welcome window, complete the following steps:

1. Click **Next**. The System Name window opens. The default entry is the storage system serial number, which is shown in Figure 11-3. The user can create a preferred system name that complies with the name convention that is used in their environment.
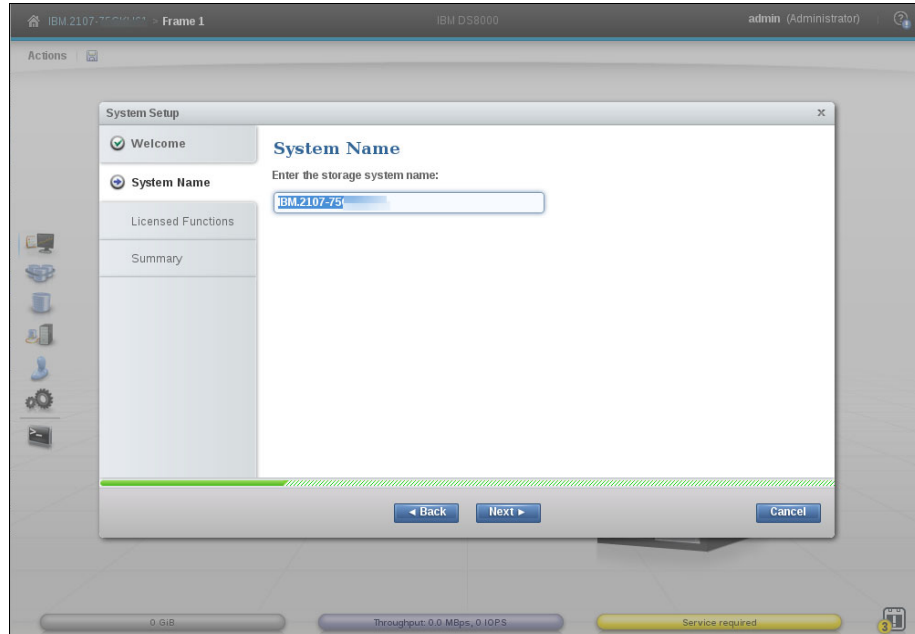
Figure 11-3   System Setup: System Name window

2. Click **Next**. The Licensed Functions window opens. Click **Activate Licensed Functions** as indicated in Figure 11-4.
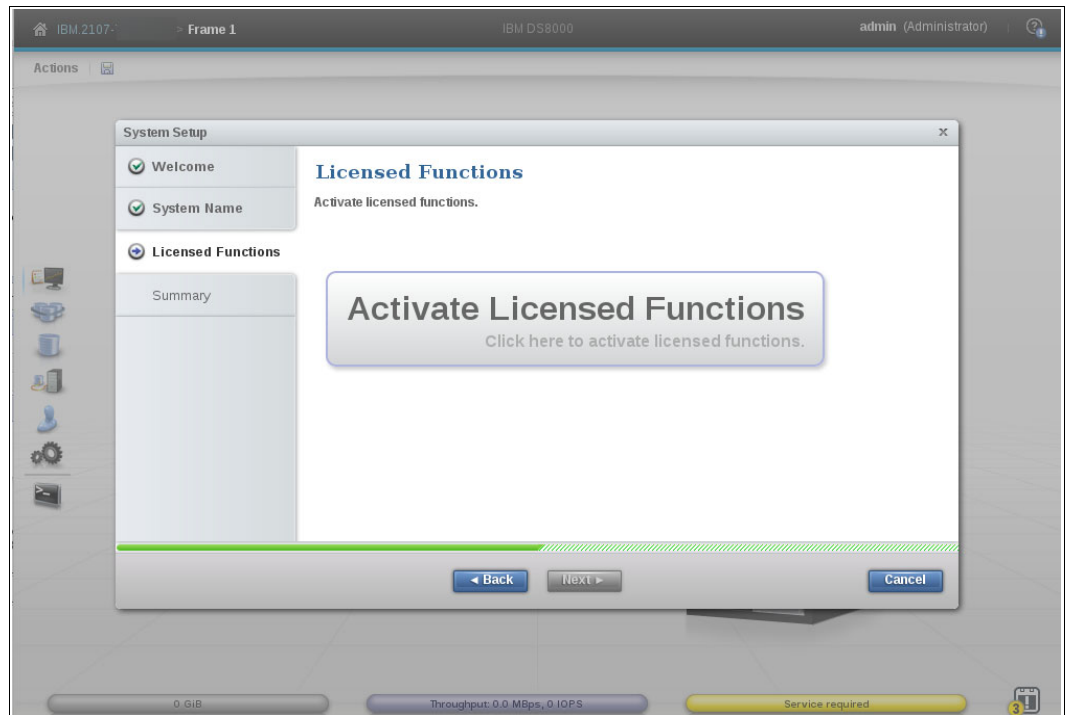
Figure 11-4   System Setup: Licensed Functions window

3. The Activate Licensed Functions window opens. Keys for licensed functions that have been purchased for this storage system need to be retrieved by using the Machine Type, Serial Number, and Machine Signature. The keys can be stored in a flat file or an XML file. Licensed function keys are downloaded from the data storage feature activation (DFSA) website:

   https://www.ibm.com/storage/dsfa

   Figure 11-5 shows the Activate License Functions window, the help information, and the activation of the licensed functions. You can import the keys by selecting the folder icon or paste the keys in the window. For more information about licensed functions, see Chapter 9, "IBM DS8880 features and licensed functions" on page 219.
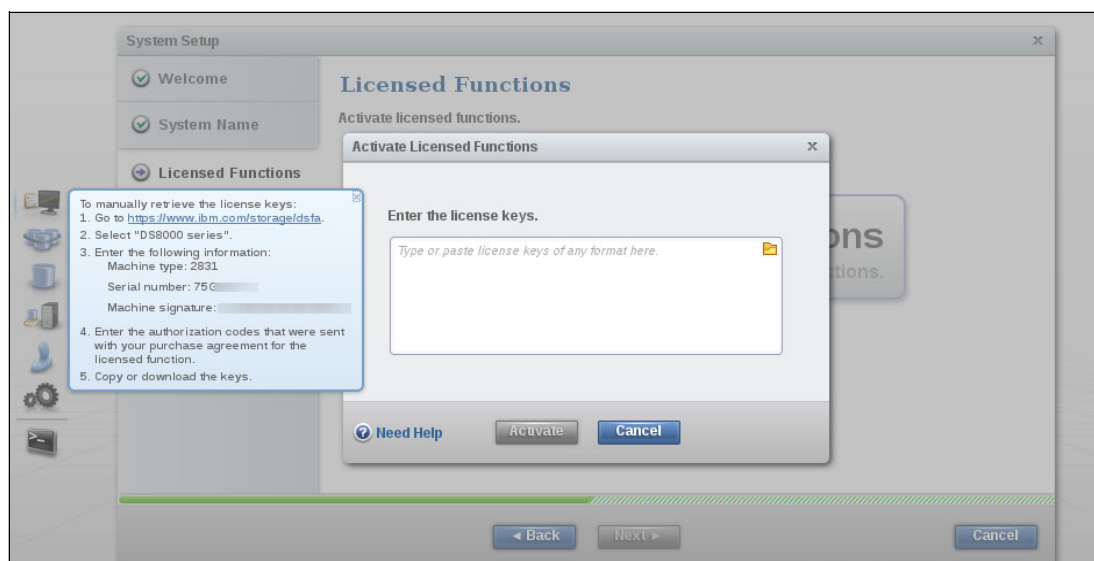
Figure 11-5   System Setup: Activate Licensed Functions

4. When the license keys are entered click **Activate** to enable the functions as illustrated in Figure 11-6.
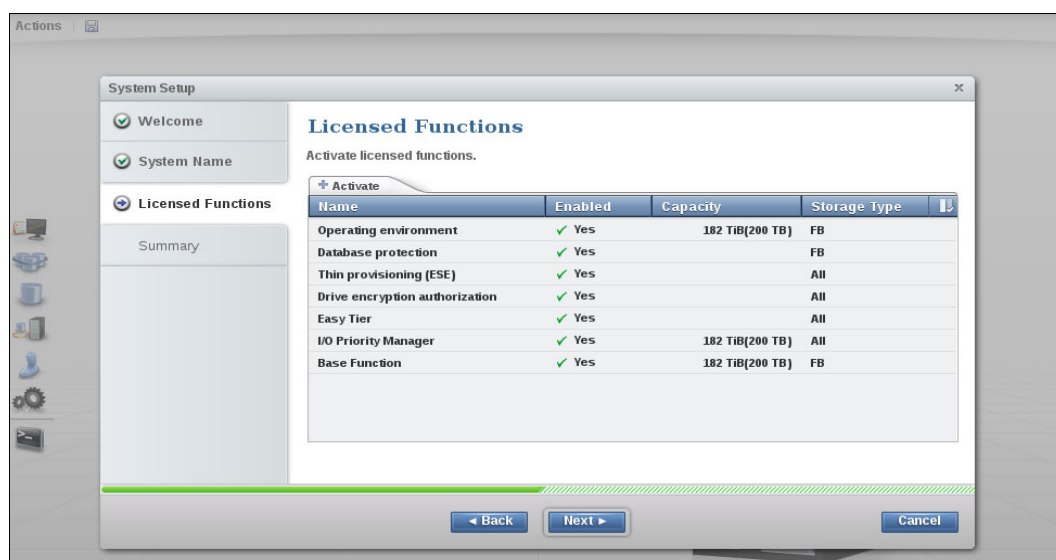
Figure 11-6   System Setup: Licensed Functions Enabled

Note: If the IBM Easy Tier license is activated, that pool pair is managed by Easy Tier by default. For more information, see *IBM DS8000 EasyTier*, REDP-4667.

5. Click **Next** to open the Summary window, which is shown in Figure 11-7. If everything looks correct, click **Finish** to exit the System Setup Wizard.

Note: The summary below shows licenses for basic functions. The list could include some advanced functions such as Copy services, z-synergy Services, and CSM (Copy Services Manager on HMC) if the licenses for them were activated.
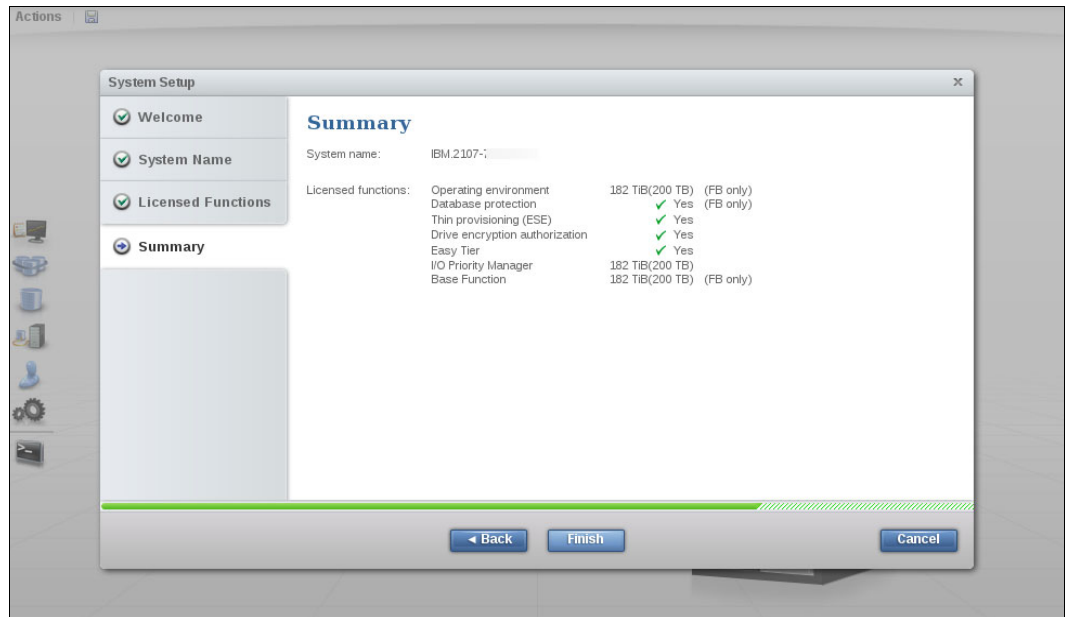
Figure 11-7   System Setup: Summary

The wizard closes and the System window opens.

11.2.3 Configuring I/O Port protocols/topologies

After initial setup is complete, you can configure the I/O port topologies if you know how each port will be attached to hosts (open system or z Systems).

Click **Settings** → **Networking** → **Fibre Channel Networks** and use the actions there to set them as shown in Figure 11-8. You can also do this configuration during logical configuration. See "Creating FB host attachments" on page 289 and "Setting the I/O port protocols for z Systems attachment" on page 321.
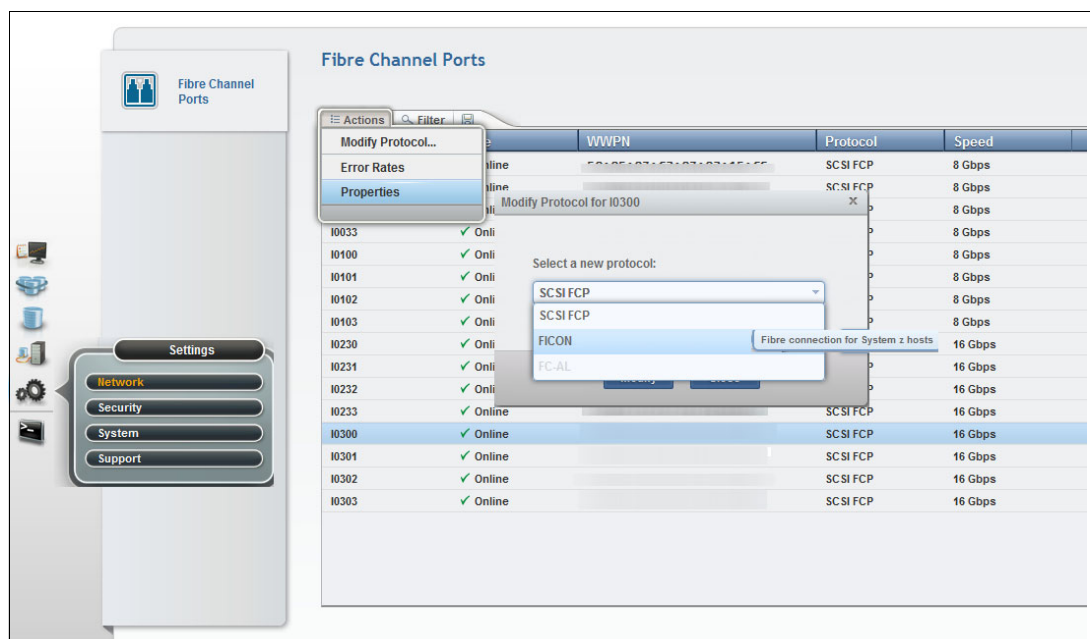
Figure 11-8   Configure I/O Port topology

11.2.4 Managing and monitoring the storage system

When the initial system setup is completed, the system page is displayed. This page is the first page that opens after a user logs in. The system page displays the major hardware components of the DS8880 storage system and shows the status of the system hardware. You can return to this page at any time by clicking the **Home/Name** icon in the upper left of the GUI.

Figure 11-9 on page 260 shows the System window, with no selected or highlighted object.

From the System window, the user can manage the system by accessing actions:

▶ Controlling how the system is powered on or off
▶ Modifying the I/O port protocols
▶ Modifying Easy Tier settings
▶ Viewing system properties
▶ Displaying performance monitoring graphs

The system administrator can create and modify the logical configuration of the system. Details can be viewed about each hardware component by right-clicking (or pointing to) the component to display an enlarged view and the current state of that component.

Figure 11-9   The DS8880 System window

Figure 11-10 presents a high-level overview of the System window and all the objects that can be accessed from this window. Not all functions display for every user role. For more information about user roles, click **Help** and search for `user role`.
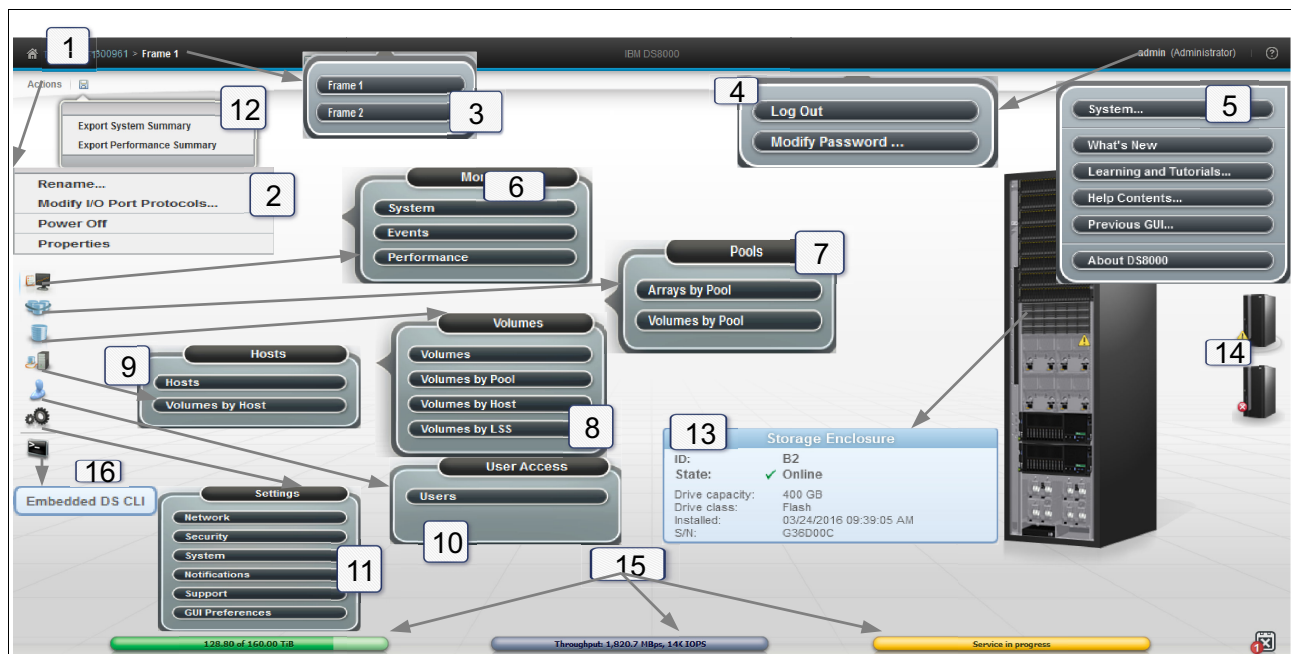
Figure 11-10   System window comprehensive view

The numbered labels in Figure 11-10 on page 260 have these meanings:

1. Home/Name icon:
   – The name of the DS8000 storage system as set by the administrator.
   – Click the Name/Home icon from anywhere to return to the System window.
2. Actions menu:
   – Rename

     Change the name of the DS8880 storage system.
   – Modify I/O Port Protocols

     Select the protocol that is used by I/O ports to connect the storage system to a host or to another storage system. The user can also display the properties of the selected port. Port error rates are not available from this view, and must be accessed by clicking **Settings** → **Network Fibre** → **Channel Ports**.
   – Power Off/On

     Initiate power off or power on of the DS8880 storage system, as required.
   – Properties

     View the properties of the DS8880 storage system shown in Figure 11-11:
     • Current state of the DS8000 storage system.
     • Release and bundle version of the IBM system software.
     • Machine type and model.
     • Unique serial number.
     • Signature (string of characters that identifies a DS8000 storage system), which is used to retrieve license keys from the IBM DSFA website.
     • Worldwide Node Name of the system.
     • Hardware component summary (such as processor type, total subsystem memory, raw data storage capacity, or number of I/O ports).
     • Last power on time stamp.
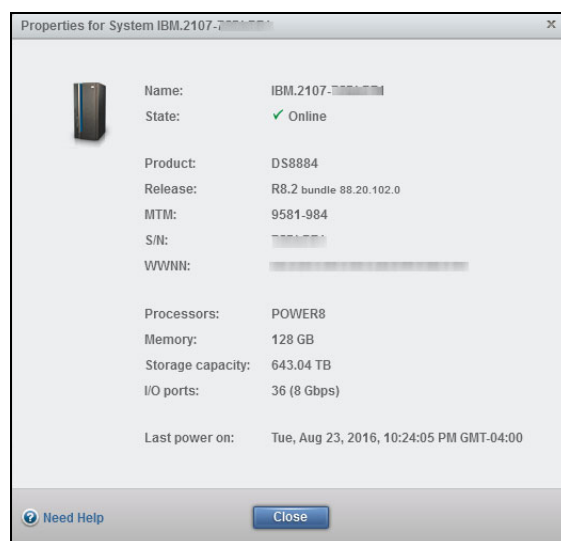
Figure 11-11   System Properties

3. Frame:

   Use the frame selector view (on the right of the system page) to select an individual frame view of a multi-frame storage system for display, if applicable. The frame selector is not displayed if the storage system has only one frame.

4. User icon:

   – Log out of the DS8000 GUI, and display the login page.

   – Modify Password: The logged in user can change their own password.

5. Help icon:

   This icon opens a drop-down menu that includes the following options:

   – The first line in this menu is context-sensitive help. The context-sensitive help pages describe the resources and actions that are available in the GUI, and they open a separate window for the DS8000 GUI IBM Knowledge Center.

   – What's New: This option provides descriptions on new features contained in this release of the GUI with a link to the DS8000 GUI IBM Knowledge Center for usage of each of the new functions. Also provided is a link to discover functions that are supported in this release, and where to find those that are not.

   – Learning and Tutorials: This option opens a separate window for the DS8000 GUI IBM Knowledge Center, at the Overview → Learning and Tutorials section. This section contains e-learning modules that introduce the DS8880 storage system family, define key concepts, and provide an overview of related installation, configuration, and administration tasks.

   – Help Contents: This option opens a separate window for the DS8000 GUI IBM Knowledge Center main page.

   – Previous GUI: This option opens a separate window to the login window of the previous generation's DS GUI. The latest release of the DS8000 Storage Management GUI does not support all of the functions of the previous GUI. The list of limitations is in the DS8000 GUI IBM Knowledge Center. (Click **Help** → **Troubleshooting** → **Limitations**.)

   – About DS8000: This option opens the DS8000 Storage Management window, which shows the following information:

     • Current release 8.x of Licensed Internal Code (LIC) bundle 88.x.x.x.

     • Product, for example, DS8888.

     • Machine type and model (MTM) 283x-98x or 958x-98x (new with Release 8.2). For information about the new models of DS8880, see Chapter 2, "IBM DS8880 hardware components and architecture" on page 23.

     • Machine Serial Number (S/N) 75xxxxx.

   For more information about DS8880 help functions, see 11.2.5, "Storage Management help functions" on page 266.

6. Monitoring menu:
   – System: This option opens the System window.
   – Events: This option opens the Events window.
   – Performance: This option opens the new Performance (statistics) window.

7. Pools menu:

Two options are available here:

– Arrays by Pool: Access this menu to view the pool configuration. This menu can be used only by users with the monitor or copy operator role.

– Volumes by Pool: Access this menu to create, modify, or delete pool pairs. This menu can be used only by users with the administrator or logical or physical operator role.

8. Volumes menu:

– Access this menu to view the volume configuration. This menu can be used only by users with the monitor or copy operator role.

– Access this menu to create, modify, or delete volumes by users with the administrator or logical or physical operator role.

9. Hosts menu:

– Access this menu to view the hosts' configurations. This menu can be used only by users with the monitor or copy operator role.

– Access this menu to create, modify, or delete hosts. This menu can be used only by users with the administrator or logical or physical operator role.

10. User Access menu:

– Only users with the administrator role can access this menu.

– This menu opens the Users window.

– A system administrator can use this menu to perform the following actions:

• Create user accounts.
• Set a user account role.
• Set temporary passwords (to be reset at first use by the new account).
• Modify an existing user account role.
• Reset an existing user account password.
• Disconnect a user account.
• Determine a user account connection (DS CLI or GUI).

11. Settings menu:

– Network:

• Modify Fibre Channel ports protocol for a selected port or group of ports.

• Display error rates for a selected port or group of ports.

• Display a single port's properties.

– Security:

• Manage encryption for the storage system.

• Manage remote Authentication.

• Manage local password rules (such as password minimum length, expiration, and age).

– System:

• Licensed Function

This window displays a summary of the activated licensed functions. When you click the Activate tab, the Activate Licensed Functions window opens, where you can activate additional licenses. For a full description of activating licensed functions, see Chapter 9, "IBM DS8880 features and licensed functions" on page 219.

- Easy Tier

  You can use this function to configure Easy Tier to improve performance by managing or monitoring volume capacity placement in pools. Easy Tier Heat Map Transfer utilities can also be enabled.

- Advanced configuration

  On the Advanced tab of the System settings page, you can configure I/O Priority Manager, enable ESSNet Copy Services, set the IBM i serial number prefix, select the power control mode for the storage system, and enable the DS Open API (CIM Agent).

– Notifications:

- Syslog

  You can define, modify, or remove Syslog servers.

– Support:

- Access

  You can enable or disable IBM service access to the Hardware Management Console (HMC) by using SSH or the DS Service GUI. An Assist On-site (AOS) connection to the HMC for IBM service is also enabled from the Access tab.

– GUI Preferences:

- Navigation

  You can enable or disable animation of the navigation dock in the GUI.

- Login Message

  You can enter a message that is displayed when users log in to either the DS GUI or DS CLI.

- General

  You can refresh the browser cache for the GUI and set the default logout timeout.

12. Export storage system and performance summary:

– As administrator, click the **Diskette** icon and select **Export System Summary** to save a DS8880 configuration to a comma-separated values (CSV) file.

– As administrator, click the **Diskette** icon and select **Export Performance Summary** to save a DS8880 performance data gathered at 1 minute per interval for up to 1 week to a comma-separated values (CSV) file.

13. Component View:

– The large graphic of one of the frames in the DS8880 system is shown.

– Use your mouse pointer to 'hover-over' an enclosure, node, HMC, or uninterruptable power supply (UPS), and a **Properties** menu opens that shows basic information about that particular component in the frame.

– If a **Magnifying Glass** icon appears when hovering over a component, left-click to bring the component to the center of the window and allow properties for sub-components to be displayed. Click anywhere on the window to return the component to the DS8880 system.

– When viewing an I/O port on a Host Adapter, you can right-click a port to modify the I/O port protocol.

The following information is typically displayed for an enclosure:

- Enclosure name.
- ID.
- State, for example, online, offline, or service required.
- Enclosure-specific data.
- Enclosure serial number.

The following information is typically displayed for a node:

- ID.
- State, for example, online, offline, and service required.
- DS8880 release.
- Data that is specific to a Power System.

– For more information about the Component view, see 11.7, "Monitoring system health" on page 321.

14. Frame Selection:

An icon exists for each frame in the system. By selecting a frame icon, that frame is then visible as the enlarged view immediately to the left, where the frame's components can be viewed.

15. Status Area:

The status area consists of the following objects:

– Capacity: Changes that affect the capacity of the storage system are displayed in the Capacity pod on the left side of the status area at the bottom of the Storage Management GUI. Click the Capacity pod to go to the Arrays by Pool window.

– Performance: I/O throughput is displayed in the Performance pod in the middle of the status area at the bottom of the Storage Management GUI. Click the Performance pod to go to the Performance page.

– System health: Changes that affect data accessibility and hardware states are displayed in the System health pod on the right side of the status area at the bottom of the Storage Management GUI. If a hardware error is displayed, click the System health pod to go to the System window, where you can observe the hardware component that needs attention. For more information about system events, see 11.7, "Monitoring system health" on page 321.

– Alerts: If any changes are generated by DS8880, an additional icon that is called **Alerts** is displayed, which shows the number of recent alerts. Click the **Alerts** icon to go to the Events window.

16. Embedded DSCLI:

Select the icon to open a DSCLI session from the DS GUI. The version of embedded DS CLI is the latest available for the microcode installed on the DS8880.

The DSCLI commands can be run conveniently from the GUI with the least amount of response time avoiding network hops when using remote DSCLI.

DSCLI scripts can also be executed from the embedded DSCLI on DSGUI. The script must be text based and located on the work station running the DS GUI session. See Figure 11-12.
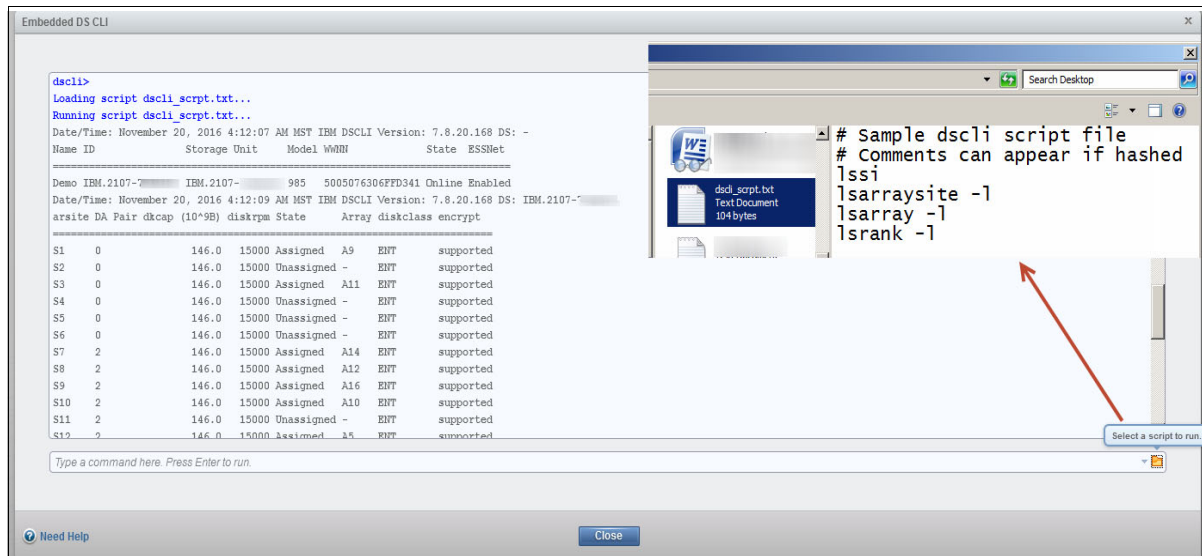
Figure 11-12 Embedded DSCLI

11.2.5 Storage Management help functions

The DS8880 Storage Management GUI provides access to comprehensive help functions. The help functions not only assist you with using the GUI, but also provide in-depth help to the overall DS8880 Storage System and functions. To access the help contents, click the **Help** icon, then select **Help Contents** to open a separate window for the DS8000 GUI IBM Knowledge Center main page, as illustrated in Figure 11-13.
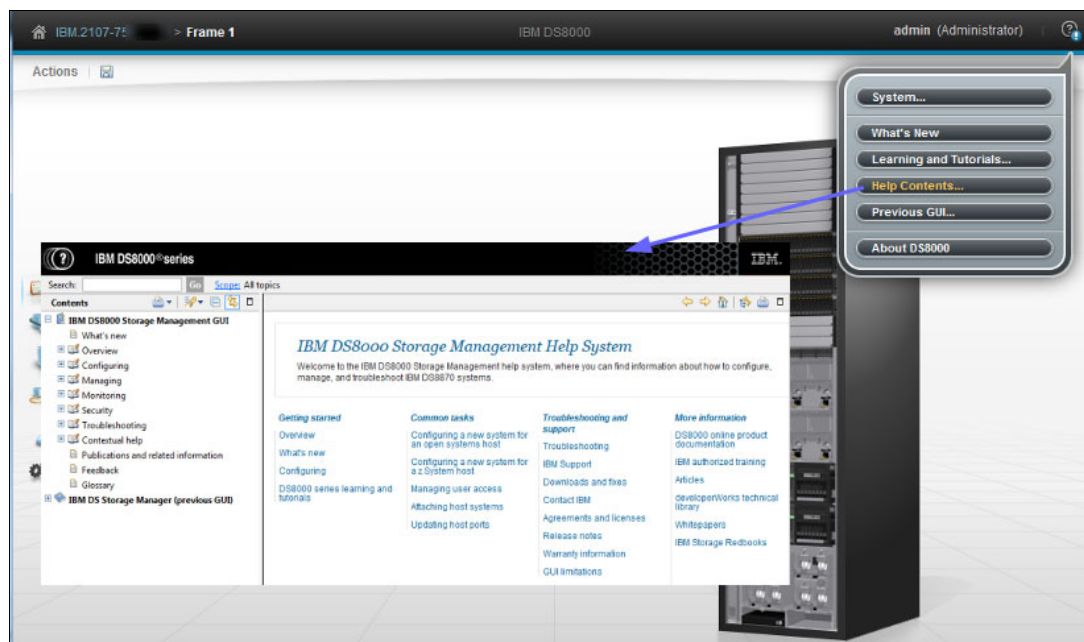
Figure 11-13   Storage Management Help System: IBM Knowledge Center

From the IBM Knowledge Center, you can discover introductory information about the DS8880 architecture, features, and advanced functions. Also, you can learn about management interfaces and tools. Many tutorials and videos can help you understand the product.

You can obtain more information about using the Storage Management GUI for common tasks, such as these:

► Logically configuring the storage system for open systems and z Systems attachment
► Configuring open systems host attachment
► Managing user access

The IBM Knowledge Center also provides links to external links for more information about IBM storage systems, and other related online documentation.

11.3 System configuration overview

After the DS8880 administrator configures the initial system, the administrator can configure additional system functions according to the storage requirements.

11.3.1 Security settings

Security is always important. Specifically, if you want to use the data-at-rest encryption capability with the DS8880, you must set it before you store any data on the system.

Encryption

To enable encryption, from the system page, click the **Settings** icon. Click **Security** to open the Security window, and click the **Encryption** tab. See Figure 11-14.

> **Important:** If you must activate encryption for the storage system, ensure that the encryption license is activated and the encryption group is configured before you begin any logical configuration on the system. When pools exist, you cannot disable or enable encryption. For more information, see *IBM DS8880 Data-at-rest Encryption*, REDP-4500.

Figure 11-14   Enable Encryption

With Release 8.2 of DS8880 microcode, you can define a custom certificate for communication between the encryption key servers (typically IBM Security Key Lifecycle Manager) and the storage system.

A system generated certificate, GEN1 or GEN2 (NIST SP 800-131a compliant), can be updated to a customer generated certificate. The custom certificate must meet all the requirements or the update will fail. For the list of requirements, see **Updating encryption and key server certificates** in **Help Contents**. Also, see "Storage Management help functions" on page 266. For more information about encryption, see *IBM DS8880 Data-at-rest Encryption*, REDP-4500.

To update the DS8000 Encryption Certificate with a customer defined certificate, complete these steps:

1. Click **Update Certificate**.
2. Select the **Customer defined** option.

3. In the Customer defined window, click the file folder to specify the location of certificate and enter your password as shown in Figure 11-15.
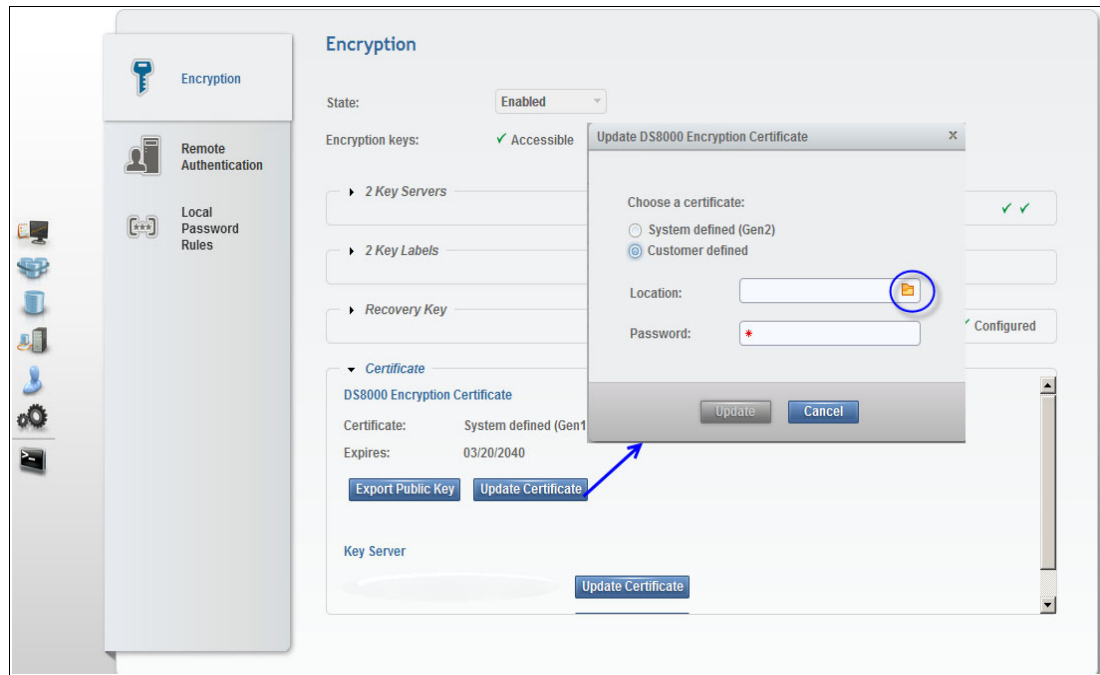
Figure 11-15   Encryption Settings Update Certificate

Remote Authentication

Customers who have restricted security policies for user logins can now enable remote authentication for their DS8880 storage system.

To enable this feature, complete these steps:

1. Click **Settings** → **Security** → **Remote Authentication**.

2. In the Remote Authentication window, click **Enable Remote Authentication**.

The Welcome page of the wizard opens with the prerequisites for enabling remote authentication as shown in Figure 11-16. This wizard guides you through the remote authentication setup for IBM DS8880.
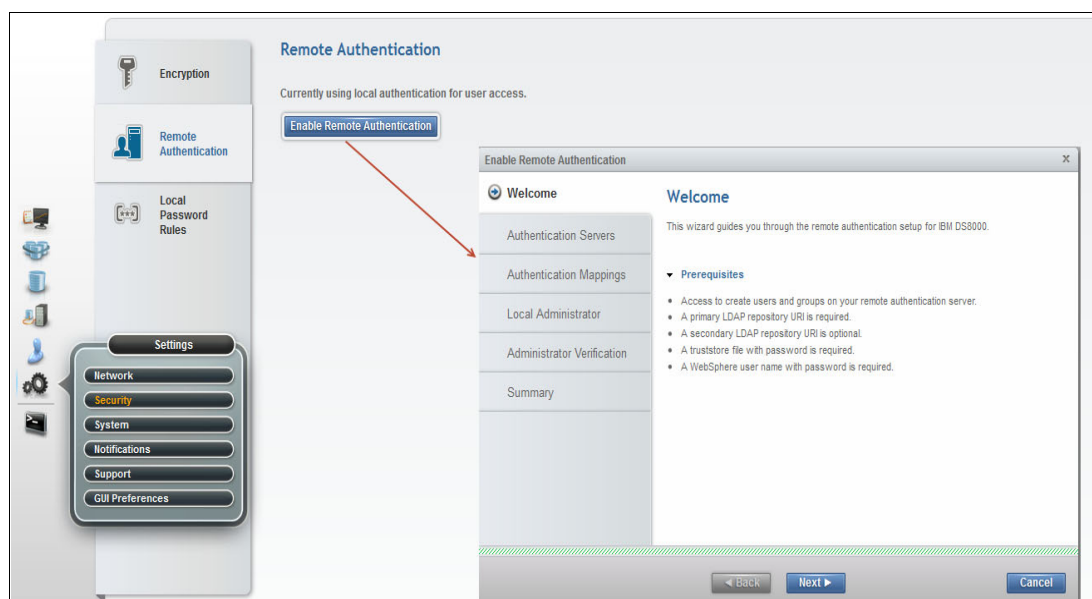
Figure 11-16   Remote Authentication

For more information about remote authentication and LDAP for the DS8880, see *IBM DS8880 Integrated Copy Services Manager and LDAP Client on the HMC*, REDP-5356.

Local password rules

Many companies have their own security policies for passwords, and want to implement the same policies for their DS8880 storage system.

To implement password rules, complete these steps:

1. From the system page, click the **Settings** icon.
2. Click **Security** to open the Security window.
3. Click the Local Password Rules tab.

Figure 11-17 shows the password configuration window.
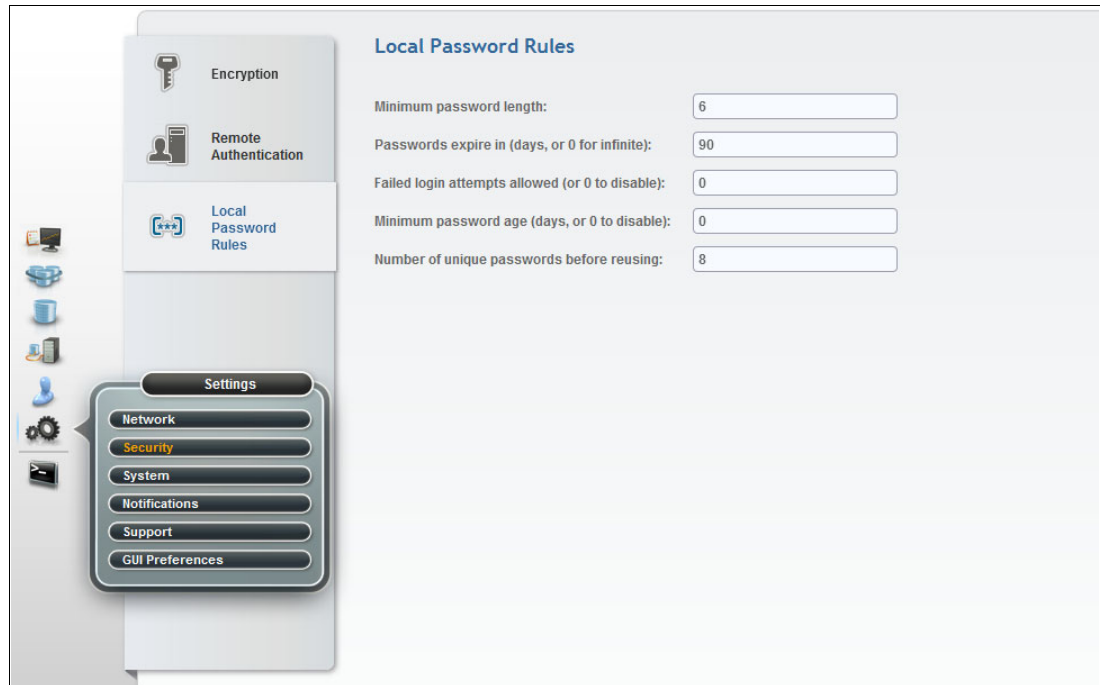
Figure 11-17   Local Password Rules

11.3.2 Easy Tier settings

No charge: Easy Tier is a no-charge feature of the IBM DS8000.

From the DS8000 GUI, you can manage Easy Tier on the system level. From the System window, select the **Settings** icon. Click **System** to open the System window, and click the Easy Tier tab to open the Easy Tier settings window, as shown in Figure 11-18.
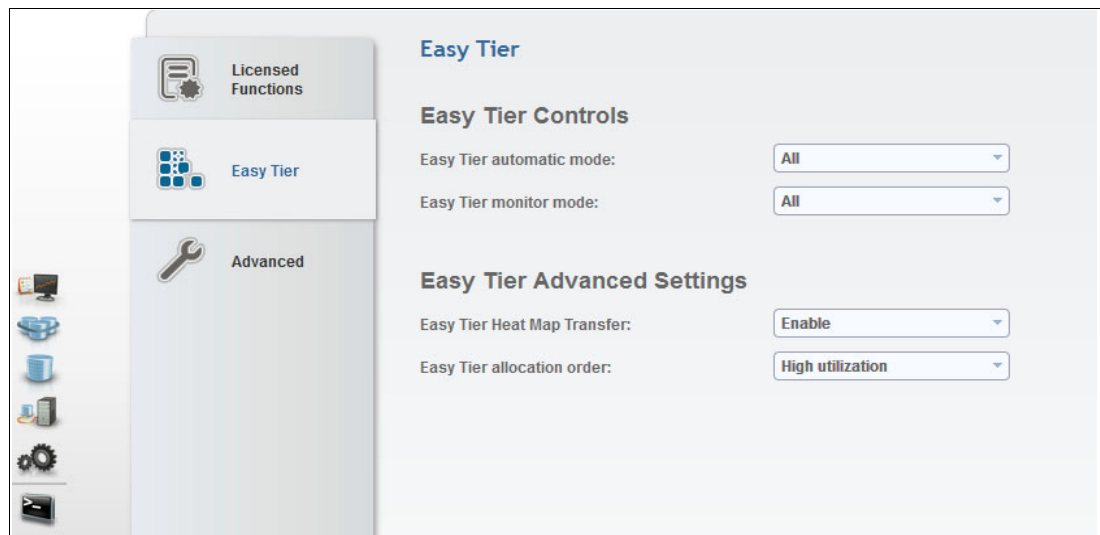
Figure 11-18   Easy Tier Settings

For more information about Easy Tier, see *IBM DS8000 EasyTier*, REDP-4667.

11.3.3 Advanced settings

By using the Advanced settings window, you can, for example, set up the I/O Priority Manager mode and the power control mode. The Advanced settings window is shown in Figure 11-9.
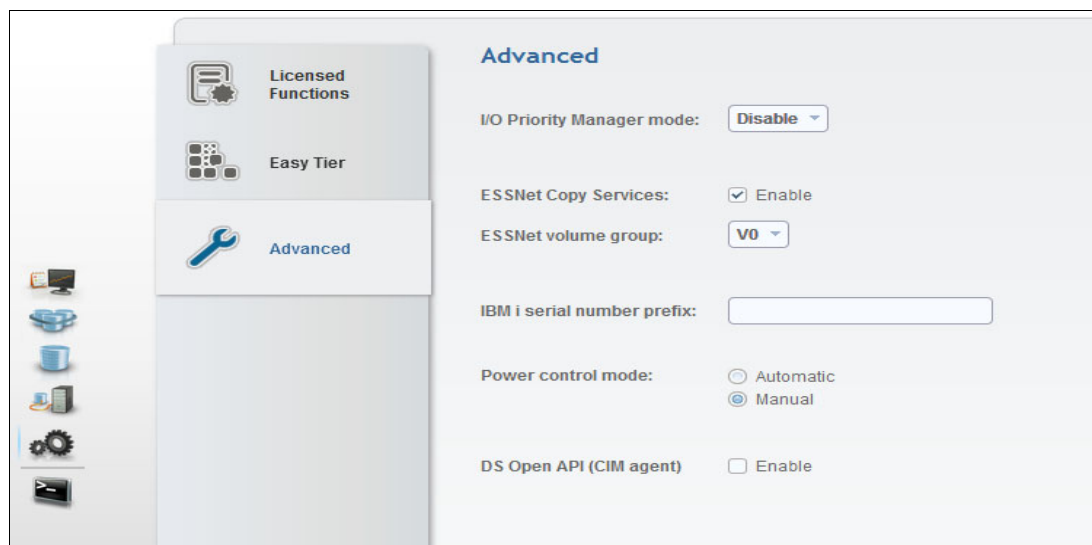
Figure 11-19   DS8880 Advanced settings

I/O Priority Manager

You can use the I/O Priority Manager to manage quality of service (QoS) levels for each application that is running on the system.

> **Note:** The I/O Priority Manager licensed function must be activated on your DS8000 storage system. The default setting for this feature is `disabled`, so it must be enabled for use.

The I/O Priority Manager aligns distinct service levels to separate workloads in the system to help maintain the efficient performance of each volume. The I/O Priority Manager detects when a higher-priority application is hindered by a lower-priority application that is competing for the same system resources. This detection might occur when multiple applications request data from the same drives. When the I/O Priority Manager encounters this situation, it delays lower-priority I/O data to assist more critical I/O data to meet its performance targets.

From the System window, click the **Settings** icon. Click **System** to open the System window. Click the Advanced tab to open the window to use to manage I/O Priority Manager mode, as shown in Figure 11-9 on page 260. You can choose I/O Priority Manager to manage or monitor capacity. You can also choose to receive Simple Network Management Protocol (SNMP) notifications about I/O Priority Manager activity.

Power control mode

You can determine how to control the power supply to the storage system. From the System window, click the **Settings** icon. Click **System** to open the System window. Click the Advanced tab to open the window to manage Power control mode (as shown in Figure 11-19 on page 272). The following options are available:

► **Automatic**: Control the power supply to the storage system through the external wall switch.

► **Manual**: Control the power supply to the storage system by using the **Power Off** action on the System window.

11.4 Logical configuration overview

Figure 11-20 illustrates the concepts of logical configuration.
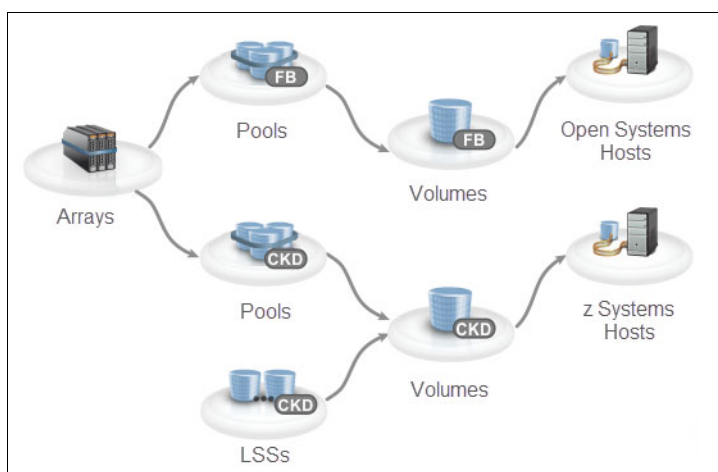
Figure 11-20   Logical configuration sequence

The following concepts apply when you use the GUI to configure the storage:

► Array: An array, which is also referred to as a managed array, is a group of storage devices that provides capacity for a pool. An array generally consists of seven or eight drives that are managed as a Redundant Array of Independent Disks (RAID).

► Pool: A storage pool is a collection of storage that identifies a set of storage resources. Pools provide the capacity and management requirements for arrays and volumes that have the same storage type, either FB for open systems or CKD for z Systems.

► Volume: A volume is a fixed amount of storage on a storage device.

► LSS: The logical subsystem (LSS) enables one or more host I/O interfaces to access a set of devices. An LSS also is known as a logical control unit (LCU) in z Systems.

► Host: The computer system that interacts with the DS8880 storage system. Hosts that are defined on the storage system are configured with a user-designated host type that enables DS8880 to recognize and interact with the host. Only hosts that are mapped to volumes can access those volumes.

The logical configuration of the DS8880 storage system begins with managed arrays.

When the storage pools are created, the arrays are assigned to the pools and then volumes are created in the pools. FB volumes are connected through host ports to an open system host. CKD volumes require that logical subsystems (LSSs) are created as well so that they can be accessed by an IBM z Systems host.

Pools must be created in pairs to balance the storage workload. Each pool in the pool pair is controlled by a processor node (either Node0 or Node1). Balancing the workload helps to prevent one node from performing most of the work and results in more efficient I/O processing, which can improve overall system performance. Both pools in the pair must be formatted for the same storage type, either FB or CKD storage. Multiple pools can be created to isolate workloads.

When you create a pool pair, all available arrays can be assigned to the pools, or the choice can be made to manually assign them later. If the arrays are assigned automatically, the system balances them across both pools so that the workload is distributed evenly across both nodes. Automatic assignment also ensures that spares and DA pairs are distributed equally between the pools.

If the storage will connect to a z Systems host, you must create the LSSs before you create the CKD volumes.

It is possible to create a set of volumes that share characteristics, such as capacity and storage type, in a pool pair. The system automatically balances the capacity in the volume sets across both pools. If the pools are managed by Easy Tier, the capacity in the volumes is automatically distributed among the arrays. If the pools are not managed by Easy Tier, it is possible to choose to use the *rotate capacity allocation method*, which stripes capacity across the arrays.

When you plan your configuration with the DS8880, all volumes, including standard provisioned volumes, consume metadata capacity when they are created, which allows the size of the reserved area to be reduced. The 1 GiB extents that are allocated for metadata are subdivided into 16 MiB subextents. The metadata capacity of each volume that is created affects the configuration planning.

Figure 11-21 shows an example of the required metadata capacity when you create volumes.
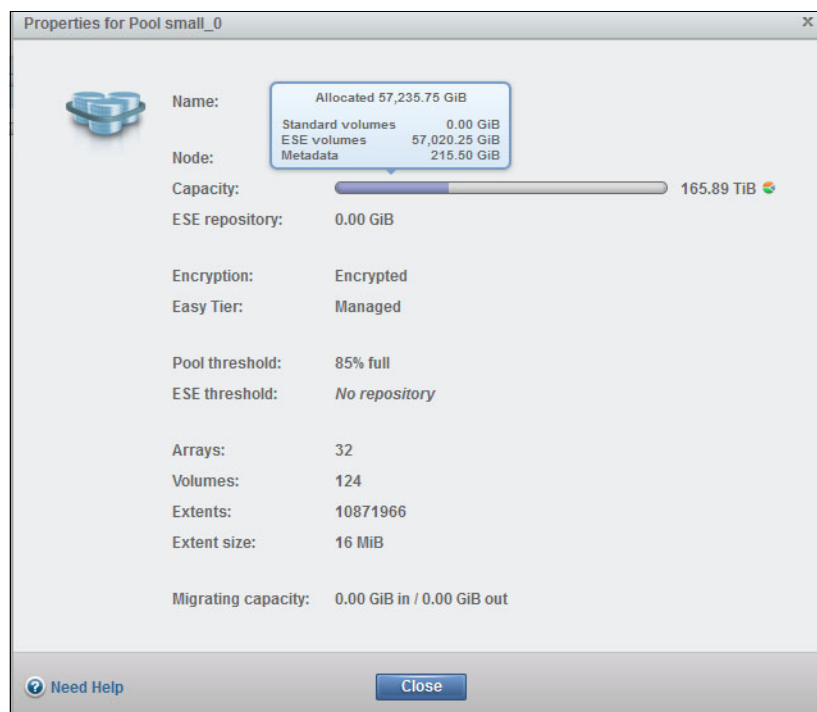
Figure 11-21   Metadata capacity that is used on volumes

If the volumes will connect to a z Systems host, the next steps of the configuration process are completed on the host. For more information about logically configuring storage for z Systems, see 11.6, "Logical configuration for count key data volumes" on page 302.

If the volumes will connect to an open system host, map the volumes to the host, and then add host ports to the host and map them to I/O ports on the storage system.

FB volumes can accept I/O only from the host ports of hosts that are mapped to the volumes. Host ports are zoned to communicate only with certain I/O ports on the storage system. Zoning is configured either within the storage system by using I/O port masking, or on the storage area network (SAN). Zoning ensures that the workload is spread correctly over I/O ports and that certain workloads are isolated from one another.

Host configuration is simplified with Release 8.2 of DS8880 microcode. Host ports are now automatically updated and mapping is allowed at volume creation of the logical configuration. In addition, host port topology can be safely changed by using the GUI and DSCLI. New host commands are available for DSCLI to make, change, delete, list, and show a host connection. For more information, see Chapter 12, "Configuration with the data storage command-line interface" on page 347.

Note: Starting with Release 8.2 of DS8880 microcode, deleting a pool with volumes is available in GUI. A warning is displayed, and the user must enter a code that is presented by the DS8880 to confirm the delete. A "Force deletion" option is also available. See Figure 11-91 on page 320.

11.5 Logical configuration for fixed-block volumes

This section describes the logical configuration for FB volumes for open systems hosts. It covers the following topics:

- ► Simple FB logical configuration flow
- ► FB pool creation
- ► Quick FB volume creation
- ► Advanced FB volume creation
- ► Creation and connection of FB volumes to open system hosts

11.5.1 Configuration flow

As with the earlier releases of DS8000 Storage Management GUI, the logical configuration for FB for open systems is greatly simplified, and can now be accomplished with a few steps. Figure 11-22 shows the logical configuration flow for FB.

The following steps are used for logical configuration for FB volumes:

1. Create a pool pair.
2. Create the FB volumes.
3. Attach to open system hosts.
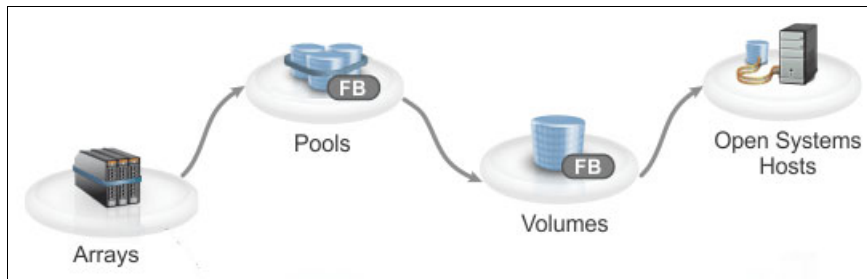
Figure 11-22   Logical configuration flow for open systems

11.5.2 Creating FB pools

For best performance and a balanced workload, create two pools. The Storage Management GUI helps the system administrator to create a balanced configuration by creating pools as a pair. The pools are configured so that one pool of the pair is managed by node 0 and the other pool of the pair is managed by node 1.

> **Note:** If the requirement is to create a single pool, see "Creating a single pool" on page 281.

To create an FB pool pair, complete these steps:

1. From the System window, click the **Pools** icon.

2. Click **Arrays by Pool** to open the Array by Pool window, as shown Figure 11-23.
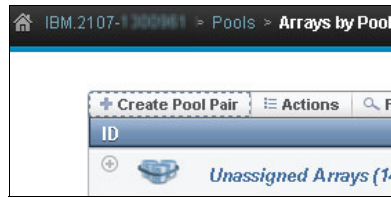
Figure 11-23   Create Pool Pair tab

> **Note:** You can automatically assign arrays when creating a pool pair. The arrays are created with the default raid type, RAID 6 for the new DS8880 models (984, 985, 986, and 988) and RAID 5 for previous models (980, 981, and 982). To configure other supported raid types, use the advanced configuration for pool creation, or assign the arrays manually to an existing storage pool from the unassigned arrays. For more information, see "Creating FB Pools: Advanced Configuration" on page 278.

3. Click the Create Pool Pair tab. The **Create Pool Pair** window opens.

4. Specify the pool pair parameters, as shown in Figure 11-24:

   – Storage type: Ensure that **Fixed Block (FB)** is selected.

   – Name prefix: Add the pool pair name prefix. A suffix ID sequence number is added during the creation process.
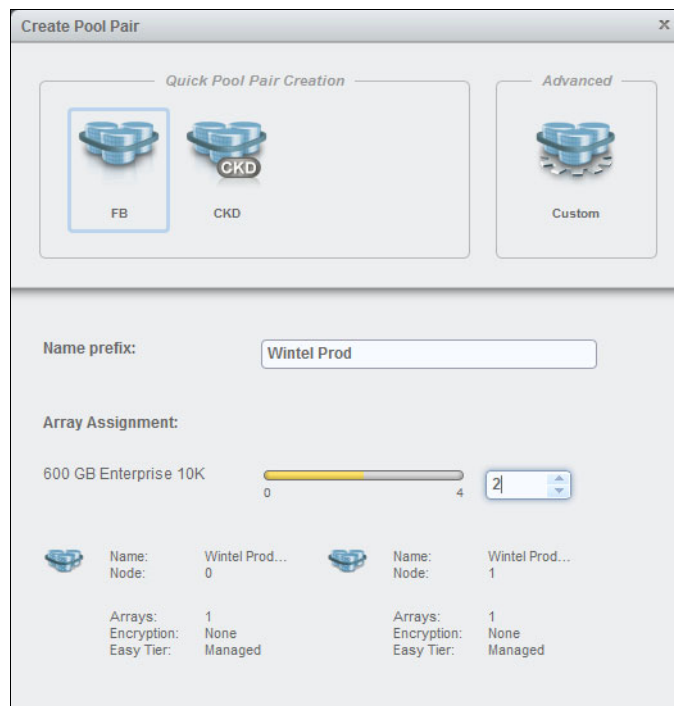
Figure 11-24   Create FB pool pair and assign arrays to a pool pair

5. Select from the listed drive types and select the number of arrays for each drive type that you want to assign to the pool pair.

> **Important:** The number of specified arrays must be even. Trying to specify an odd number results in a message that states "`Arrays must be spread evenly across the pool pair`". The GUI increases the number of arrays by one to achieve an even number.

6. When pool pair parameters are correctly specified, click **Create** to proceed. Figure 11-25 shows a pool pair that is created and assigned arrays.

ID		State	Capacity	Drive Capacity	Drive Class
⊕	Unassigned Arrays (2)				
⊕ CKD	Large_ext_P0_0			16.06 KMod1 in Node 0	
⊕ CKD	Large_ext_P1_1			16.04 KMod1 in Node 1	
⊕ CKD	Small_ext_P2_2			15.35 KMod1 in Node 0	
⊕ CKD	Small_ext_P3_3			15.35 KMod1 in Node 1	
⊖	Wintel Prod_4			3.60 TiB in Node 0	
	MA6	✓ Normal	3,688.00 GiB	600 GB	Enterprise 10K
⊖	Wintel Prod_5			3.60 TiB in Node 1	
	MA5	✓ Normal	3,688.00 GiB	600 GB	Enterprise 10K

Figure 11-25   Pool pair that is created and assigned arrays

Creating FB Pools: Advanced Configuration

When creating FB storage pools, you can also specify the extent size and RAID level. With Release 8.1 of DS8880 microcode and later, you can specify large 1 GiB, or small 16 MiB extents for a storage pool or pool pair. The RAID level for the arrays can also be specified at creation time by using the advanced (Custom) configuration option.

To create a custom FB pool pair, complete these steps:

1. From the System window, click the **Pools** icon.

2. Click **Arrays by Pool** to open the Array by Pool window, as shown previously in Figure 11-23.

3. Click the **Create Pool Pair** tab. The Create Pool Pair window opens.

4. Specify the pool pair parameters, as shown in Figure 11-26 on page 279:

   – Storage type: Ensure that **Fixed Block (FB)** is selected.

   – Select the **Custom** option in the Advanced section.

   – Name prefix: Add the pool pair name prefix. A suffix ID sequence number is added during the creation process.

   – Extent size: Select **1 GiB** for large extents, or **16 MiB** for small extents.

5. Select from the listed drive types and select the number of arrays for each drive type that you want to assign to the pool pair.

6. Choose the RAID level for the selected arrays. See Figure 11-26 for all DS8880 Models 980, 981, and 982. For DS8880 Models 984, 985, 986, and 988, if RAID 5 is selected, you need to accept the disclaimer that appears. See Figure 11-27 on page 280. A time stamped record with user information is created for audit purposes.

> **Note:** With new DS8880 Models 984, 985, 986, and 988, the default RAID type is RAID 6. RAID 5 is allowed by using advanced configuration for supported drive types less than 1 TB.
>
> RAID level 5 for drive types greater than 1 TB is not recommended and is not allowed. If your environment requires RAID 5, contact IBM. However, you will have to acknowledge your understanding of the risk with RAID 5, as shown in Figure 11-27 on page 280.
>
> For more information about the supported drive types and available RAID levels, see Chapter 2, "IBM DS8880 hardware components and architecture" on page 23.

7. When pool pair parameters are correctly specified, click **Create** to proceed.

Figure 11-26   Configure FB storage pools with Advanced Options

Figure 11-27 shows configuring FB storage pools with RAID 5.
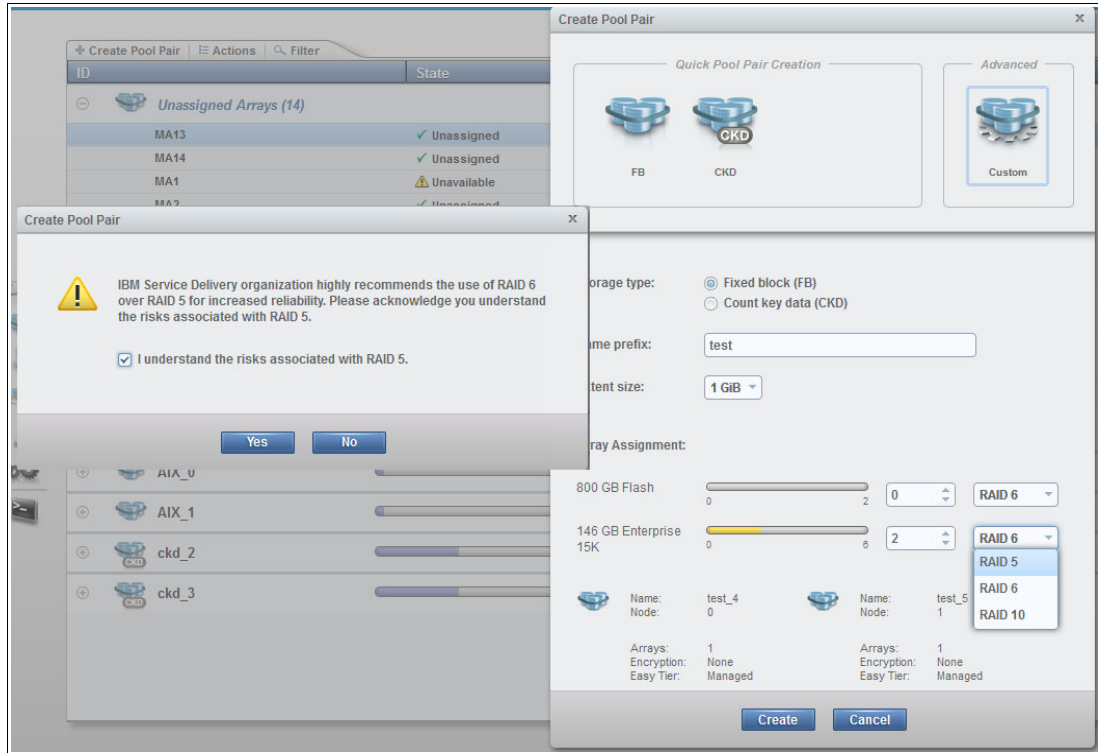
Figure 11-27   Configure FB storage pools with Advanced Options RAID 5

Manually assigning arrays to existing pools

The storage administrator can manually unassign arrays or reassign assigned arrays to existing pools when the current configuration needs to be modified, such as when the administrator adds storage capacity. To manually assign arrays, complete these steps:

1. Select an array and click **Assign** or **Reassign**, as shown in Figure 11-28. This action opens the Assign Array window.
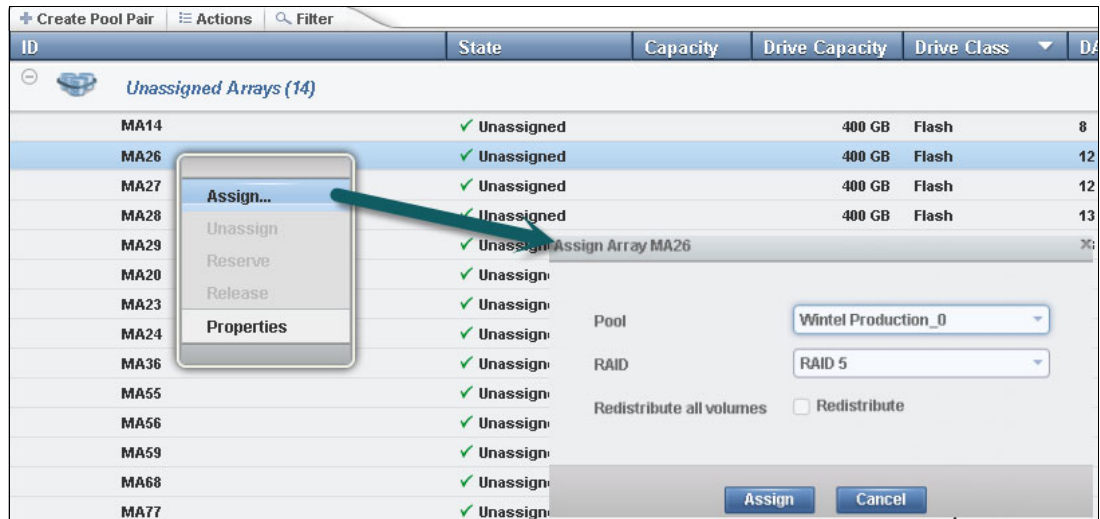
Figure 11-28   Manually assign an array to an existing pool

2. Select the target pool from the drop-down list, and the RAID level that you want.

3. Select the **Redistribute** check box to redistribute all existing volumes across the pool, including the new array.

4. Click **Assign**.

> **Note:** In a pool that is managed by Easy Tier, redistributing volumes across the pool is automatic. This redistribution is called *Dynamic Pool Reconfiguration*. For more information, see *IBM DS8000 EasyTier*, REDP-4667.

Creating a single pool

Occasionally, you are required to create a single pool, as opposed to creating a pool pair for balancing a workload. To create a single storage pool, complete these steps:

1. Create a pool pair, as shown in Figure 11-29. However, do not assign any arrays to new pool pair.
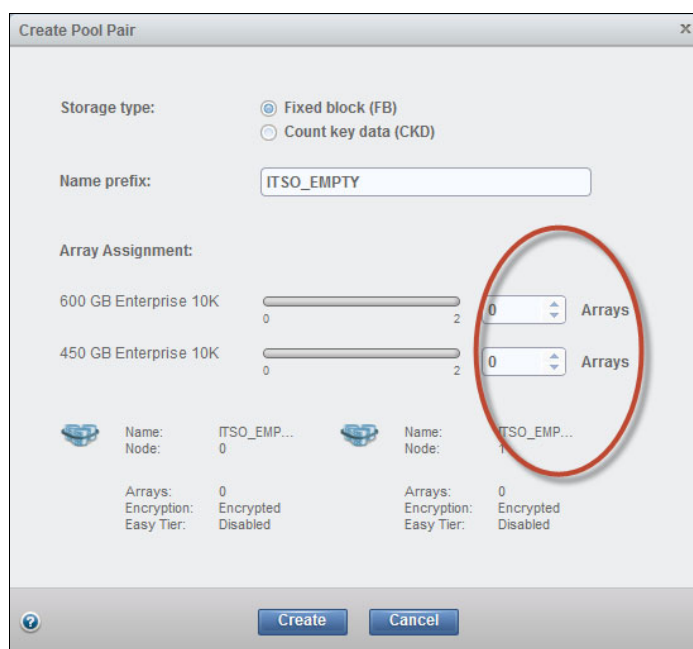
Figure 11-29   Create an empty pool pair with no assigned arrays

2. Click **Create**.

3. Choose one of the pools from the recently created pool pair to delete, as shown in Figure 11-30.
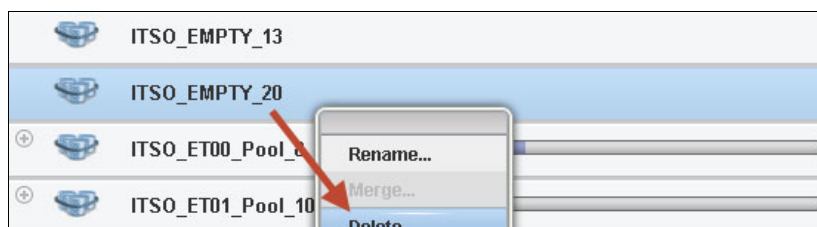
Figure 11-30   Delete one pool of the pool pair

4. Assign one or more arrays to the single pool, as shown in Figure 11-31.
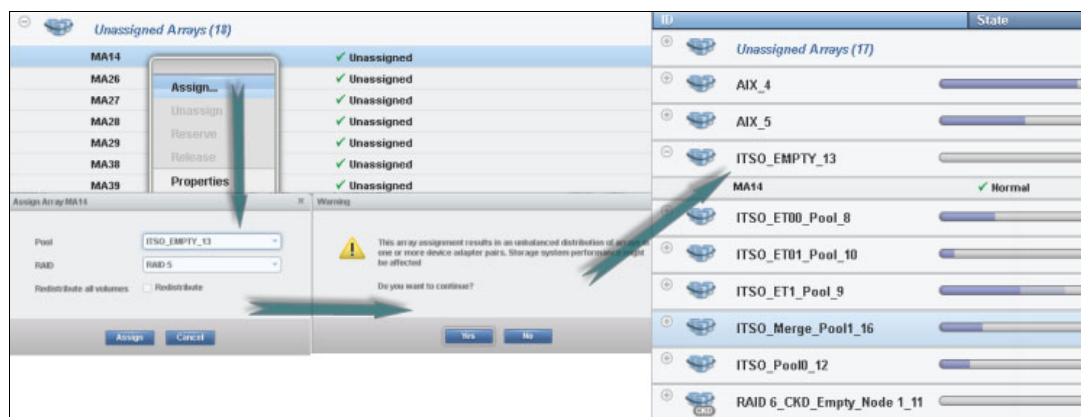
Figure 11-31   Assign an array to a single pool

11.5.3 Creating FB volumes

Create a set of FB volumes by using the Create Volumes tab on either the Volumes window or the Volumes by Pool window. The maximum capacity for an FB volume is 16 TiB. The Storage Management GUI automatically distributes capacity across the two pools.

From the system page, select the **Volumes** icon. Four options are provided as shown in Figure 11-32:

► **Volumes** (All volumes are visible in single view.)
► **Volumes by Pool** (Volumes are grouped by pool.)
► **Volumes by Host** (Volumes are grouped by host.)
► **Volumes by LSS** (Volumes are grouped by LSS.)

Figure 11-32   Create FB volumes

Select any option to open the Create Volumes window (Figure 11-33).

The Storage Management GUI provides two options to create FB volumes:

► Quick Volume Creation

The intent of the Quick Volume Creation option is to simply create volumes with standard provisioning that are balanced across pool pairs.

► Advanced

This option is used for custom configuration when you either need to create thin-provisioned volumes, specify the extent allocation method (EAM), or select specific LSSs.

Create FB volumes: Quick volume creation

To used the Quick Volume Creation method, complete these steps:

1. On the Create Volumes window, click the **FB** icon, under Quick Volume Creation, as shown in Figure 11-33.
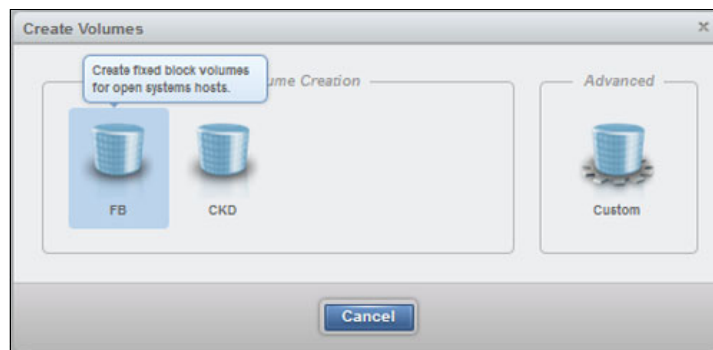
Figure 11-33   Create FB volumes by using the Quick Volume Creation option

2. The Create Volumes configuration window is displayed. By default, the Storage Management GUI tries to balance the volumes across the pool pair so that the workload is balanced across both nodes or central processor complex (CPC).

3. Enters the following user-specified values:

   – Name prefix: User-defined name for volumes (a suffix ID sequence number will be added during the creation process).

   – Quantity: Number of volumes to be created in selected pools.

   – Capacity: The capacity of the volumes to be created. Volumes can be configured in the following increments:

     • GiB or TiB.
     • Blocks.
     • IBM System i volumes.

Note: IBM System i volumes are chosen from a list of fixed capacities. Currently, the GUI is limited because it cannot create variable-sized IBM i volumes.

In the example that is shown in Figure 11-34, 20 volumes of 55 GiB are created. Both pools of the pool pair are selected by default. The system allocates 10 volumes to each pool, balancing the workload. Select a host or cluster from the list to assign the volumes to it.

Note: With Release 8.2 of DS8880 microcode, you can create volumes and map them to a target host or cluster in one configuration step.
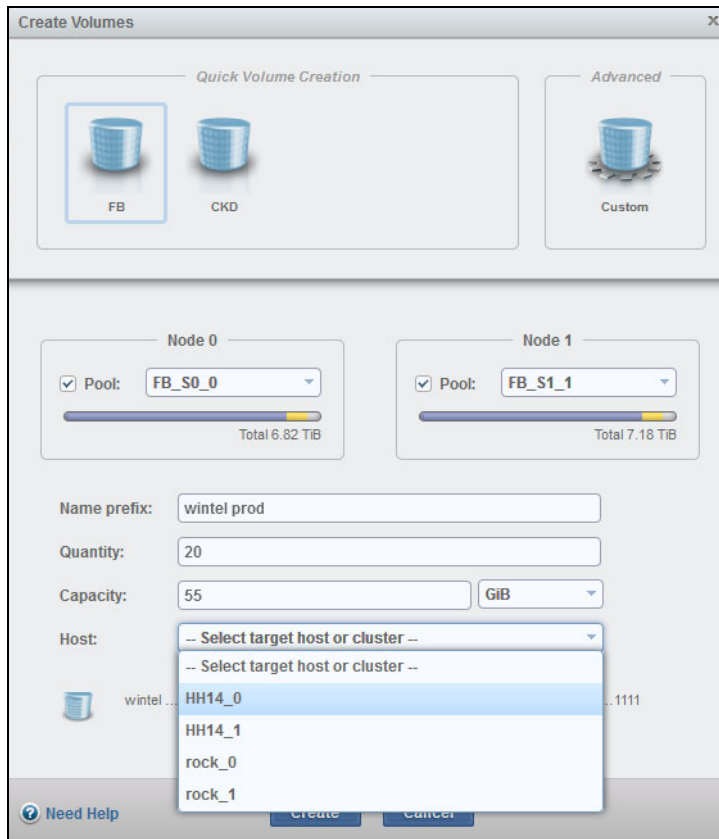
Figure 11-34   FB volume creation with host mapping window

Create FB volumes: Advanced configuration

To use the advance configuration option, complete these steps:

1. From the **Create Volumes** window, select **Advanced** (Custom), as shown in Figure 11-35.
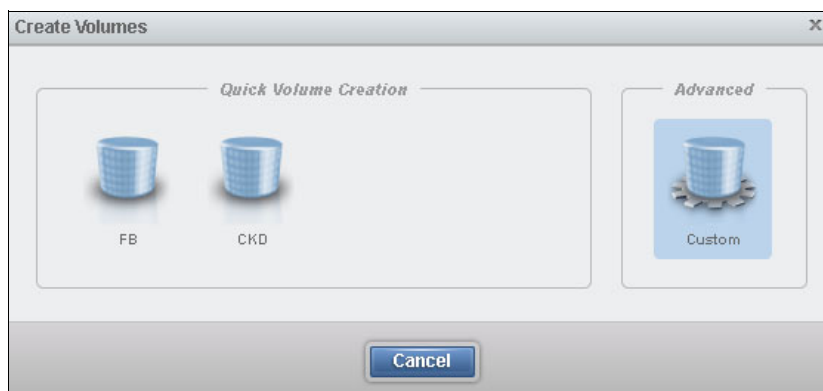
Figure 11-35   Create FB Volumes: Advanced

2. The Type window appears, as shown in Figure 11-36. By default, the Storage type that is selected is **Fixed block (FB)**. For FB volume creation, two volume definition modes are available:

   – Volume creation by volume sets.
   – Volume creation by LSS.
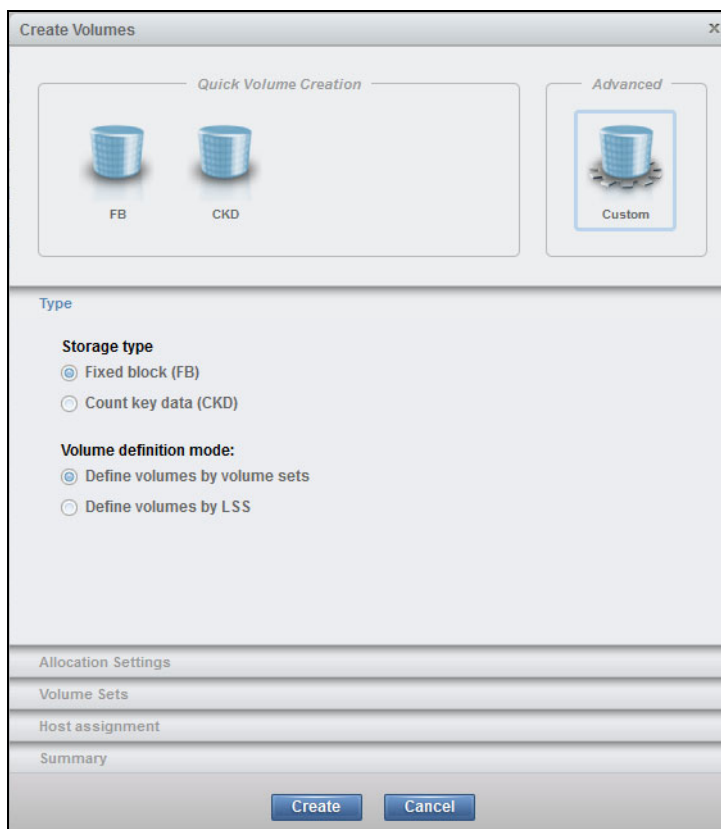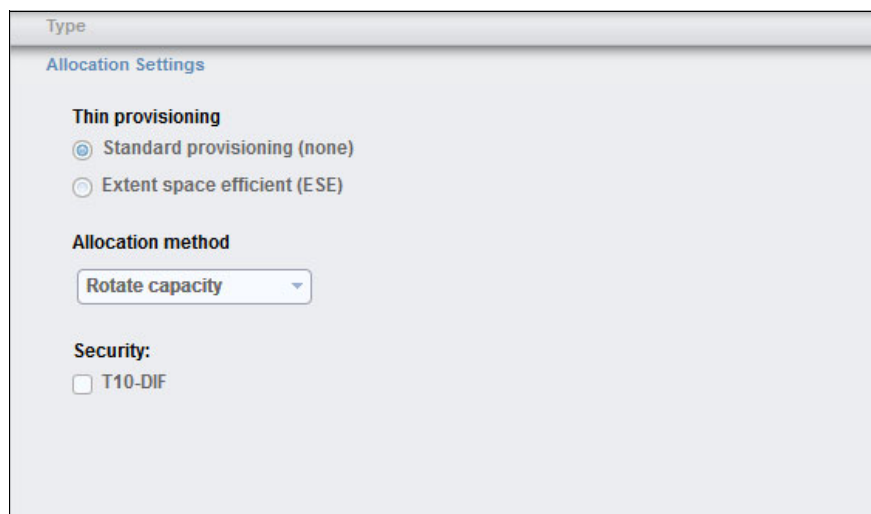
Figure 11-36   Advanced: Define FB volumes by volume sets

3. Select the option that you want, then click **Allocation Settings** to open the Allocation Settings window that is shown in Figure 11-37.

Figure 11-37   Advanced: FB volume creation: Allocation method

4. Select the type of thin provisioning (**Standard provisioning (none)** or **Extent space efficient (ESE)** for FlashCopy.

5. If you select **Standard provisioning** or **Extent space efficient**, you can then select the extent allocation method (**Rotate capacity** or **Rotate volumes**).

6. For FB volumes, you can enable the T10 Data Integrity Feature (DIF/PI). You can enable the T10 - data integrity feature (DIF) only for FB volumes.

Note: When FlashCopy is used on FB volumes, the source and the target volumes must have the same protection type of either T10-DIF or standard.

7. If the volume definition mode was set to **Define volume by volume sets**, click the Volume Sets tab to display the Volume Sets window, as shown in Figure 11-38, and specify the following information:

   – Pool selection: By default, a pool pair is selected to balance the workload across the nodes.

   – Volume characteristics: The storage administrator defines the following user-determined values:

     • Name prefix: User-defined name for volume set. (A suffix ID sequence number will be added during the creation process).

     • Quantity: Number of volumes to create.

     • Capacity: The capacity of the volumes to create. Volumes can be configured in increments of GiB, TiB, blocks, or IBM System i volumes.

Figure 11-38   Advanced: FB volume creation: Volume sets

8. If the volume definition mode was set to **Define volume by LSS**, click the Volumes by LSS tab to display the Volumes by LSS window, as shown in Figure 11-39 on page 288, and specify the following information:

– Pool selection: By default, a pool pair is selected to balance the workload across the nodes.

– LSS range: The LSS IDs to create. If only a single LSS is required, make the range a single ID. For example, LSS range 30 - 30 creates a single LSS.

– Volume characteristics: The storage administrator defines the following user-determined values:

• Name prefix: User-defined name for volume in the LSSs. A suffix ID sequence number is added during the creation process.

• Quantity: Number of volumes to create for each LSS.

• Capacity: The capacity of the volumes to create. Volumes can be configured in increments of GiB, blocks, or IBM System i volumes.

Figure 11-39   Advanced: FB volume creation: Volumes by LSS

9. In the Host Assignment section, select a host from the list to map the volumes to it. This is an enhancement with Release 8.2 of DS8880 microcode as shown in Figure 11-40.

If wanted, volumes can be attached to hosts manually at the later time as explained in 11.5.5, "Assigning FB volumes" on page 301.
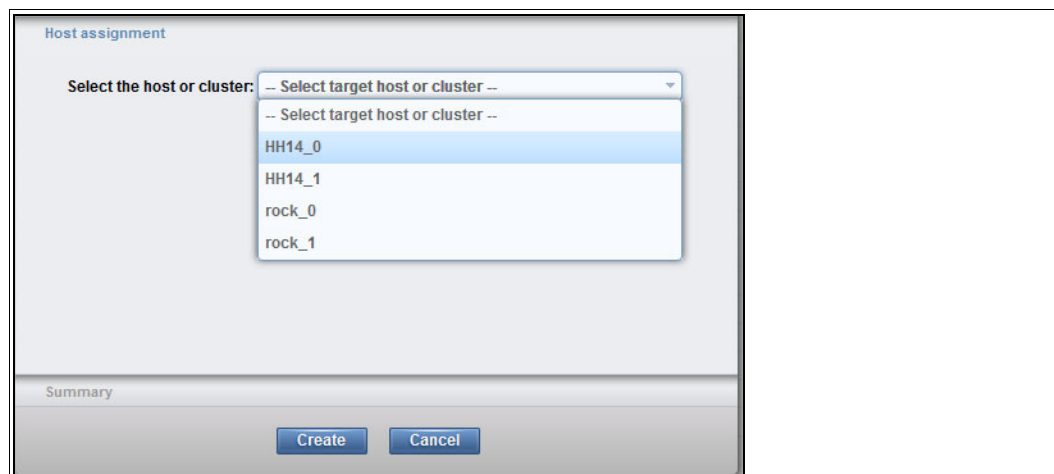
Figure 11-40   Advanced FB Volume by Volume Set Host Mapping

10. Review the summary, as shown in Figure 11-41, and then click **Create**.
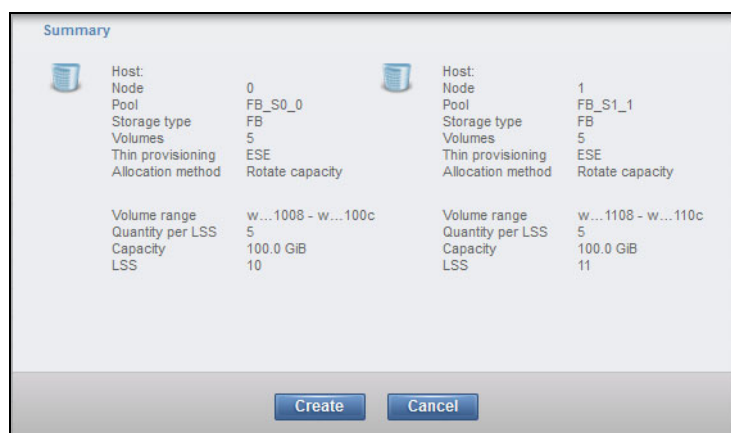
Figure 11-41   Advanced FB volume by volume set summary

11.5.4 Creating FB host attachments

This section describes the steps that are required to attach FB volumes to open system hosts. You use this sequence:

1. Setting I/O port topology
2. Creating open systems clusters
3. Creating open systems hosts
4. Assigning host ports to hosts
5. Assigning FB volumes to open systems hosts

Setting I/O port topology

For an open system host to access FB volumes that are configured on the DS8880, the host must be connected to the DS8880 through a Fibre Channel connection (FICON). The Fibre Channel Protocol (FCP) must be configured on the I/O port so that the host can communicate.

DS8880 has two kinds of host adapters (HAs): 4-port 16 Gbps and 4-port or 8-port 8 Gbps Fibre Channel (FC) adapters. Each port can be independently configured to one of the following Fibre Channel topologies:

► Fibre Channel Arbitrated Loop (FC-AL): Only for 8 Gbps host adapters

► FCP: Also known as *Fibre Channel-switched fabric* (which is also called *switched point-to-point*) for open system host attachment, and for Metro Mirror, Global Copy, Global Mirror, and Metro/Global Mirror connectivity

► FICON: To connect to z Systems hosts, and for z/OS Global Mirror connectivity

Note: FC-AL is not supported by 16 Gbps adapters.

For open system hosts, the I/O port must be configured as either FCP or FC-AL. To set the I/O port topology, complete these steps:

1. From the Storage Management GUI System window, click the **Actions** menu and select **Modify I/O Port Protocols** to open the Modify I/O Port Protocols window (Figure 11-42).
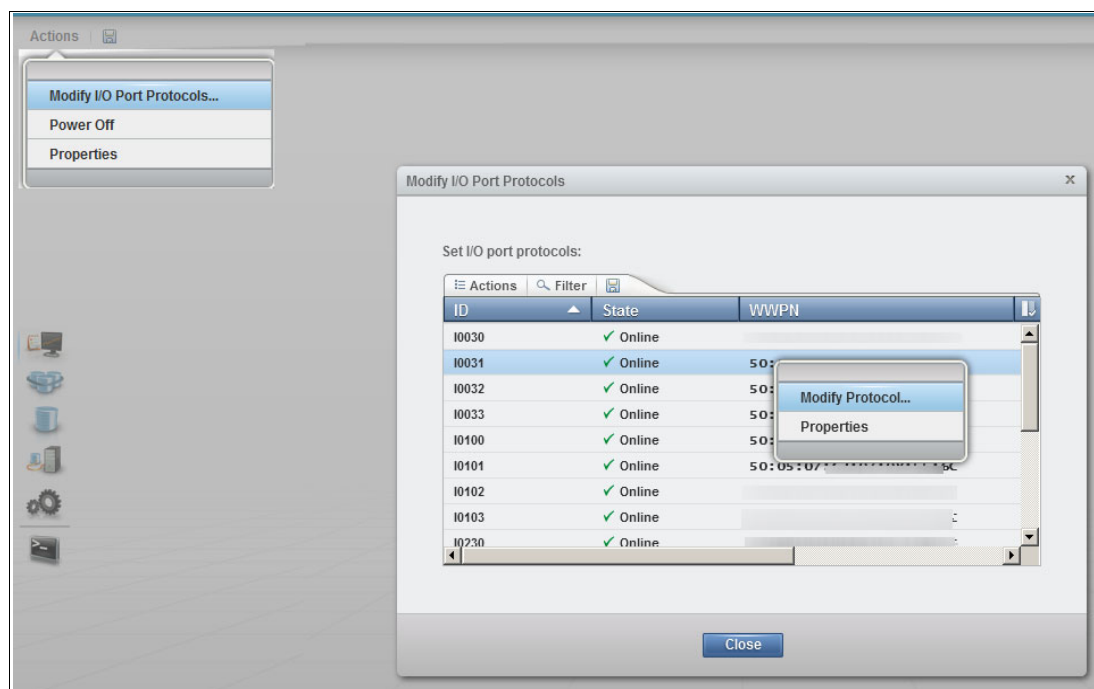
Figure 11-42   Modify I/O Port Protocols window

2. Select the port to modify. Multiple ports can be selected by using the **Shift** or **Ctrl** key.

3. From the Actions tab, click **Modify Protocol** to open the Modify Protocol window, as shown in Figure 11-43.
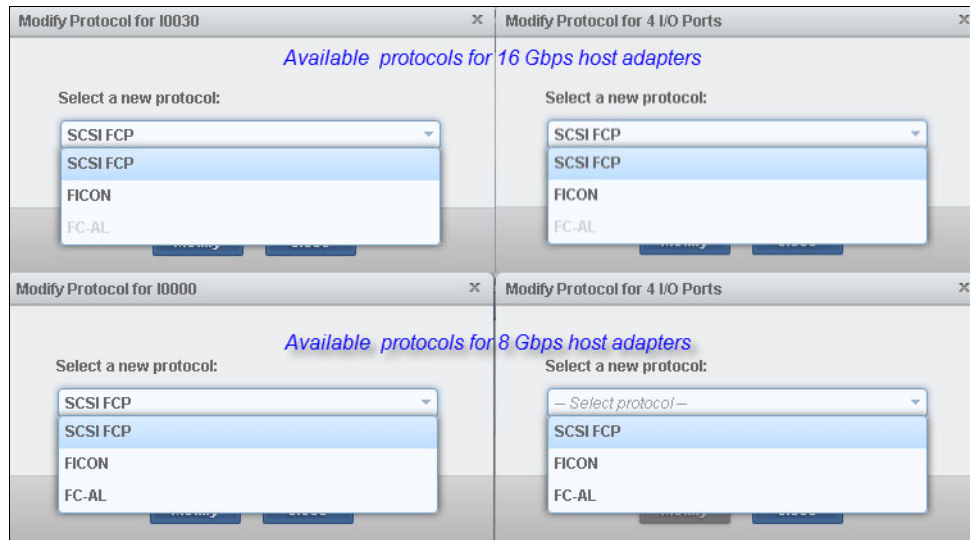
Figure 11-43   Modify I/O Port: Select a new protocol: Single or multiple 16 and 8 Gbps ports

Figure 11-43 is divided into two main sections:

– Available protocols for 16 Gbps host adapters (for single or multiple ports that are selected)

– Available protocols for 8 Gbps host adapters (for single or multiple ports that are selected)

For open system host attachment, select either one of the following items:

– Small Computer System Interface (SCSI) FCP: For open system hosts that connect to a SAN

– FC-AL: For open system hosts that are connected directly to the DS8880 (only for 8 Gbps)

4. Click **Modify** to perform the action.

Creating open system clusters and hosts

Starting with DS8880 Release 8.1 of the microcode, hosts of the same type can be grouped into a cluster for simplified volume mapping operations and management. Private volume mappings are supported, along with support of all cluster events and notifications. This section describes how to configure clusters and hosts by using the Storage Management GUI. For reference, a host port is the Fibre Channel port of the HBA Fibre Channel adapter that is installed on the host system. An I/O port is the Fibre Channel port of the HA that is installed in the DS8880.

To configure an open system cluster, complete the following steps:

1. Create a cluster: Configure a cluster object to access the storage system.

2. Create a host: Configure a host object to access the storage system.

3. Assign hosts: Assign hosts to a cluster object.

4. Assign a host port: Assign a host port to a host object by identifying one of the worldwide port names (WWPNs) of the HBA that is installed on the host system.

5. Modify the I/O port mask: Modify the I/O port mask to allow or disallow an I/O port that the host can use to communicate with the DS8880.

To configure an open system host, complete the following steps:

1. Create a host: Configure a host object to access the storage system.

2. Assign a host port: Assign a host port to a host object by identifying one of the WWPNs of the HBA that is installed on the host system.

3. Modify the I/O port mask: Modify the I/O port mask to allow or disallow an I/O port that the host can use to communicate with the DS8880.

Creating clusters

To configure a cluster object, complete these steps:

1. Click the **Hosts** icon in the Storage Management GUI System window.

2. Click **Hosts** as shown in Figure 11-44.

Figure 11-44   Configure hosts

3. The Hosts window opens. Click the **Create Cluster** tab as shown in Figure 11-45.
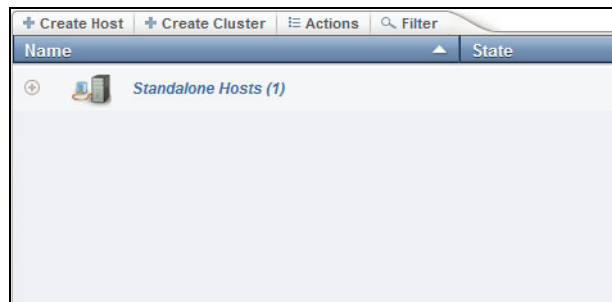
Figure 11-45   Create Cluster

4. The Create Cluster window opens, as shown in Figure 11-46. Specify the name of the cluster, and click **Create**.
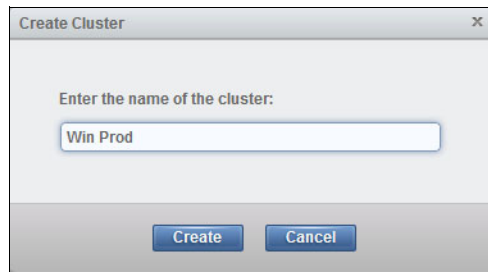
Figure 11-46   Create Cluster window

Creating hosts

To configure a host object, complete these steps:

1. Click the **Hosts** icon in the Storage Management GUI System window.

2. Click **Hosts** as shown in Figure 11-47.

Figure 11-47   Configure hosts

3. With Release 8.2 of DS8880 microcode, the host ports are automatically updated. If there are any detected unassigned host ports, a suggested task window opens with two options as shown in Figure 11-48.

> **Note:** This window always appears whenever definitions in the DSCLI have been made, which are not yet fully reflected in the GUI. So the GUI offers to move over the CLI definitions fully into the GUI.
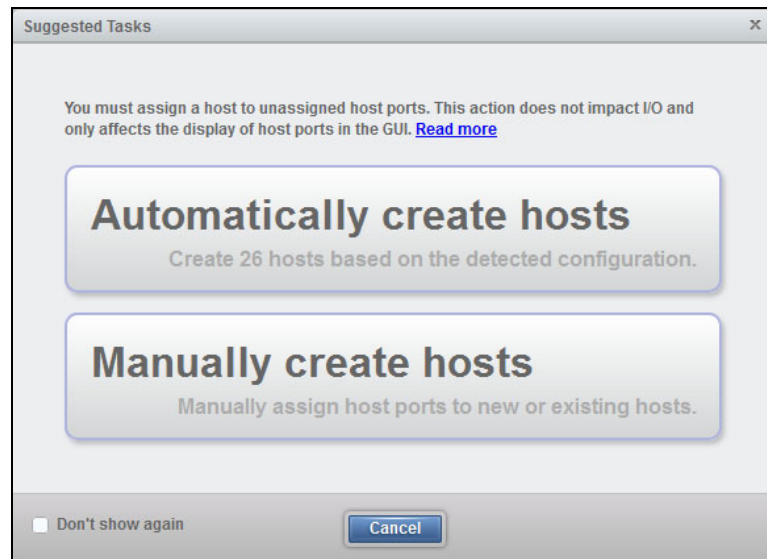
Figure 11-48   Hosts Suggested Tasks

> **Tip:** If canceled or closed, the Suggested Tasks window can be reopened by clicking the attention message as shown in Figure 11-49.
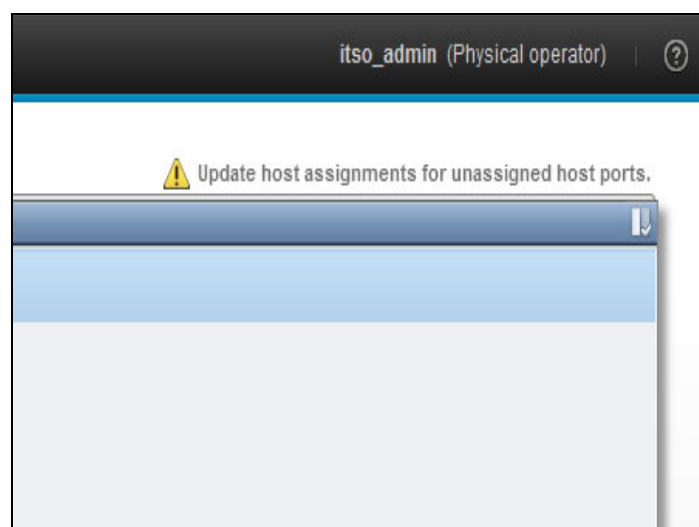
Figure 11-49   Update host assignment attention

4. Click **Automatically Create Hosts**. A list of detected hosts and unassigned host ports are displayed as shown in Figure 11-50. Verify the list and click **Create** to complete the assignment task automatically.
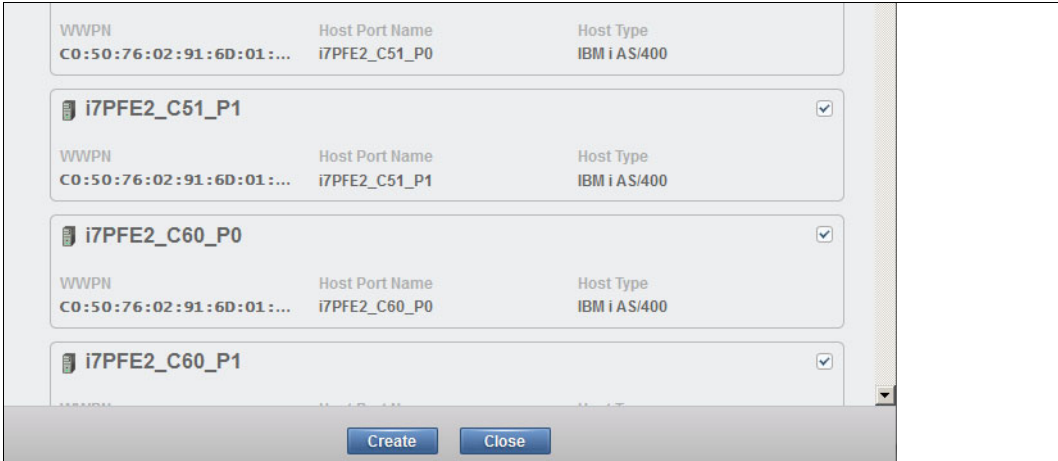
Figure 11-50   Update Unassigned Host Ports

5. Select **Manually Create Hosts** to manually assign host ports to the existing hosts or click **Create Host** to create hosts to assign the ports (Figure 11-51).
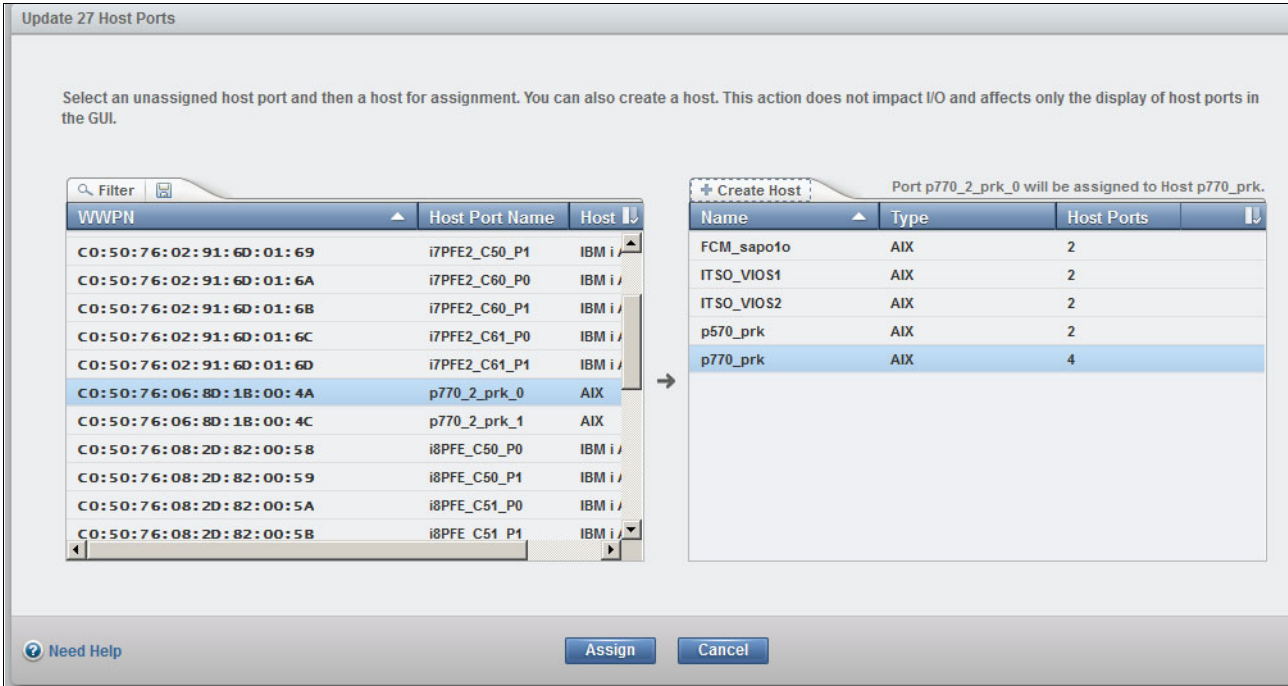
Figure 11-51   Manual Host Ports Update

6. In the Create Host window, click the **Create Host** tab, as shown in Figure 11-52.
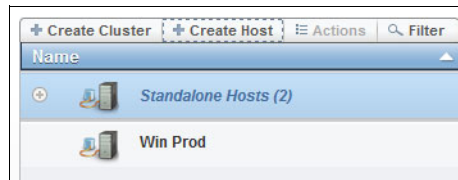
Figure 11-52   Create Host

7. The Add Hosts window opens, as shown in Figure 11-53. Specify the following items:

   – Name: The user-defined name for the host to add.

   – Type: This item represents the operating system of the host to add. Figure 11-53 shows the available host types.
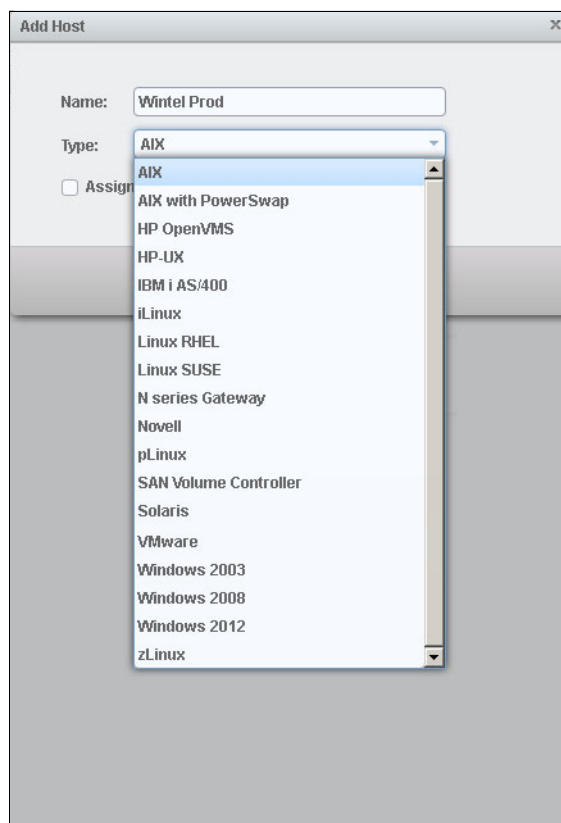
Figure 11-53   Add Host window that shows all available host types

Assigning hosts to a cluster

If required, after a host is created, it can be assigned to a defined host cluster. To do so, complete these steps:

1. From the Hosts window, select the host to be assigned to a cluster.

2. Either right-click, or from the **Actions** tab, select **Assign to Cluster**, as shown in Figure 11-54.
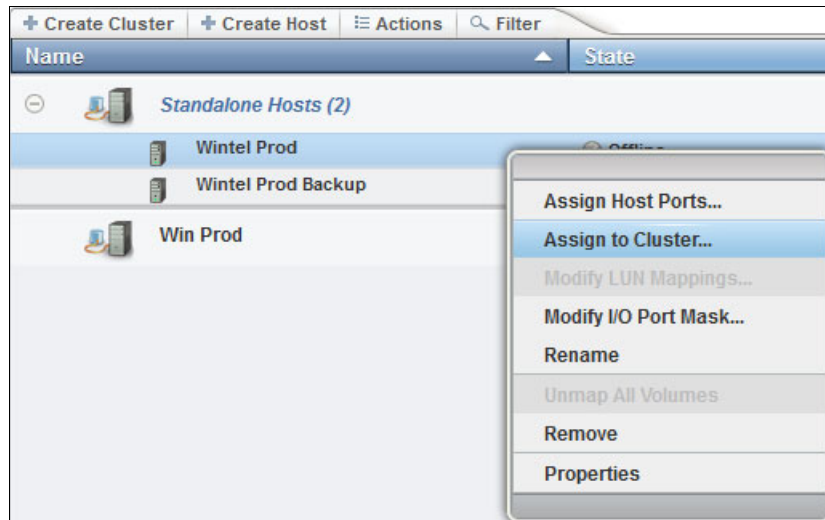
Figure 11-54   Assign to Cluster

3. The Assign Host window opens as shown in Figure 11-55. From the drop-down list, select the cluster to add the host to.
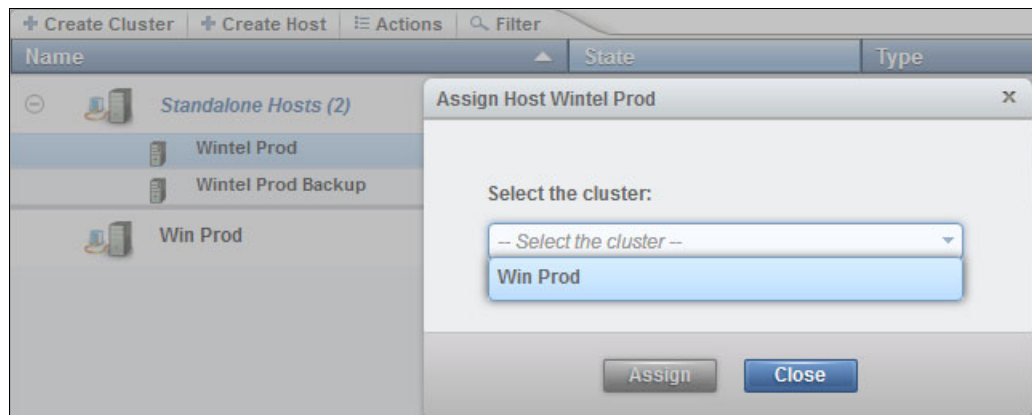
Figure 11-55   Assign Host window

4. Click **Assign** to complete the action.

5. Repeat the previous actions to add all hosts that are required in the cluster. After you complete all the hosts, assigned hosts will be listed under the cluster in the Hosts window as shown in Figure 11-56.
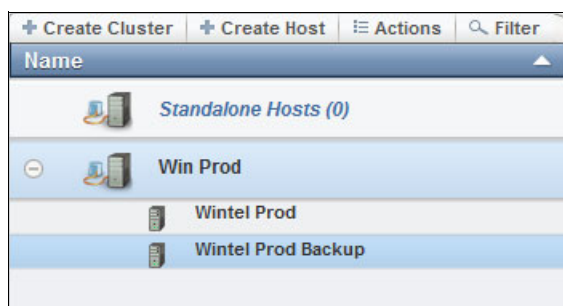
Figure 11-56   Hosts window showing the created cluster

Assigning host ports

After the host is added, host ports must be assigned to the defined host. To do so, complete these steps:

1. From the Hosts window, select the host to which to assign the host port. Either right-click, or from the **Actions** tab, select **Assign Host Port**, as shown in Figure 11-57.
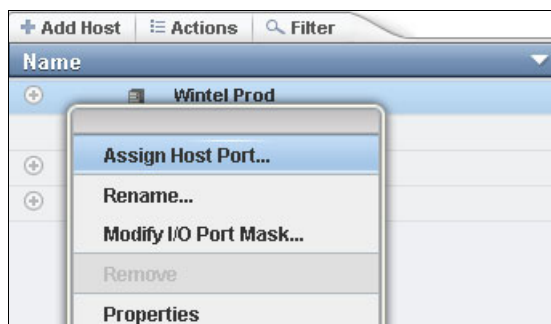
Figure 11-57   Assign Host Port

2. This action opens the Assign Host Port to Host window. If the host that is being added is effectively connected to the DS8880, the available WWPNs are listed in the drop-down. If the host that is being added is not connected to the DS8880, the WWPN of the host HBA must be added manually.

Select from one of the WWPNs that are shown in the drop-down list in Figure 11-58, or manually enter the WWPN of the HBA of the host that you are adding. Click **Assign**.
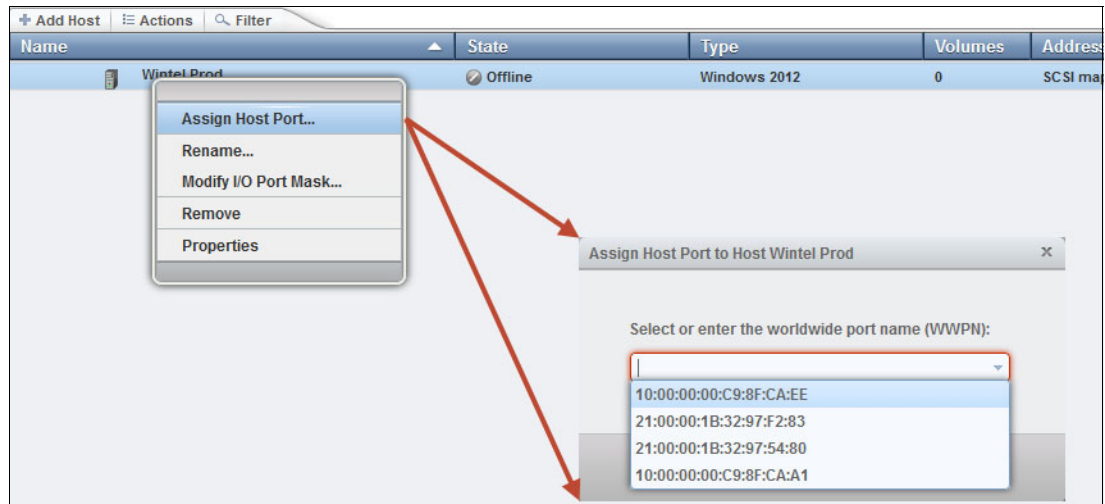
Figure 11-58   Add a host port to an existing host

Typically, most open system hosts have multiple FC connections to the DS8880 for redundancy and performance. Ensure that all additional host WWPNs that are connected for this host are defined to the host object by using the same procedure.

Modifying the I/O port mask

When the host is configured, by default it has access to all I/O ports. To see this property, right-click a host object and select **Properties**, as shown in Figure 11-59.

Figure 11-59   Host properties with all I/O ports that are allowed for a host

If the system administrator wants to restrict the I/O ports that can communicate with the host, *I/O port masking* needs to be defined. Modify the I/O port mask to allow or disallow an I/O port that the host can use to communicate with the DS8880. To do so, complete these steps:

1. From the Hosts window, select the host to modify the I/O port mask. Right-click, or from the Actions tab, select **Modify I/O Port Mask**.

2. A list of the DS8880 I/O ports is displayed. Select one or more ports to disallow. You can use the **Ctrl** or *Shift* key for multiple selections.) Click **Save**. See Figure 11-60.
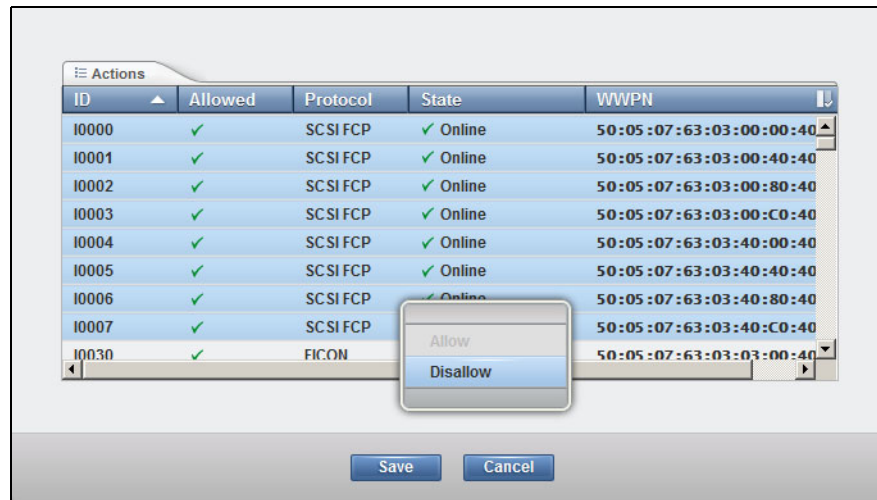
Figure 11-60   Modify I/O Port Mask: Disallow ports

The properties of the selected host now reflect the number of I/O ports that have access, as shown in Figure 11-61.
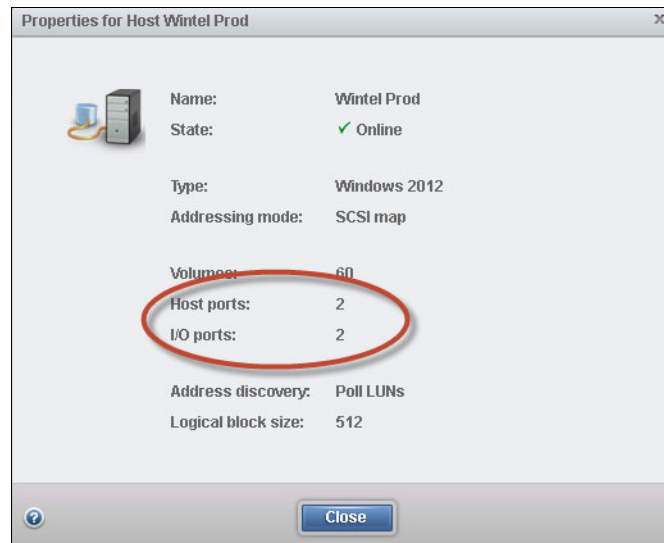
Figure 11-61   Host properties with I/O port masking

I/O ports can be selectively modified from disallowed to allowed by using the previous procedure.

11.5.5 Assigning FB volumes

FB volumes can accept I/O only from the host ports of hosts that are mapped to FB volumes. To map a volume to a host or a cluster, complete these steps:

1. From the Storage Management GUI system page, select **Volumes by Host** from either the **Volumes** icon or the **Hosts** icon, as shown in Figure 11-62.
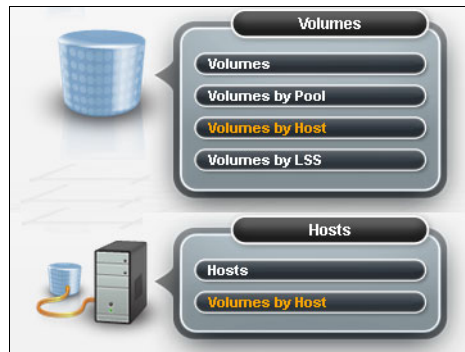
Figure 11-62   Volumes by Host from the Volumes menu or the Hosts menu

2. The Volumes by Hosts menu opens. Select the volumes to map from the **Unmapped Volumes** list, as shown in Figure 11-63. From the Actions tab, or by right-clicking, select **Map to Host or Cluster**.
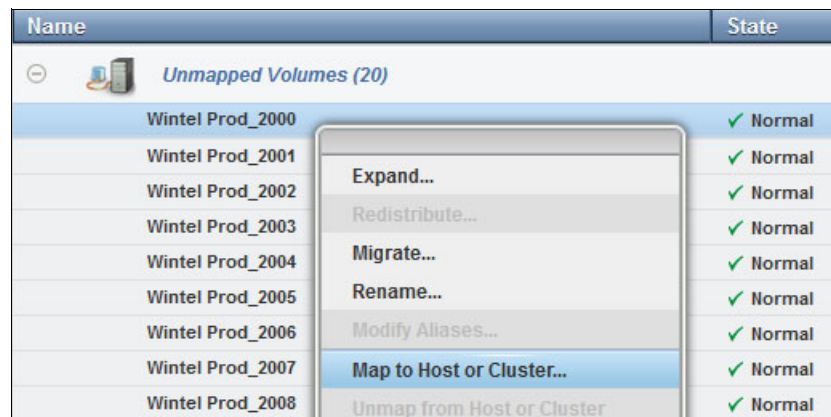
Figure 11-63   Map unmapped volumes to a host or cluster

3. The Map Volume to Host or Cluster window opens. Select the host or cluster from the list of configured stand-alone hosts and clusters, then click **Map**, as shown in Figure 11-64.
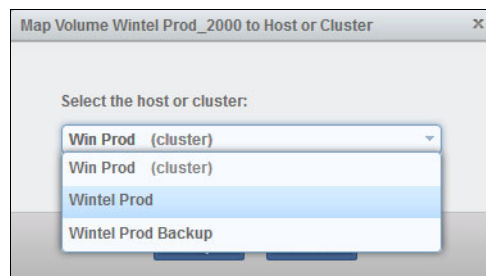
Figure 11-64   Map volumes to the selected host or cluster

Note: When mapping volumes to a cluster, volumes mapped to the cluster are public volumes that are seen by all hosts in the cluster. Volumes mapped to a single host in a cluster are private volumes.

It is the responsibility of the system administrator to ensure that the correct clustering software is implemented to ensure data integrity when a volume is mapped to more than one host.

Figure 11-65 shows a mix of public and private volumes mapped to a cluster.

Name	State	Host Type	LUN	Mapping	Capacity
⊕ Unmapped Volumes (16)					
Standalone Hosts (0)					
⊖ Win Prod	⊘ Offline				
⊖ Wintel Prod	⊘ Offline	Windows 2012			
Wintel Prod_2000	✓ Normal		00	Private	300.0 GiB
Wintel Prod_2002	✓ Normal		01	Public	300.0 GiB
Wintel Prod_2003	✓ Normal		02	Public	300.0 GiB
⊖ Wintel Prod Backup	⊘ Offline	Windows 2012			
Wintel Prod_2001	✓ Normal		00	Private	300.0 GiB
Wintel Prod_2002	✓ Normal		01	Public	300.0 GiB
Wintel Prod_2003	✓ Normal		02	Public	300.0 GiB

Figure 11-65   Mapping volumes to a cluster

11.6 Logical configuration for count key data volumes

This section describes the logical configuration for CKD volumes for z Systems.

11.6.1 Configuration flow

As with the historical Storage Management GUI, the logical configuration for CKD volumes is greatly simplified and can now be accomplished with a few steps and within a minimum amount of time.

Figure 11-66 shows the logical configuration flow for CKD.
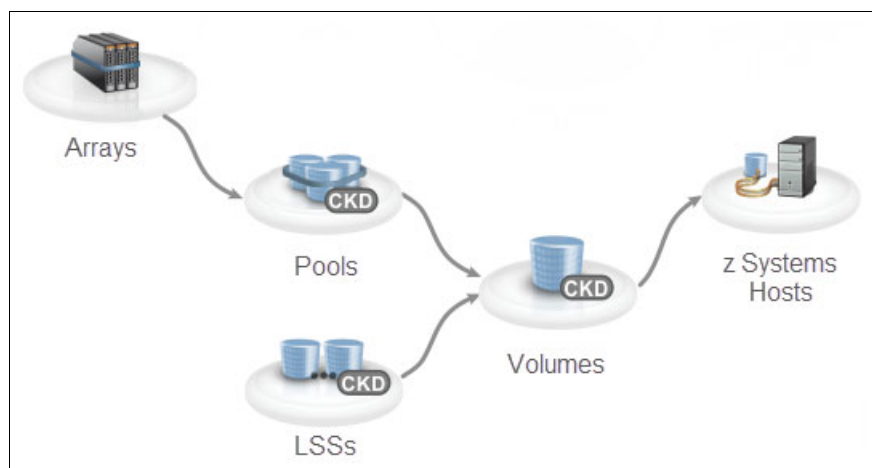
Figure 11-66   Logical configuration flow for z Systems

Use these steps to logically configure CKD:

1. Create a CKD pool pair.
2. Create the CKD LSSs.
3. Create the CKD volumes.
4. Configure the parallel access volumes (PAVs).
5. Configure the I/O ports for FICON.

11.6.2 Creating CKD storage pools

For the best performance and a balanced workload, two pools must be created. The Storage Management GUI helps the system administrator to create a balanced configuration by creating pools as a pair. The pools are configured so that one pool of the pair is managed by node 0, and the other pool of the pair is managed by node 1.

To create a CKD pool pair, complete these steps:

1. From the Storage Management GUI window, select the **Pools** icon.

2. Click **Arrays by Pool** to open the Arrays by Pool window.

3. Click the Create Pool Pair tab.

4. In the Create Pool Pair window that is shown in Figure 11-67, select the number of arrays to assign to the pool pair.
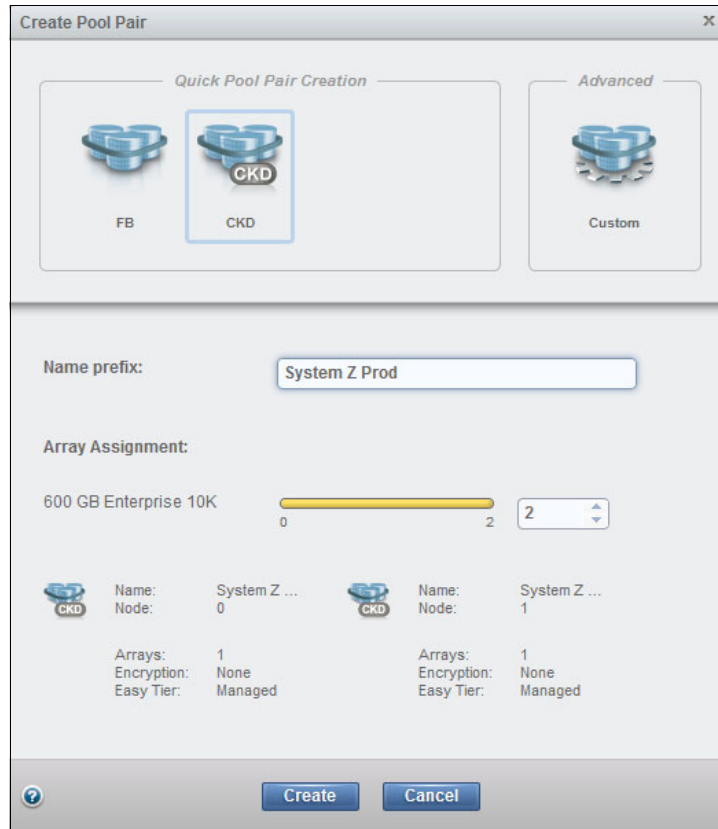
Figure 11-67   Create the CKD pool pair and assign arrays to the pool pair

5. If multiple drive classes are installed on the storage system, decide how many arrays of each drive class are required in each pool.

6. Ensure that storage type **CKD** is selected.

7. Assign a name to the pool pair. This name is used as the prefix for the pool pair ID.

8. Click **Create**.

> **Important:** The CKD LSSs cannot be created in an address group that already contains fixed-block (FB) LSSs. The address groups are identified by the first digit in the two-digit LSS ID.

9. After the pool pair creation is complete, the arrays are assigned to the pool pair, as shown in Figure 11-68. The Storage Management GUI configures the selected arrays for CKD storage and distributes them evenly between the two pools.

> **Note:** You can automatically assign arrays when creating a pool pair. The arrays are created with the default raid type, RAID 6 for the new DS8880 models (984, 985, 986, and 988) and RAID 5 for previous models (980, 981, and 982). To configure other supported RAID types, you must use the advanced configuration for pool creation, or assign the arrays manually to an existing storage pool from the unassigned arrays. For more information, see "Creating CKD Pools: Advanced Configuration" on page 305.

ID			State	Capacity	Drive Capacity	Drive Class	▼
		Unassigned Arrays (0)					
⊖		System Z Prod_6			3.54 KMod1 in Node 0		
		MA4	✓ Normal	3,540.00 Mod1	600 GB	Enterprise 10K	
⊖		System Z Prod_7			3.54 KMod1 in Node 1		
		MA3	✓ Normal	3,540.00 Mod1	600 GB	Enterprise 10K	

Figure 11-68   CKD pool pair created and arrays assigned to the pool pair

Creating CKD Pools: Advanced Configuration

When creating CKD storage pools, you can also specify the extent size and RAID level. With Release 8.1 of DS8880 microcode and later, you can specify large 1,113 cylinder, or small 21 cylinder extents for a storage pool or pool pair. The RAID level for the arrays can also be specified at creation time by using the advanced (Custom) configuration option.

To create a custom CKD pool pair, complete these steps:

1. From the System window, click the **Pools** icon.

2. Click **Arrays by Pool** to open the Array by Pool window as shown in Figure 11-23.

3. Click the **Create Pool Pair** tab. The **Create Pool Pair** window opens:

4. Specify the pool pair parameters, as shown in Figure 11-69:

   – Storage type: Ensure that **Count Key Data (CKD)** is selected.

   – Select the **Custom** option in the Advanced section.

   – Name prefix: Add the pool pair name prefix. A suffix ID sequence number is added during the creation process.

   – Extent size: Select **1,113 cylinders** for large extents, or **21 cylinders** for small extents.

5. Select from the listed drive types and select the number of arrays for each drive type that you want to assign to the pool pair.

6. Choose the RAID level for the selected arrays.

7. When pool pair parameters are correctly specified, click **Create** to proceed.

Figure 11-69 shows custom configuration options.
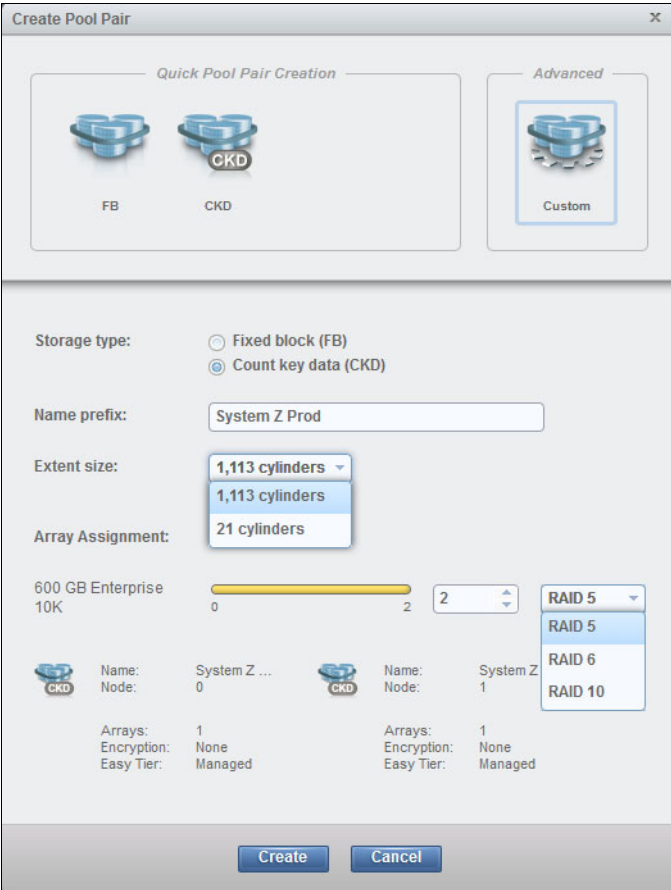
Figure 11-69   Configure CKD Storage Pools with Advanced Options

With new DS8880 models (984,985,986, and 988), RAID level 5 is not recommended for drives larger than 1 TB and the option is not available in advanced configuration as shown in Figure 11-70. If your environment requires RAID 5, contact IBM.
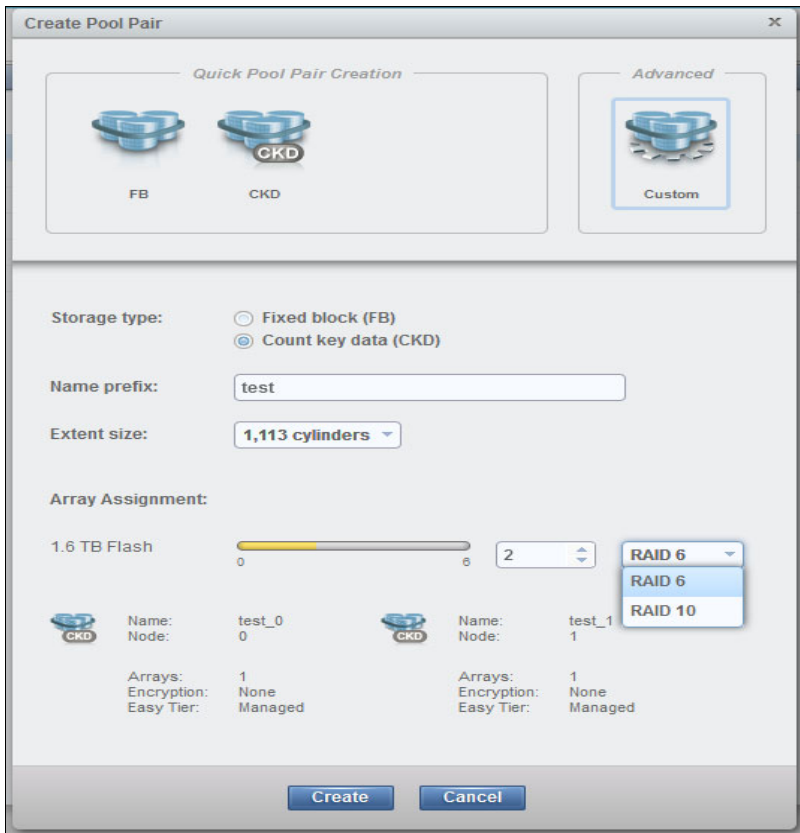
Figure 11-70   Configure CKD Storage Pools with advanced option Model 985

RAID 5 is allowed for drives less than 1 TB and when selected you need to accept the disclaimer acknowledging that you understand the risks that are associated with RAID 5. A Time stamped record with user information is created for audit purposes. See Figure 11-71.

For more information about the supported drive types and available RAID levels for DS8880 models, see Chapter 2, "IBM DS8880 hardware components and architecture" on page 23.
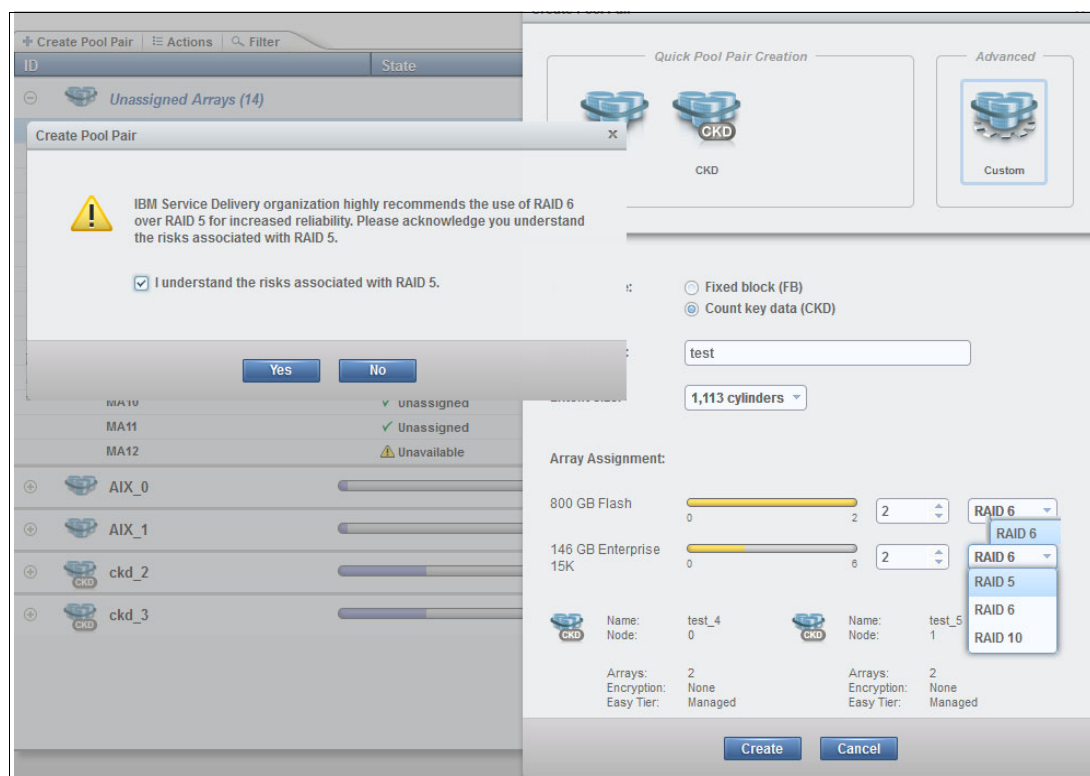
Figure 11-71   Configure CKD Storage Pools with advanced option

Manually assigning arrays to existing pools

The system administrator can manually unassign arrays or reassign assigned arrays to existing pools when the current configuration needs to be modified, such as when the system administrator adds storage capacity. To do so, complete these steps:

1. From the Arrays by Pool window, select the array that you want to assign, then either right-click the array, or select **Actions** → **Assign**.

2. When you manually assign arrays, choose from an existing storage pool and define the RAID type.

3. Click **Assign**.

Figure 11-72 shows the Arrays by Pool window showing how to assign the arrays.
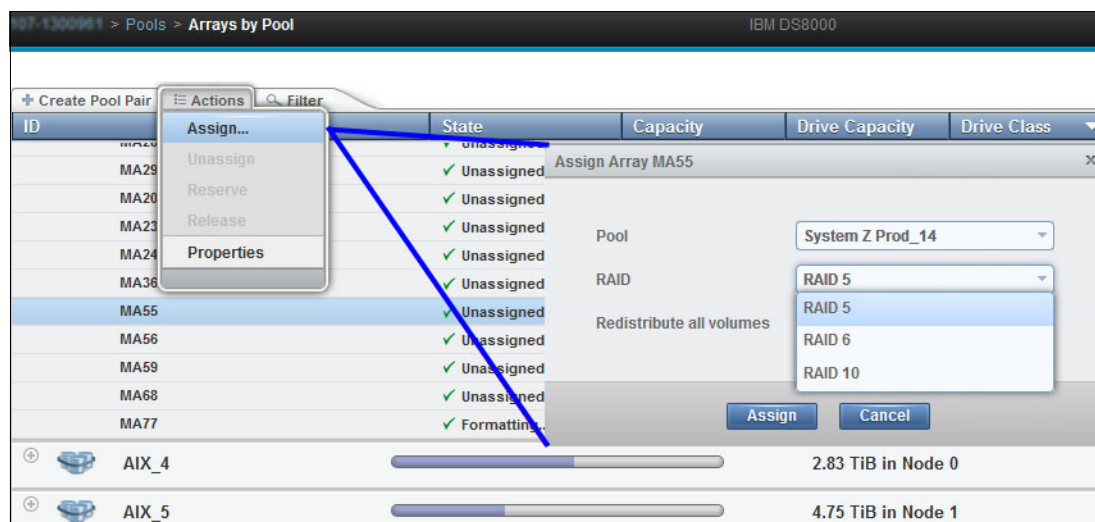
Figure 11-72   Manually assigning arrays to an existing pool

Creating a single pool

Occasionally, you are required to create only a single pool, as opposed to creating a pool pair for balancing the workload. To create a single storage pool, see "Creating a single pool" on page 281.

11.6.3 Creating CKD logical subsystems

The DS8880 LSS emulates a CKD storage control unit image. A CKD LSS must be created before CKD volumes can be associated to the LSS.

To create CKD LSSs, complete these steps:

1. Click the **Volumes** icon in the system page.

2. Select **Volumes by LSS** from the menu.

3. Click the **Create CKD LSSs** tab from the Volumes by LSS page as shown in Figure 11-73.
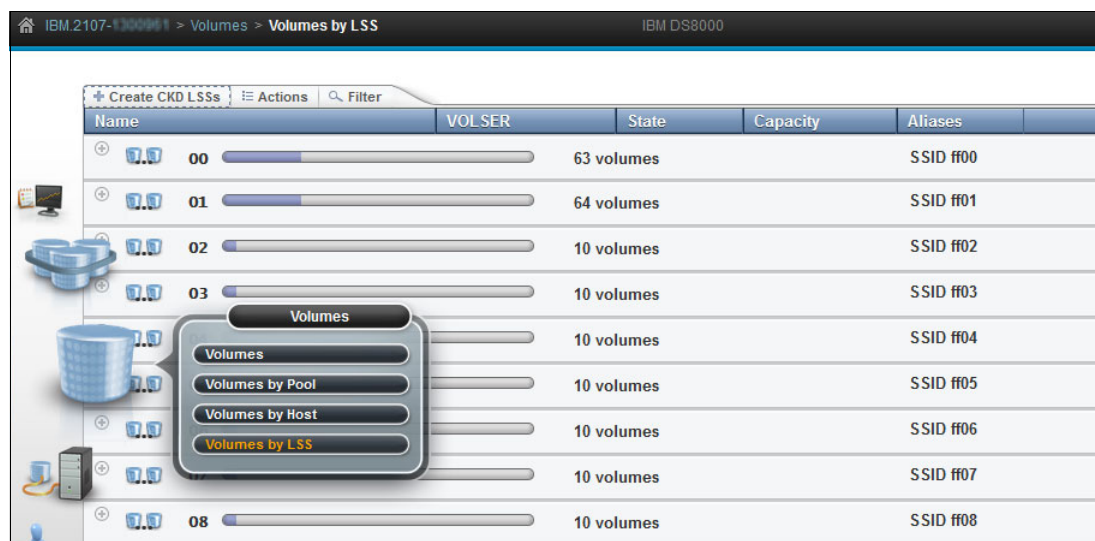
Figure 11-73   Create CKD volumes by LSSs

4. The Create CKD LSSs page opens, as shown in Figure 11-74. Complete the required information. After you enter the values for the LSS range, subsystem identifier (SSID) prefix, and LSS type, click **Create**. The Need Help icon displays information about how the unique SSID for each LSS is determined based on the SSID prefix that is provided.
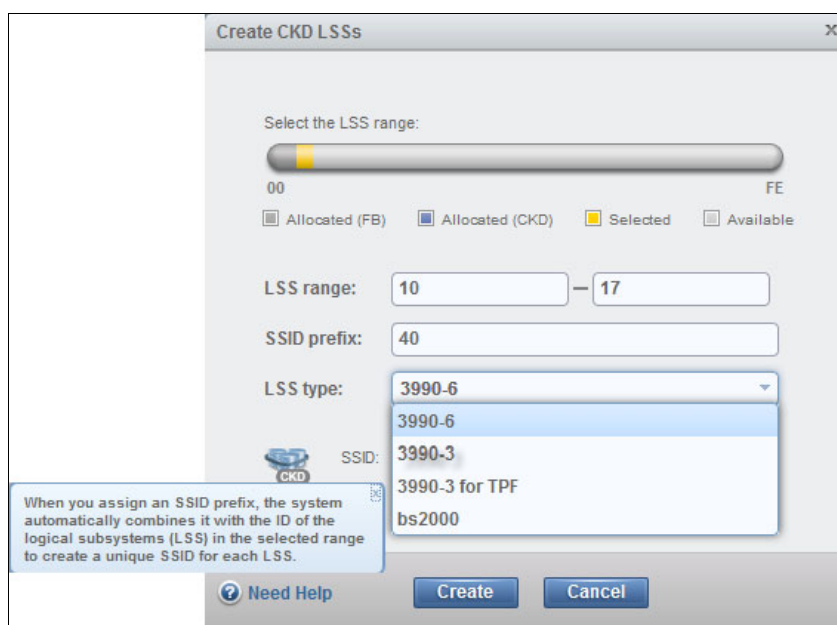
Figure 11-74   Defining CKD LSSs

> **Note:** The CKD LSSs cannot be created in an address group that already contains FB LSSs. The address groups are identified by the first digit in the two-digit LSS ID.

Figure 11-75 shows the resulting eight LSSs that are created.

Figure 11-75   CKD LSSs that are created

5. The unique SSID for each LSS is automatically determined by combining the SSID prefix with the ID of the LSS. The SSID can be modified if needed as shown in Figure 11-76.

> **Important:** This situation is important in a z Systems environment where the SSIDs were previously defined in the input/output definition files (IODFs) and might differ from the SSIDs that are automatically generated by the Storage Management GUI.
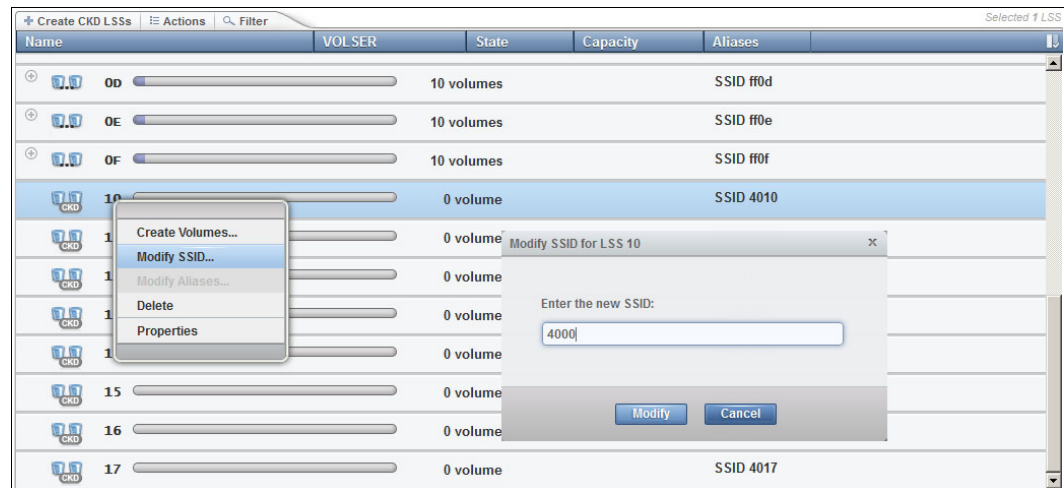
Figure 11-76   Modify the CKD LSS SSID

> **Note:** Occasionally, the DS8880 Storage Management GUI view does not immediately update after modifications are made. After you modify the SSID, if the view is not updated, refresh the GUI cache to reflect the change by clicking **Settings** → **GUI Preferences** → **General: Refresh GUI cache**.

11.6.4 Creating CKD volumes

To create CKD volumes, complete these steps:

1. Select the **Volumes** icon from the system page, then select **Volumes** or **Volumes by Pool** from the menu as shown in Figure 11-77.
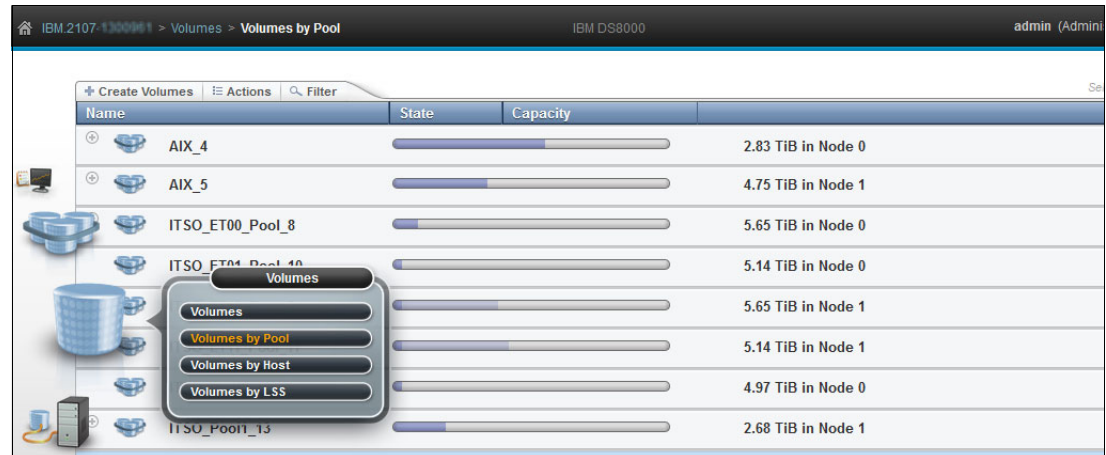
Figure 11-77   Select pools to add CKD volumes

2. Select the **Create Volumes** tab. The Create Volumes window opens.
3. Select **CKD** under Quick Volume Creation to display the configuration menu that is shown in Figure 11-78.
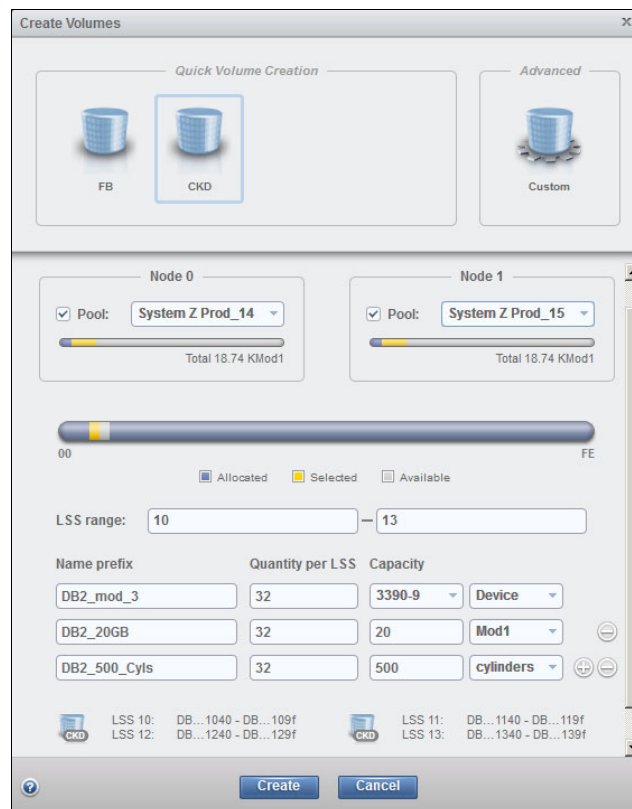
Figure 11-78   Create multiple groups of CKD volumes

4. Determine the LSS range for the volumes to create. Determine the name prefix and the quantity of volumes to create for each LSS.

5. Multiple groups of CKD volumes in an LSS range for the same pool pair can be created at the same time. Each group of created volumes uses the same name prefix and capacity. To configure multiple groups of columns, click the plus sign (**+**) icon to add and the minus sign (**-**) icon to delete rows. The Storage Management GUI automatically distributes the volumes across the two pools.

In the volume creation configuration example that is shown in Figure 11-78 on page 312, three groups of 32 volumes each are in the LSS ID range of 10 - 13. You enter a prefix name and a capacity for each group. The capacity can be specified in three ways:

► Device: Select from the list (3380-2, 3380-3, 3390-1, 3390-3, 3390-9, 3390-27, and 3390-54). These device types have a fixed capacity that is based on the number of cylinders of each model:

  – A 3390 disk volume contains 56,664 bytes for each track, 15 tracks for each cylinder, and 849,960 bytes for each cylinder.

  – The most common 3390 model capacities are shown:

    • 3390-1 = 1113 cylinders
    • 3390-3 = 3339 cylinders
    • 3390-9 = 10017 cylinders
    • 3390-27 = 32760 cylinders
    • 3390-54 = 65520 cylinders

► Mod1: Emulates a 3390 Model 1. However, the number of cylinders is variable from 1 to the maximum available capacity in the pool. The required capacity is specified in gibibytes (GiBs).

► Cylinders: Enter the number of cylinders for capacity based on a 3390 cylinder @ 849,960 bytes per cylinder.

> **Note:** The DS8880 Storage Management GUI *quick volume creation* method automatically configures volumes with thin provisioning.

Create CKD volumes: Advanced configuration

To create volumes with Advanced configuration, complete these steps:

1. From the **Create Volumes** window, select **Advanced** (Custom), as shown in Figure 11-79.
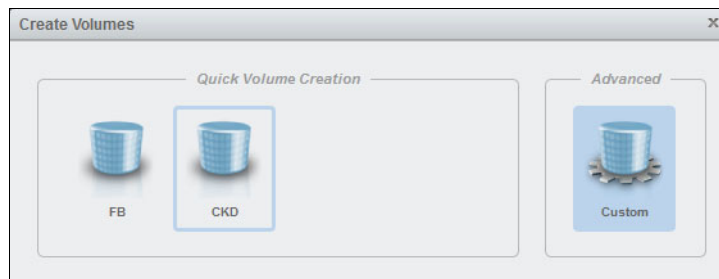
Figure 11-79   Create CKD Volumes: Advanced

2. The Type selection window appears, as shown in Figure 11-80. By default, the Storage type that is selected is **Fixed block (FB)**, so select the **Count Key Data (CKD)** radio button for Storage Type. Next, select the Data Type, **3380** or **3390** from the drop-down options.
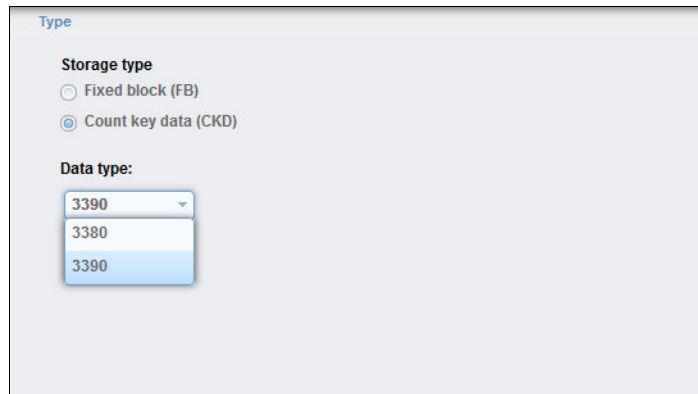
Figure 11-80   Create CKD Volumes: Advanced: Type Selection

3. Click **Allocation Settings** to open the Allocation Settings window that is shown in Figure 11-81.
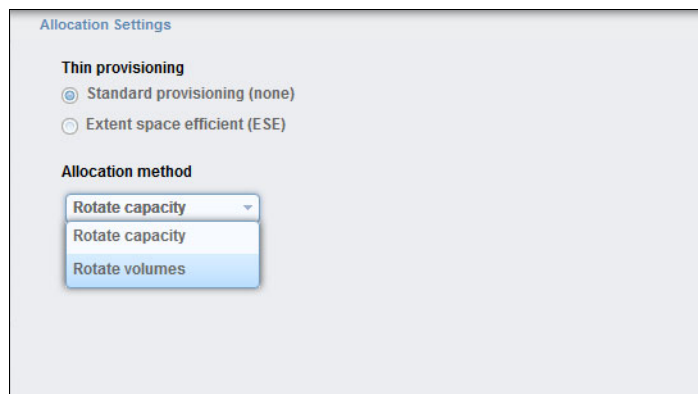
Figure 11-81   Create CKD Volumes: Advanced: Allocation Settings

4. Select the type of thin provisioning **(Standard provisioning (none)** or **Extent space efficient (ESE)**.

5. Select the extent allocation method, **Rotate capacity** or **Rotate volumes**).

6. Now click the **Volumes by LSS** tab to display the Volumes by LSS window, as shown in Figure 11-82.
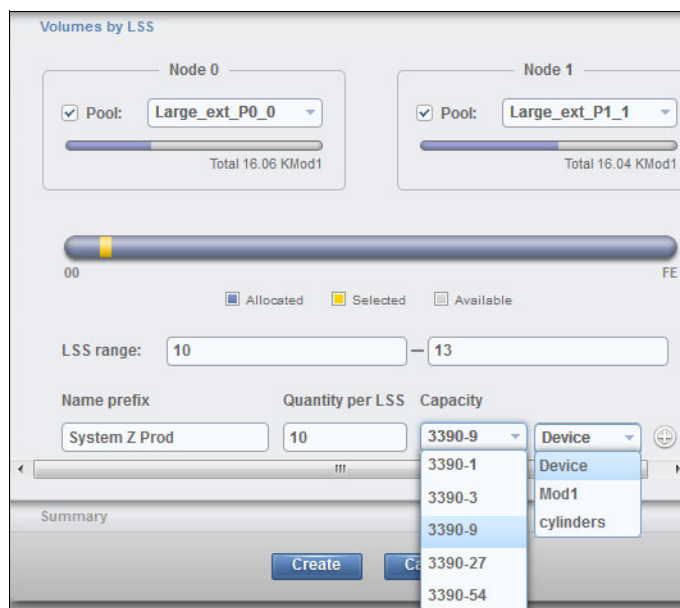
Figure 11-82   Create CKD Volumes: Advanced: Volumes by LSS

7. Specify the following information:

   – Pool selection: By default, a pool pair is selected to balance the workload across the nodes.

   – LSS range: The LSS IDs to create. If only a single LSS is required, make the range a single ID. For example, LSS range 30 - 30 creates a single LSS.

   – Volume characteristics: The storage administrator defines the following user-determined values:

     • Name prefix: User-defined name for volume in the LSSs. A suffix ID sequence number is added during the creation process.

     • Quantity: Number of volumes to create for each LSS.

     • Capacity: There are three methods of definition here.

     • Device: Options are based on device type chosen in the Type selection window, 3380 (3380-2, 3380-3) or 3390 (3390-1, 3390-3, 3390-9, 3390-27, 3390-54).

     • Mod: Specify the capacity that is required in GiB.

     • Cylinders: Specify the capacity based on a 3390 cylinder @ 849,960 bytes per cylinder.

8. Click the **Summary** tab, as shown in Figure 11-83, review details based on your selected options, and then click **Create**.

Figure 11-83   Create CKD Volumes: Advanced: Summary

11.6.5 Creating CKD parallel access volumes

The DS8880 storage system supports the configuration and usage of PAVs. PAVs allow the definition of additional unit control blocks (UCBs) to the same logical device, each using an additional alias address. For example, a DASD device at base address 1000 can have alias addresses of 1001, 1002, and 1003. Each of these alias addresses has its own UCB. Because now four UCBs exist to a single device, four concurrent I/Os are possible. Writes to the same *extent* (an area of the disk that is assigned to one contiguous area of a file) are still serialized, but other reads and writes can occur simultaneously.

In the first version of PAV, the disk controller assigns a PAV to a UCB (static PAV). The second version of PAV processing, Workload Manager (WLM), reassigns a PAV to new UCBs from time to time (dynamic PAV).

The DS8880 supports HyperPAV. For each I/O, an alias address can be selected from a pool of alias addresses within the same LSS.

Within the same LSS, with SuperPAV support enabled, aliases can even be selected from foreign LSSs.

The restriction for configuring PAVs is that the total number of base and alias addresses for each LSS cannot exceed 256 (00 - FF). These addresses need to be defined in the IODF so that they match the correct type, base, or alias.

Typically, when you configure PAVs in the IODF, the base addresses start at 00 and increment toward FF. Alias addresses typically are configured to start at FF and decrement (decrease) toward 00. A system administrator might configure only 16 or 32 aliases for each LSS. However, no restrictions exist other than the total of 256 addresses that are available to the LSS (bases and aliases).

The Storage Management GUI configures aliases in this manner, starting at FF and descending. The storage administrator can either configure many aliases against the LSS, in which case, those aliases are assigned to the lowest address in the LSS. Alternatively, the system administrator can define any number of aliases to any specific base address. For more information about PAVs, see *IBM DS8870 and IBM z Systems Synergy*, REDP-5186.

Configuring parallel access volume: Aliases

To configure aliases (PAVs), complete these steps:

1. Select the **Volumes** icon from the System window and then click **Volumes by LSS**.

2. Select the LSS for which to create aliases and click **Actions** → **Modify Aliases** or right-click the LSS, as shown in Figure 11-84.

Figure 11-84   Create aliases for the LSS

3. In the Modify Aliases for LSS *xx* (where *xx* = 00 - FE) dialog box, enter the number of aliases for the LSS. For example, enter 16 aliases for LSS 10, as shown in Figure 11-85.
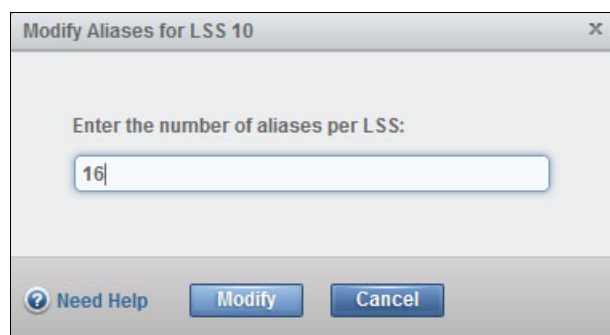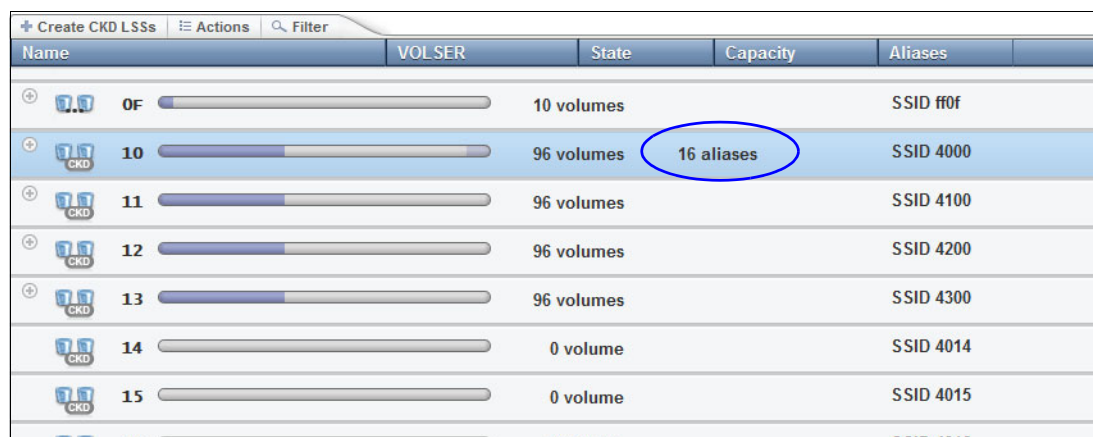
Figure 11-85   Modify 16 aliases for LSS 10

4. Click **Modify**. The aliases are created for LSS 10, as shown in Figure 11-86.
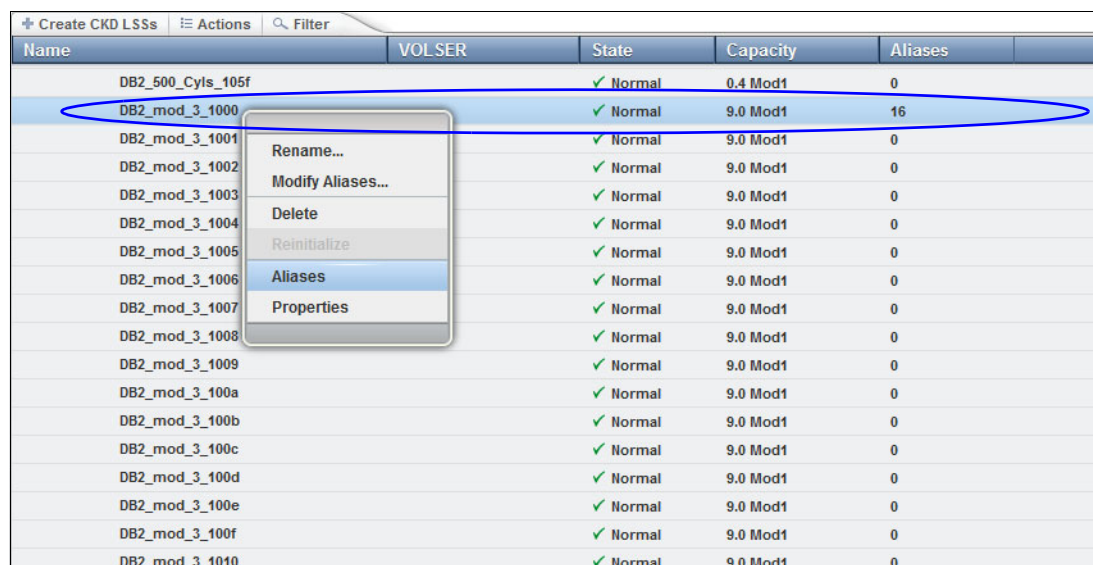
Figure 11-86   Sixteen aliases are created for LSS 10

5. Click the plus sign (**+**) icon to the left of the LSS to expand the LSS and display the base volumes that are assigned to the LSS. An example in Figure 11-87 shows the list of 96 volumes for LSS 10.

   The aliases are automatically created against the lowest base volume address in the LSS first. For example, in Figure 11-87, the 16 aliases are created against the lowest base volume address 1000 (`DB2_mod_3_1000`).

Figure 11-87   Sixteen aliases against the lowest base volume address

6. To display the aliases, select the base volume with those aliases assigned to it and then click **Action** → **Aliases**.

The list of aliases that are assigned to the base volume is displayed with the logical unit number (LUN) that is reported to the host (Alias ID) for each alias. See the sample display in Figure 11-88.
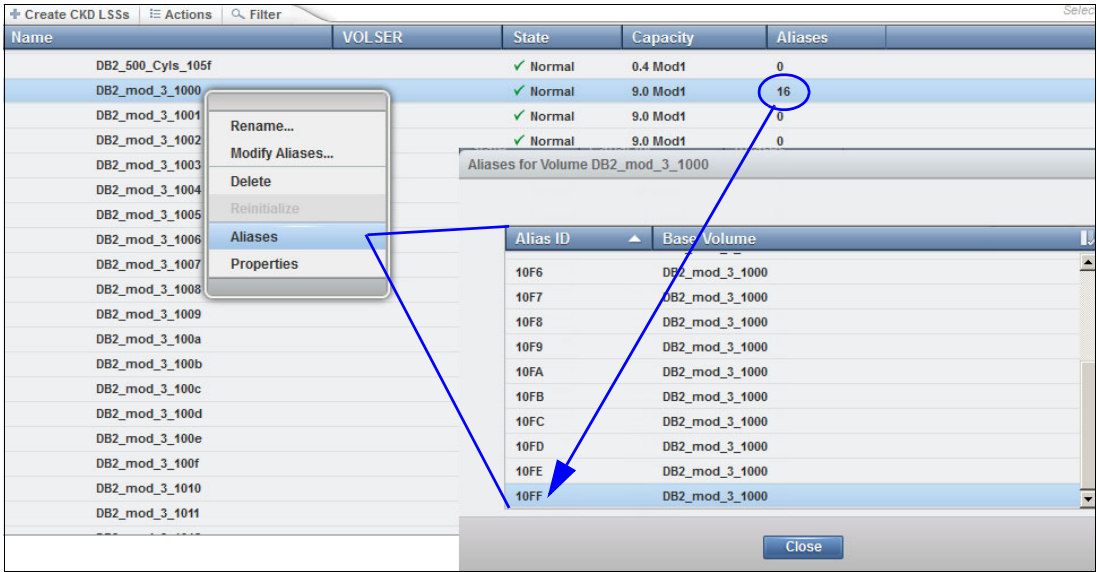
Figure 11-88   List of aliases with their alias IDs starting at FF and in descending order

Note: The alias IDs start at FF and they are in descending order, as shown in Figure 11-88.

7. Aliases can also be created for a single base volume by selecting the base volume, right-clicking, and selecting **Action** → **Modify Aliases**. Enter the number of aliases that you want for the base volume, as shown in Figure 11-89.
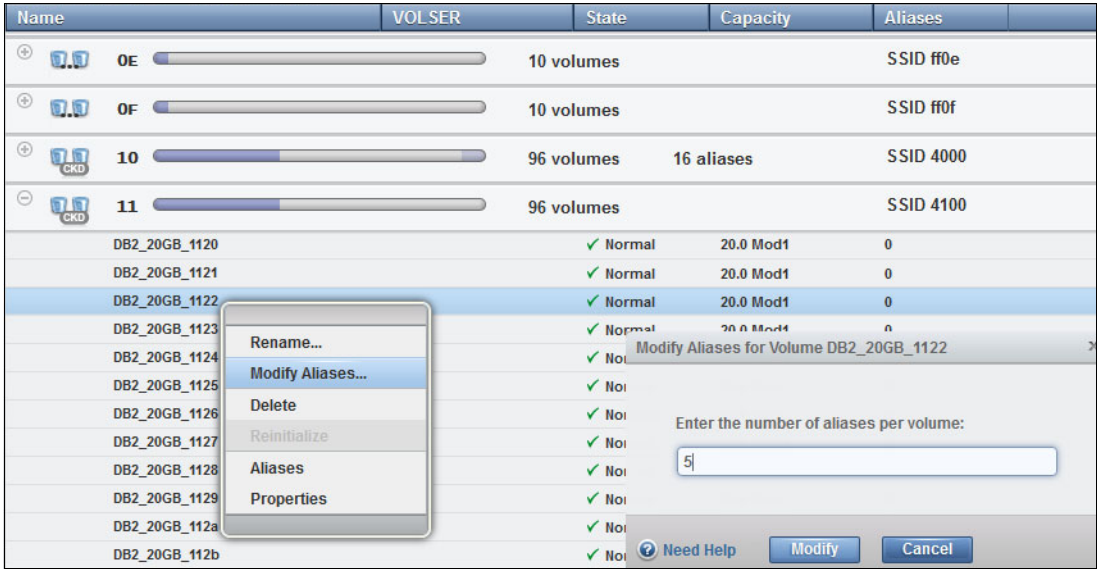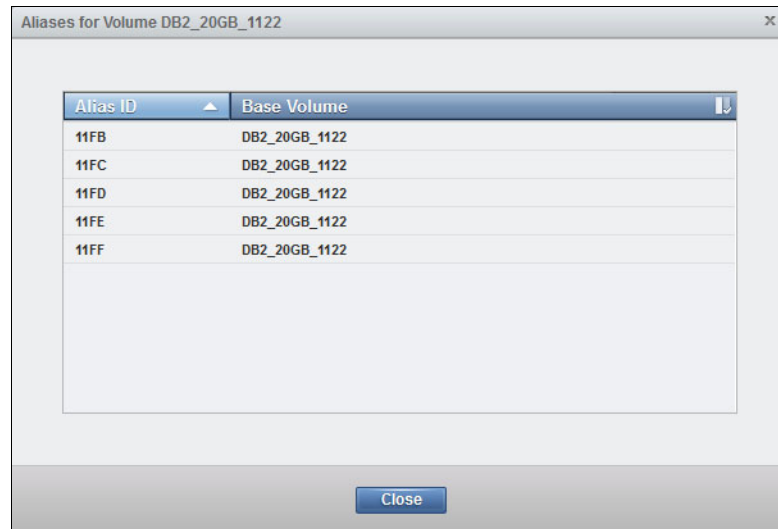
Figure 11-89   Configure aliases for a single base volume

The five aliases for a single base volume address (`DB2_20GB_1122`) are created with a starting address of FF and end with FB in descending order, as shown in Figure 11-90.

Figure 11-90   List of five aliases that are created for a single base address

Note: Starting with Release 8.2 of DS8880 microcode, deleting a pool with volumes is available in the GUI. A warning is displayed, and the user must enter a code that is presented by the DS8880 to confirm the delete. A **Force deletion** option is also available as shown in Figure 11-91.
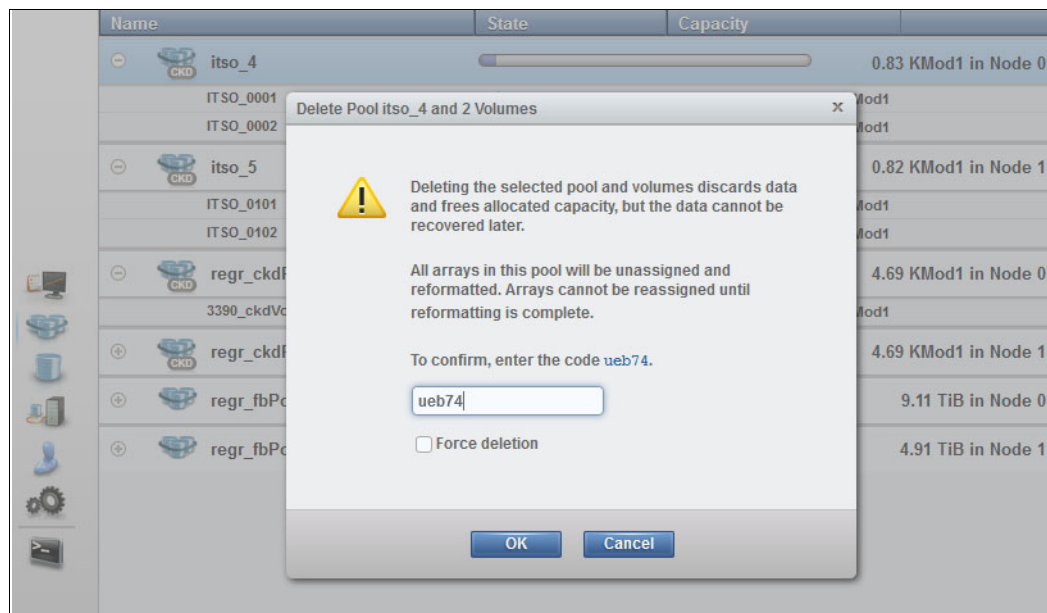
Figure 11-91   Deleting Pools with Volumes

11.6.6 Setting the I/O port protocols for z Systems attachment

For a z Systems host to access assigned CKD volumes, the host must connect to the DS8880 HA over Fibre Channel. The protocol of the host adapter port must be set to FICON.

To set the I/O port protocols of the I/O ports that the host uses to communicate with the DS8880, complete these steps:

1. In the System window, click **Actions** → **Modify I/O protocols**.

2. In the Modify I/O protocols window, select one or multiple ports to modify and click **Actions** → **Modify**.

3. The Modify Protocol for *xx* I/O Ports page opens. Select the **FICON** protocol.

4. Click **Modify** to set the topology for the selected I/O ports.

Figure 11-92 shows an example of modifying 24 I/O ports.
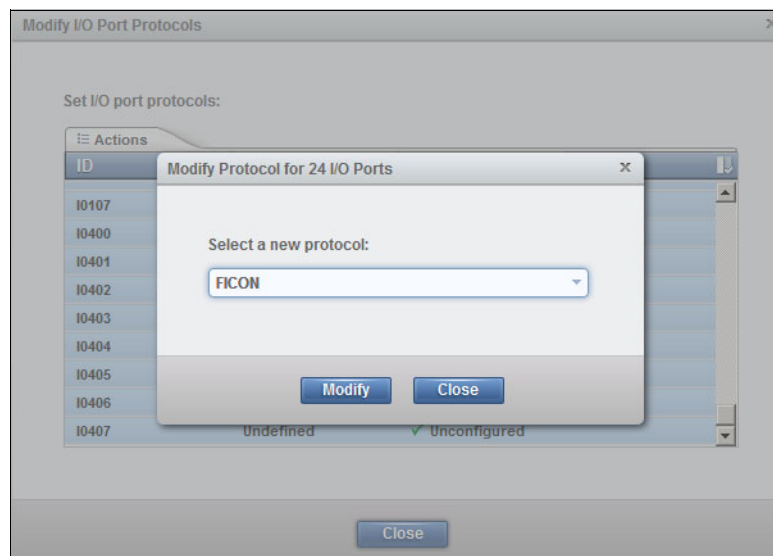
Figure 11-92   Modify 24 I/O ports to the FICON protocol

11.7 Monitoring system health

The DS8880 uses advanced predictive analysis to predict error conditions. If a system failure occurs, the system automatically provides notification. The system monitoring window provides a visual representation of the DS8880 storage system, including all system actions and system hardware states.

The DS8880 Storage Management GUI provides tools to help monitor the health of the storage system in real time. The following tools provide the status of hardware components, errors, and alerting events in the DS8880 Storage Management GUI:

► Hardware resource alerts

   The state of each hardware component of the storage system is displayed on the system page. Hardware components that need attention are highlighted.

► Capacity

   Changes that affect the capacity of the storage system are displayed in the capacity pod on the left side of the status area at the bottom of the DS8000 Storage Management GUI. Click the capacity pod to go to the Arrays by Pool page.

► System health

Changes that affect data accessibility and hardware states are displayed in the system health pod on the right side of the status area at the bottom of the DS8000 Storage Management GUI. If a hardware error is displayed, click the system health pod to go to the system page, where you can observe the hardware component that needs attention.

► Alerting events

Error and warning alerts are displayed as badges on the event status icon in the lower-right corner of the DS8000 Storage Management GUI. Click the **Event Status** icon to see the alerts. Click the specific alert to view the corresponding event in the Events page.

Figure 11-93 shows an example of the overall status of a system with one alert and three warnings.
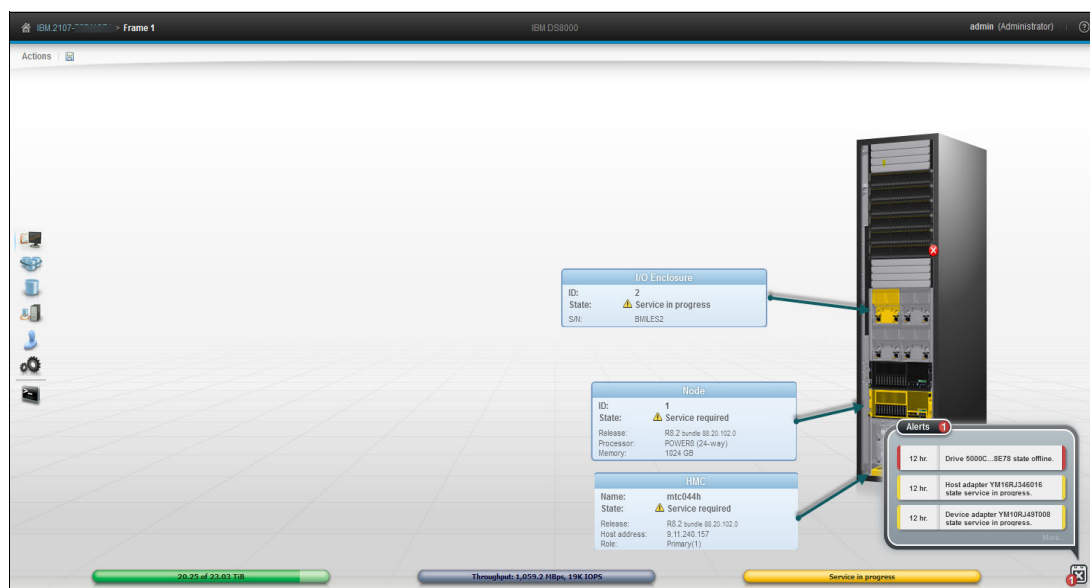
Figure 11-93   System Monitoring window

Point to different system components for detailed information. Click the component to create a zoomed view, and display details and the current state of that component.

11.7.1 Hardware components: Status and attributes

The DS8880 Storage Management GUI offers the ability to identify different types of hardware components on the storage enclosure. Five of these hardware components are listed as an example:

► Processor nodes
► HMCs for management
► Storage enclosures, including drives
► I/O enclosures, including I/O adapters
► Uninterruptible power supplies (UPSs) for power

This section provides more information about these hardware components and how details are displayed from the system page.

Processor nodes

Two processor nodes exist that are named ID 0 and ID 1. Each node consists of a CPC and the LIC that runs on it. Point to each node to display detailed information about it, as shown in Figure 11-94.
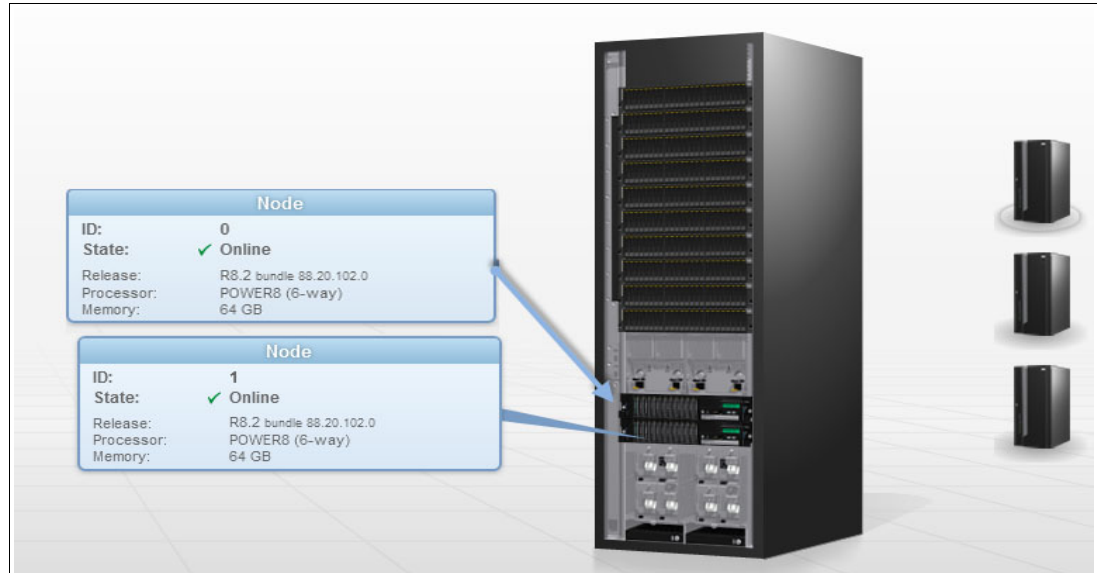
Figure 11-94   Detailed views of the processor nodes

The following are the node attributes shown in Figure 11-95:

▶ ID: The node identifier, which is node 0 or node 1.

▶ State: The current state of the node is shown:

– Online: The node is operating.

– Initializing: The node is starting or not yet operational.

– Service required: The node is online, but it requires service. A call home was initiated to IBM Hardware Support.

– Service in progress: The node is being serviced.

– Drive service required: One or more drives that are online require service. A call home was initiated.

– Offline: The node is offline and non-operational. A call home was initiated.

▶ Release: The version of the licensed machine code (LMC) or hardware bundle that is on the node.

▶ Processor: The type and configuration of the processor that is on the node.

▶ Memory: The amount of raw system memory that is installed in the node.

Hardware Management Console

The HMC provides a standard interface on a dedicated console to control managed systems. One HMC is inside the base frame. A second HMC can be ordered to provide redundancy. Point to the HMC component to display the detailed attributes for both HMCs.

The following are the attributes that are displayed for the HMC component:

► Name: The name of the HMC as defined by the user.
► State: The status of the HMC is shown:
  – Online: The HMC is operating normally.
  – Code updating: The HMC software is being updated.
  – Service required: The HMC is online, but it requires service. A call home was initiated to IBM Hardware Support.
  – Offline with redundancy: The HMC redundancy is compromised. A call home was initiated to IBM Hardware Support.
  – Offline: The HMC is offline and non-operational.
► Release: The version of the licensed machine code that is installed on the HMC.
► Host address: The IP address for the host system of the HMC.
► Role: The primary or secondary HMC.
► Location: The physical location of the HMC. If the HMC is in the base frame, the name of the storage system is displayed. If the HMC is external, the location is identified as `off-system`.

Storage enclosures

A *storage enclosure* is a specialized chassis that houses and powers the drives and flash cards in the DS8880 storage system. The storage enclosure also provides the mechanism to allow the drives to communicate with one or more host systems. The DS8880 has two types of storage enclosures:

► High-performance flash enclosures (HPFEs): Contain flash cards, which are Peripheral Component Interconnect Express (PCIe)-connected to the I/O enclosures
► Standard drive enclosures: Contain flash drives and spinning disk drives, which are fiber-connected to the device adapters in the I/O enclosures

The attributes of the storage enclosure are described:

► ID: The storage enclosure number.
► State: The current state of the storage enclosure is shown:
  – Online: The storage enclosure is operating normally.
  – Service requested: A service request to IBM was generated for one or more drives within the storage enclosure.
  – Service required: The storage enclosure is online, but it requires service.
  – Service in progress: The storage enclosure is being repaired.
  – Offline: The storage enclosure requires service.
► Drive capacity: The raw capacity of the drives or flash cards that are installed in the storage enclosure.
► Drive class: The type and speed in revolutions per minute (RPM) (if specified) of the drives of a drive or flash card in the storage enclosure. Examples include enterprise 10K, enterprise 15K, nearline 7.2K, and flash drives.
► Installed: The time and date when the enclosure was installed.
► S/N: The serial number of the storage enclosure.

A *drive* or *flash card* is a data storage device. From the GUI perspective, a drive can be either a magnetic disk drive, which is also called a hard disk drive (HDD), or a flash drive or high performance flash card. To see the data storage devices and their attributes, left-click a storage enclosure when the magnifying glass pointer appears. This action shows the storage enclosure and installed storage devices in more detail.

The attributes for the drives are described:

► S/N: The serial number of the drive.

► State: The current state of the drives is described:

– Online: Operational and normal.

– Initializing: The drive is being prepared for operation.

– Service: The drive is being serviced.

– Offline: The drive requires service.

► Array: The ID of the array that the drive belongs to.

► Capacity: The raw capacity of the drive.

► Class: The class and speed of the drive.

► Interface: The type and speed of the interface that is used by the drive.

► Firmware: The level of firmware that is installed on the drive.

► WWNN: Worldwide node name of the drive.

I/O enclosures

The *I/O enclosure* contains I/O adapters that are installed in the DS8880 storage system. The I/O adapters transfer data into and out of the system. Point to an I/O enclosure in the frame to get information about the enclosure.

The attributes for the I/O enclosure are described:

► ID: The enclosure ID.

► State: The current state of the I/O enclosure is one of the following states:

– Online: Operational and normal.
– Offline: Service is required. A service request to IBM was generated.
– Service: The enclosure is being serviced.

► S/N: The serial number of the enclosure.

To see the I/O adapters and their attributes, left-click an I/O enclosure when the magnifying glass pointer appears. This action shows the I/O enclosure adapter view (rear of the enclosure).

Two types of adapters are in the I/O enclosures:

► Device adapter (DA)
► Host adapter (HA)
► New device adapter function for connection to the new flash enclosure, DA pair 18 shown in Figure 11-95 on page 326.

I/O ports

The *I/O ports* are ports on a host adapter that connect the DS8880 to hosts, switches, or another storage system either directly or through a switch.

Point to a port, and detailed information about the port is displayed. The port topology can be modified by selecting the port, right-clicking, and selecting **Modify I/O Port Protocol**, as shown in Figure 11-95.

The following are the attributes for the I/O ports:

► ID: The identification number of the I/O port.

► State: The current state of the I/O port is one of the following states:

  – Online: The I/O port is operating normally.

  – Unconfigured: The I/O port protocol is not configured.

  – Offline: The I/O port requires service.

  – Protocol: Use one of three Fibre Channel protocols:

    • FC-AL: A Fibre Channel arbitrated loop where devices are connected in a one-way loop ring topology. The bandwidth on the loop is shared among all ports. Only two ports can communicate at a time on the loop. One port wins arbitration, after which it can open one other port.

    • SCSI FCP: A method for transferring SCSI commands with Fibre Channel Protocol.

    • FICON: Fibre Channel connection is a high-speed I/O interface for z Systems host connections.

  – WWPN: The unique 16-digit hex number that represents the worldwide port number of the I/O port.
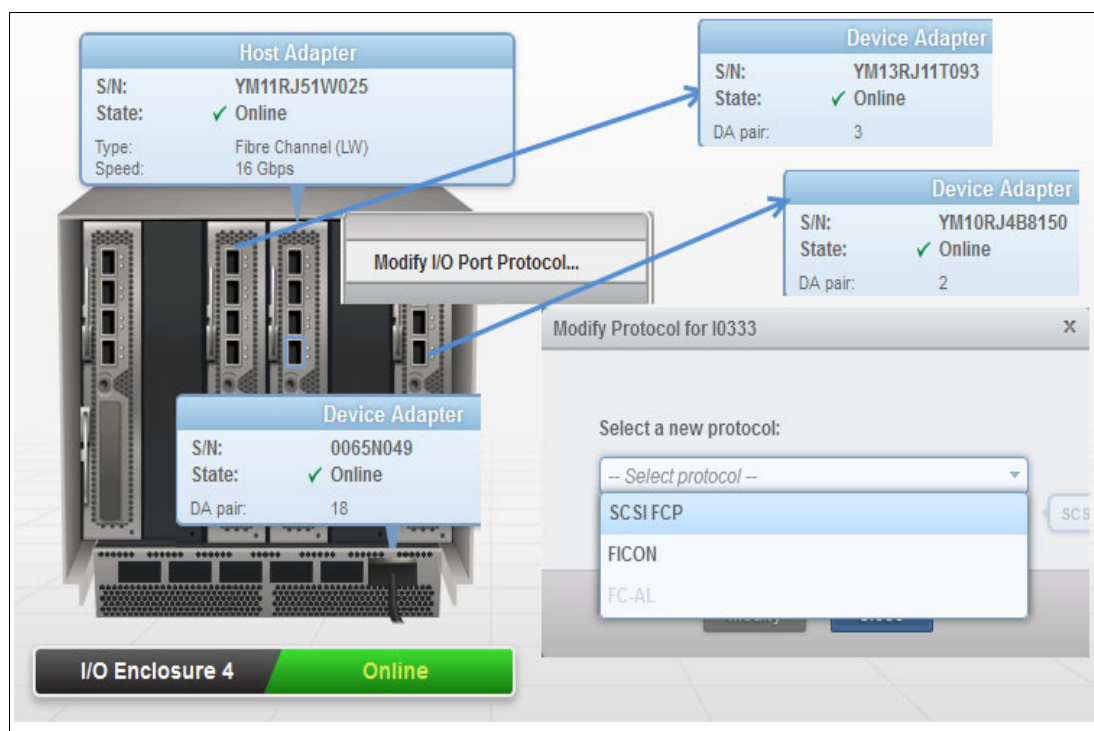
Figure 11-95   View I/O port: Modify port topology

To restore the view of a component to its position in the frame, click it in the frame.

Uninterruptible power supply

The UPS provides power to a DS8880 frame. Two UPSs exist for each frame for redundancy. The UPS maintains the storage system function if a commercial power failure occurs. If the remaining battery power in the UPS reaches a certain level, the unit initiates an orderly shutdown of system processing as required. A detailed view of each UPS can be displayed by pointing to the components.

The following are the attributes for the UPS component:

► ID: The enclosure number for the UPS.

► State: The current state of the UPS component is one of the following states:
  – Online: Operational and normal.
  – Battery service required: The UPS is online but one or more batteries require service.
  – Battery service in progress: Battery service module is being serviced.
  – Power communication: Service is required. The UPS is online but it requires service for power communication.
  – Power communication service in progress: The UPS is being serviced to restore power communication.
  – Service in progress: The UPS is being serviced.
  – Offline: The UPS requires service.

► Input power: The type of electrical power that is supplied to the DS8880.

► Extended power line disturbance (EPLD) feature: This option enables a storage facility to continue operation through a power outage that lasts either 4 seconds or 40 seconds, depending on the following settings:
  – Default: Operation can continue through a power outage that lasts up to 4 seconds.
  – EPLD: Operation can continue through a power outage that lasts up to 40 seconds.

► Local/remote switch: The setting of the physical local/remote power control switch that is on the back of the DS8880 frame. The switch controls whether the system is to turned off locally or by remote means.

► S/N: The serial number of the UPS.

11.7.2 Monitoring system events

The Events page displays all events that occurred within the storage system, whether they are initiated by a user or by the system.

The Events table updates continuously so that you can monitor events in real time and track events historically.

Events are categorized by five levels of severity:

► Error
► Warning
► Inactive error
► Inactive warning
► Information

To access the Events page, click **Monitor** → **Events**, as shown in Figure 11-96. The Events page can also be displayed by clicking the **Event Status** icon in the lower-right corner of the DS8880 Storage Management GUI system page.
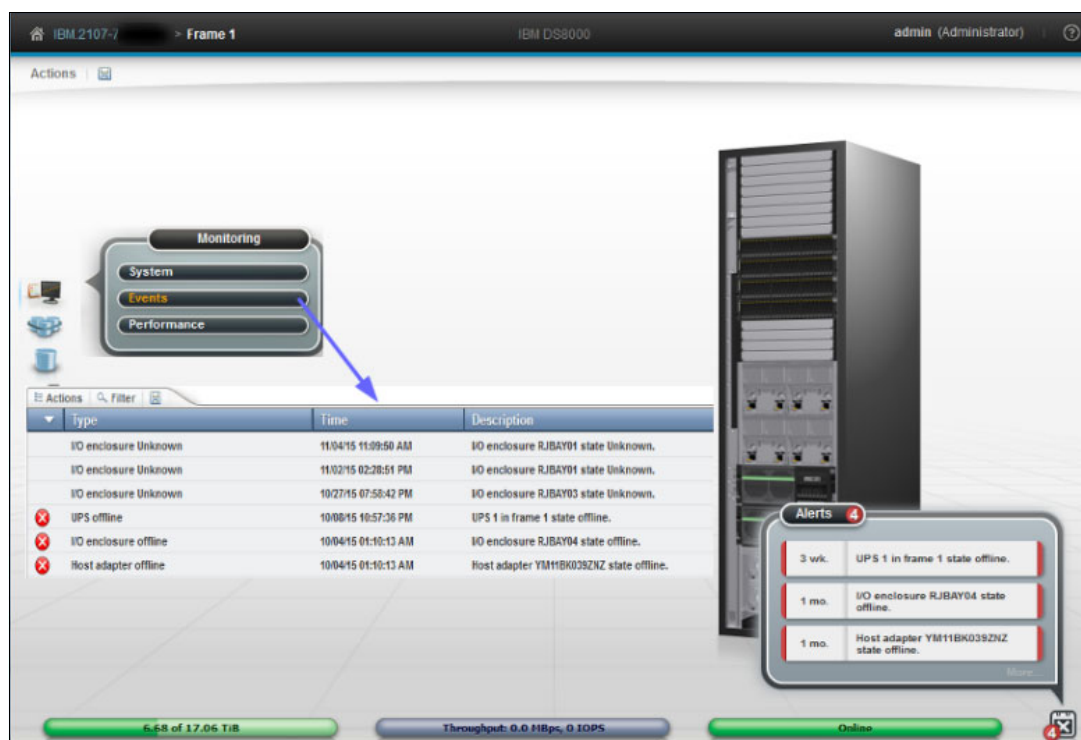
Figure 11-96   The Events page displays the following attributes for each event that occurred

The events can be exported as a comma-separated values (CSV) file by selecting **Export Table** on the Events page. The **Export Table** action creates a CSV file of the events that are displayed in the Events table with detailed descriptions.

11.7.3 Audit Logs and Syslog Notifications

Audit logs provide a record for auditing purposes to determine when changes were made to a storage system and by which user.

The audit log is an unalterable record of all actions and commands that were initiated by users on the system through the DS8880 Storage Management GUI, DS CLI, DS Network Interface (DSNI), or IBM Spectrum Control. The audit log does not include commands that were received from host systems or actions that were completed automatically by the storage system. The audit log is downloaded as a compressed text file.

An audit log for the DS8880 can also be exported by selecting **Export Audit Log** on the Events window, as shown in Figure 11-97.
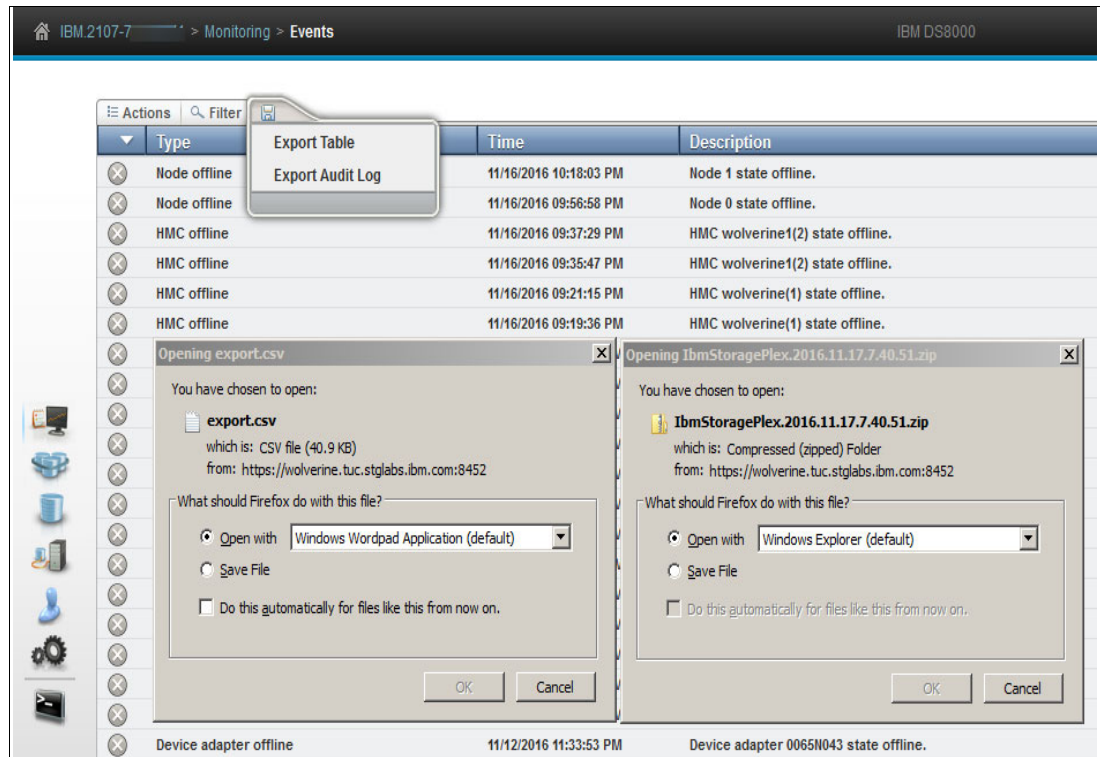
Figure 11-97   Export Table and Export Audit Log options

From DS8880 Release 8.1 microcode, up to eight external syslog servers can be configured to receive forwarded events such as user login/logout, all commands issued by using the GUI or DSCLI while the user is logged in, and remote access events.

Events are sent from Facility 19, and are logged as level 6.

To configure one or more syslog servers, complete these steps:

1. Click **Settings** → **Notifications**, as shown in Figure 11-98.

> **Note:** Syslog uses TCP protocol by default. New with Release 8.2, the Transport Layer Security (TLS) protocol can be enabled to allow secure syslog traffic.
>
> You must enable TLS before you add any syslog servers. If the storage system includes syslog servers that use TCP, you must delete these servers before you can enable TLS. Each syslog server must use the same TLS certificate.
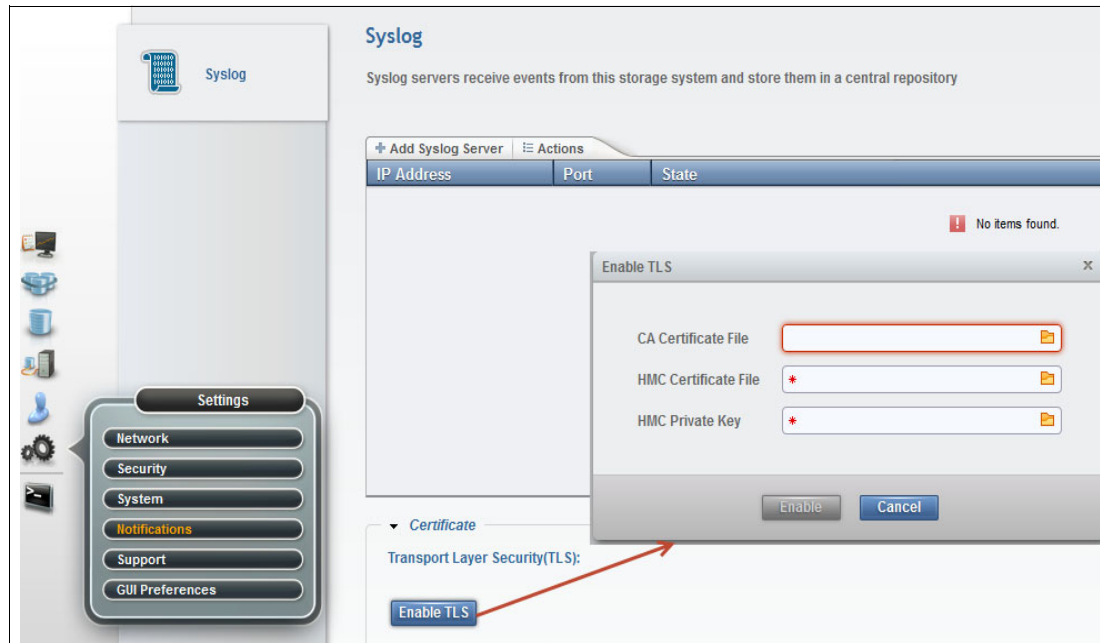
Figure 11-98   Notification Syslog Configuration

2. To enable TLS, in the Certificates area, click **Enable TLS**.

3. In the Enable TLS window, browse for the following certificate files:

   a. The certificate authority (CA) certificate file.

   b. HMC Certificate file, the storage system certificate signed by the CA.

   c. HMC Private Key, the private key for the storage system.

4. Click **Enable** to configure TLS.

5. On the Notifications window, select **Add Syslog Server**.

6. Enter the following parameters as shown in Figure 11-99:
   – **IP Address**: IP address of the external syslog server.
   – **Port**: TCP port for the external syslog server (default is 514).
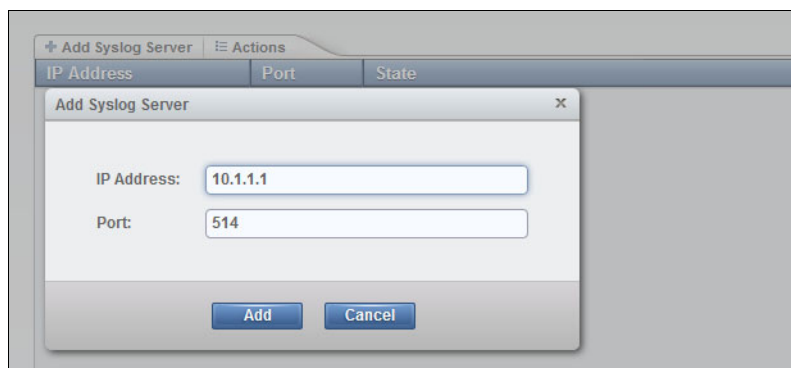
Figure 11-99   Notifications: Add Syslog Server

7. After you review the details, click **Add** to create the syslog server entry.

8. After the required syslog servers are created, right-clicking a syslog server entry allows you to **Modify**, **Test**, **Activate**, **Deactivate**, and **Remove** the selected syslog server as shown in Figure 11-100.

Note: New with Release 8.2, TLS can be enabled to allow secure syslog traffic with TLS.
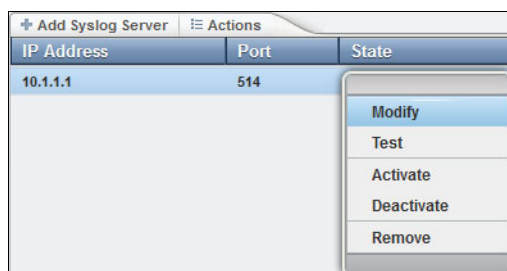
Figure 11-100   Notifications: Syslog Server Options

11.8 Performance monitoring

You can monitor storage system performance to evaluate workload processing and response times. To monitor performance, complete these steps:

1. From the System window, select the **Monitoring** icon.

2. Click **Performance** to display graphs that track performance indicators, as shown in Figure 11-101.
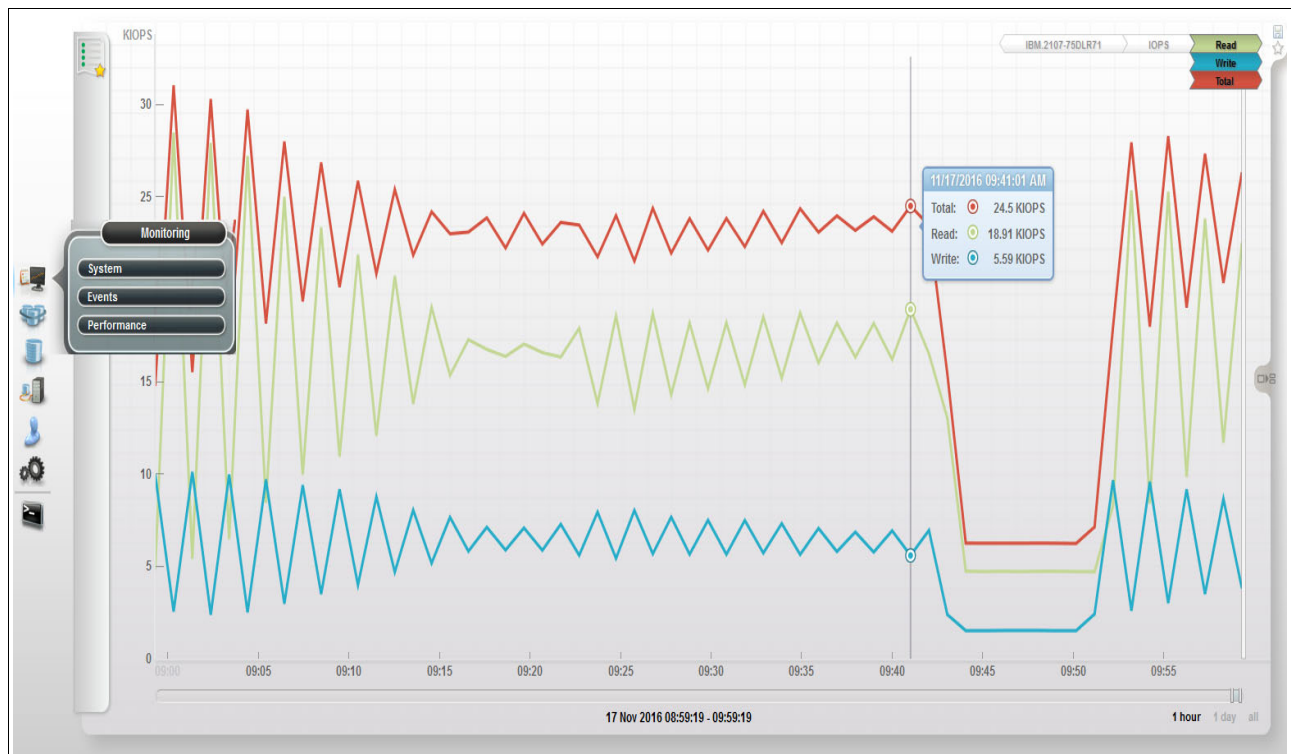
Figure 11-101   Performance monitoring window

The performance indicators include I/O operations per second (IOPS), latency, bandwidth, cache read hits, and write delays.

3. You can create your own performance graphs for the storage system, pools, and I/O ports. You can use predefined graphs and compare performance statistics for multiple pools or I/O ports.

Figure 11-102 shows a comprehensive view of the available performance functions.
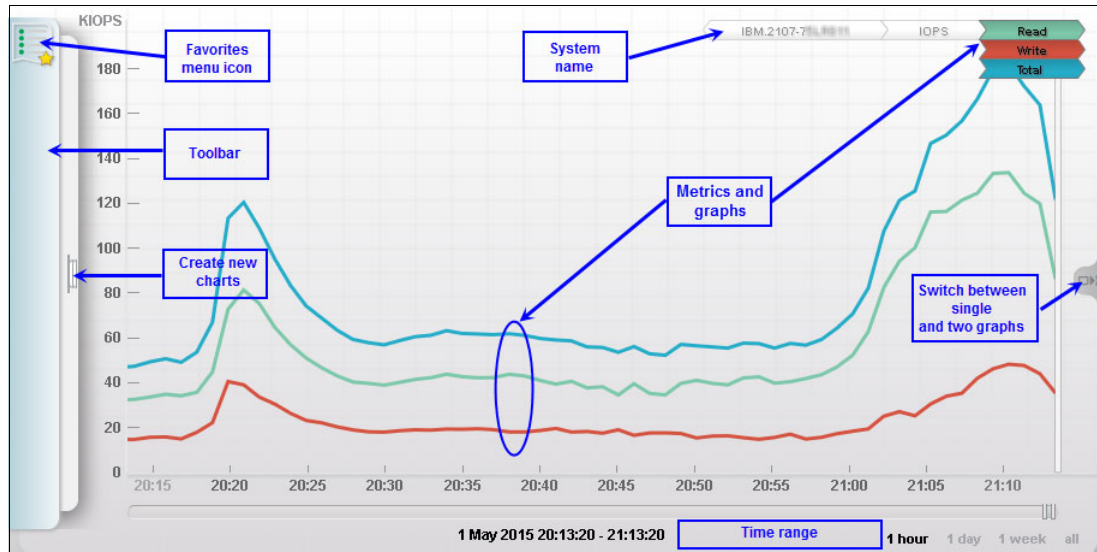
Figure 11-102   Performance monitoring graph comprehensive view

To learn how to obtain statistics from the DS CLI, see 12.5, "Metrics with DS CLI" on page 394.

11.8.1 Performance statistics

Important: All of the listed performance statistics are averaged over one minute. The performance graphs cover data that is collected for the last seven days. For long-term performance statistics, use IBM Spectrum Control.

Users can monitor the following resources and performance metrics:

► System performance metrics:

  – IOPS

    Total number of processed requests in thousand I/O operations per second (KIOPS) for the selected I/O operations.

  – Latency

    Response time in milliseconds (ms) for the selected I/O operations.

  – Transfer Size

    Number of kilobytes per I/O operation (KB/operation) for the selected I/O operations.

  – Bandwidth

    Number of megabytes per second (MBps) for the selected bandwidth type.

  – Cache

    Percentage of read hits (read I/O operations that were fulfilled from the system cache or memory) and write delays (I/O operations that were delayed because of write cache space constraints or other conditions) during one minute.

  – Capacity

    Total capacity and allocated capacity of the storage system in gibibytes (GiB).

- – Power

  Power usage of the storage system in watts.

- – Temperature

  Average temperature of the storage system in degrees Celsius.

► Array performance metrics (introduced with Release 8.1):

- – Back-end IOPS

  Number of requests that were processed on the selected arrays, in KIOPS for the selected I/O operation type.

- – Back-end Latency

  Response time in ms for the selected I/O operations that were processed on the chosen arrays.

- – Transfer Size

  Number of KB/operation for the selected I/O operations (read, write, or average of both) that were processed on the selected arrays.

- – Back-end Bandwidth

- – Number of megabytes per second (MBps) for the selected bandwidth type (read, write, or combined total) that was processed on the selected arrays.

► Pool performance metrics:

- – Back-end IOPS

  Number of requests that were processed on the arrays in the pool, in KIOPS for the selected I/O operation type.

- – Back-end Latency

  Response time in ms for the selected I/O operations that were processed on the arrays in the pool.

- – Transfer Size

  Number of KB/operation for the selected I/O operations (read, write, or average of both) that were processed on the arrays in the pool.

- – Back-end Bandwidth

  Number of MBps for the selected bandwidth type that was processed on the arrays in the selected pools.

- – Capacity

  Allocated capacity, and defined volume capacity, in GiB, in the pool.

- – Over-provisioning

  Over-provisioned capacity of the pool, as a ratio of defined capacity over total capacity.

► I/O port performance metrics:

- – IOPS

  Number of processed requests in KIOPS for the selected I/O operations (read, write, and total) on the I/O port.

- – Latency

  Response time in ms for the selected I/O operations on the I/O port.

– Transfer Size

Number of KB/operation for the selected I/O operations on the I/O port.

– Bandwidth

Number of MBps for the selected bandwidth type on the I/O port.

You can use these performance metrics to define and save your own graphs. Additionally, the Performance window includes four preset graphs, which are listed in the **Favorites** menu.

As part of the performance functionality of the DS8880, the storage sample performance logs can be exported as a CSV file for analysis by IBM, as shown in Figure 11-103.
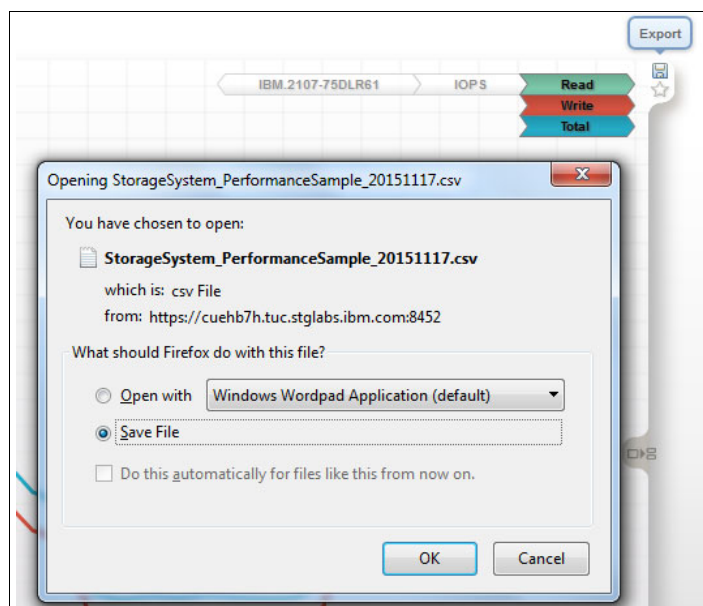
Figure 11-103   StorageSystem_PerformanceSample export CSV file

With Release 8.2 of DS8880 microcode, the performance data gathered at one minute intervals for up to one week can now be exported for IBM support analysis as shown in Figure 11-104. These include CSV offload of the system summary (full hardware and logical configuration) and exporting individual tables (Volumes, IO Ports, Users, and Volumes by LSS), and exporting individual performance charts.
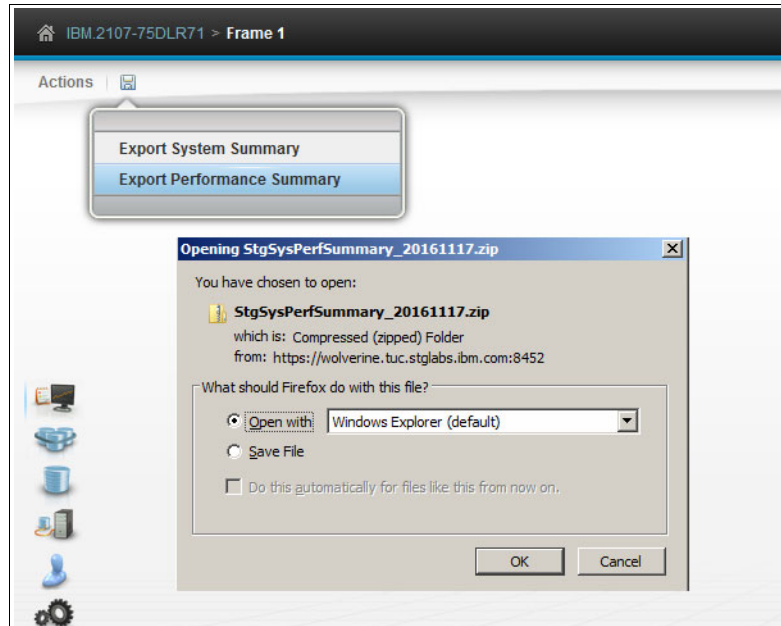
Figure 11-104   Export Performance Summary

Note: Individual table offload of volumes by LSS is new with Release 8.2 of DS8880 microcode as shown in Figure 11-105.
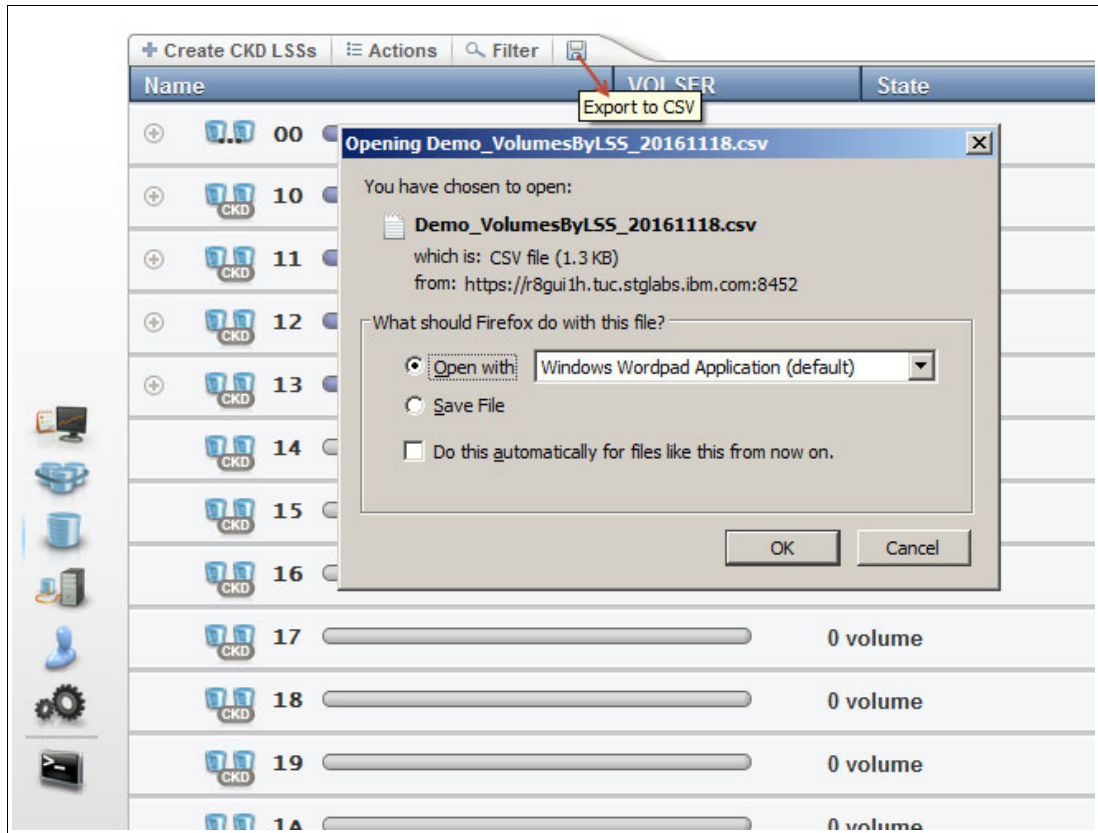
Figure 11-105   Table Export Volumes by LSS

11.8.2 Working with customized performance graphs

DS GUI offers four predefined performance graphs, which are available in the Favorites menu as shown in Figure 11-106.
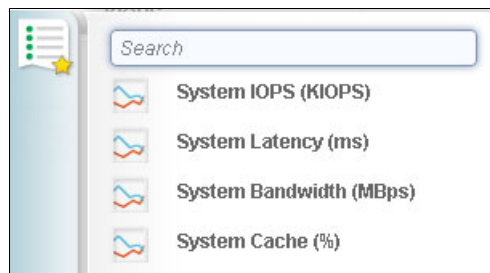
Figure 11-106   Predefined graphs

For detailed performance analysis, you can define more detailed statistics and graphs, which can help identify and isolate problems.

You can perform the following actions:

► Define your own performance graphs on demand.
► Add defined graphs to the Favorites menu.
► Pin defined graphs to the toolbar.
► Set defined graphs as a default in the Performance window.
► Rename or delete your graphs. You cannot delete predefined graphs.
► Change the time range of displayed graphs.

From the System window, click the **Monitoring** icon and click **Performance** to open the Performance Graphs window, where you can define your own graphs that are based on performance metrics that are available for each resource.

Creating array performance graphs

To create a graph from the System window, click the **Monitoring** icon. Click **Performance** to open the Performance Graphs window. Figure 11-107 demonstrates how to complete the following steps:

1. Select **Array** from the resources to monitor.
2. Select the system to monitor.
3. Select the metrics that you want.
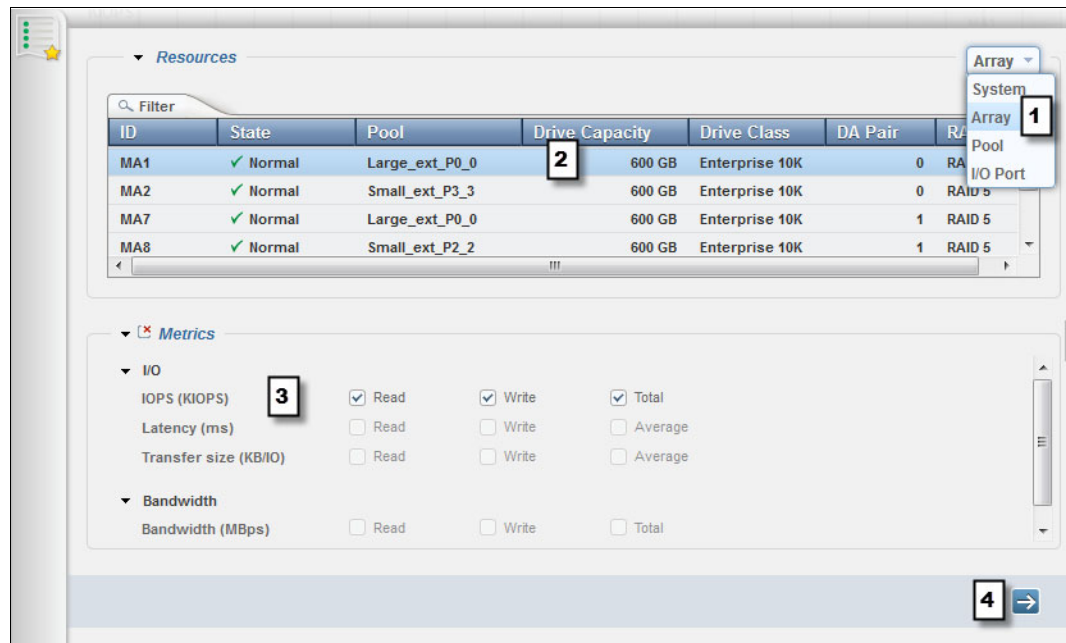
Figure 11-107   Array graph definition window

4. Draw the graph, which is shown in Figure 11-108.
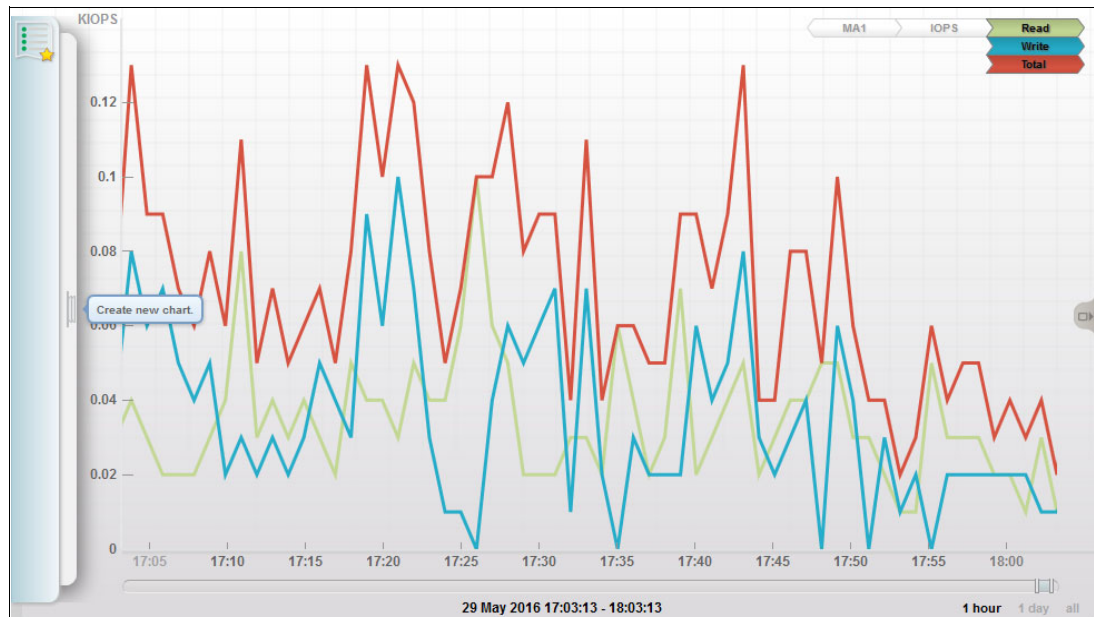
Figure 11-108   Array performance graph

Creating pool performance graphs

To create a graph from the System window, select the **Monitoring** icon. Click **Performance** to open the Performance Graphs window. Figure 11-109 shows how to complete these steps:

1. Select **Pool** from the resources to monitor.
2. Select the pool name to monitor.
3. Select the metrics that you want.
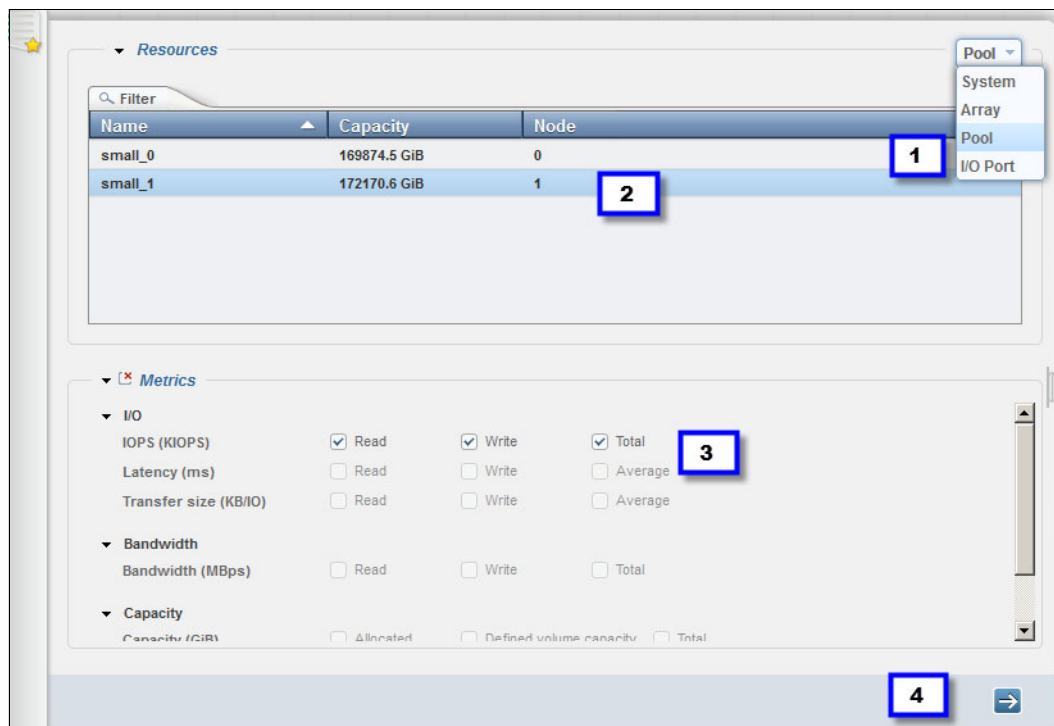
Figure 11-109   Pool graph definition window

4. Draw the graph, which is shown in Figure 11-110.
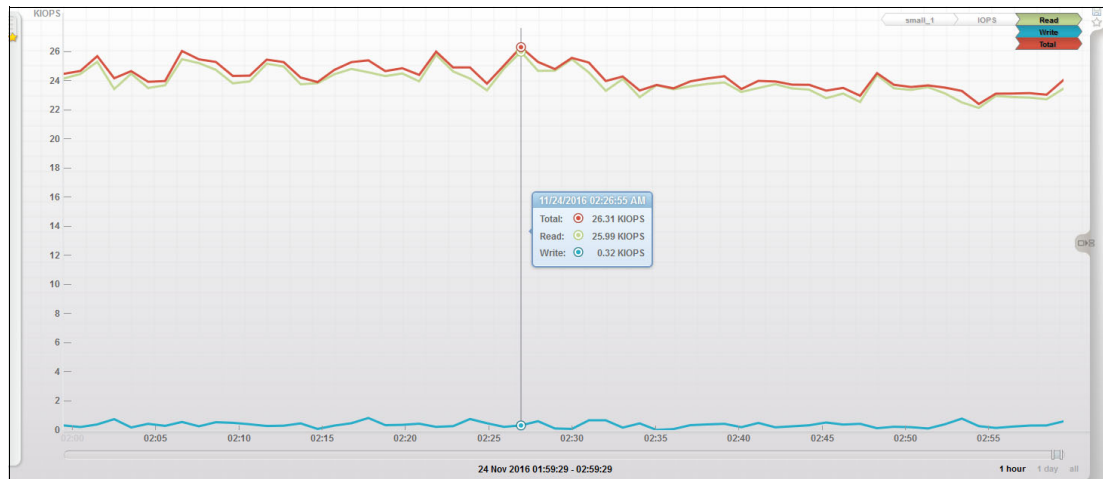
Figure 11-110   Pool performance graph

Creating system performance graphs

To create a graph from the System window, select the **Monitoring** icon. Click **Performance** to open the Performance Graphs window. Figure 11-111 shows how to complete these steps:

1. Select **Storage System** from the resources to monitor.
2. Select the system to monitor.
3. Select the metrics that you want.
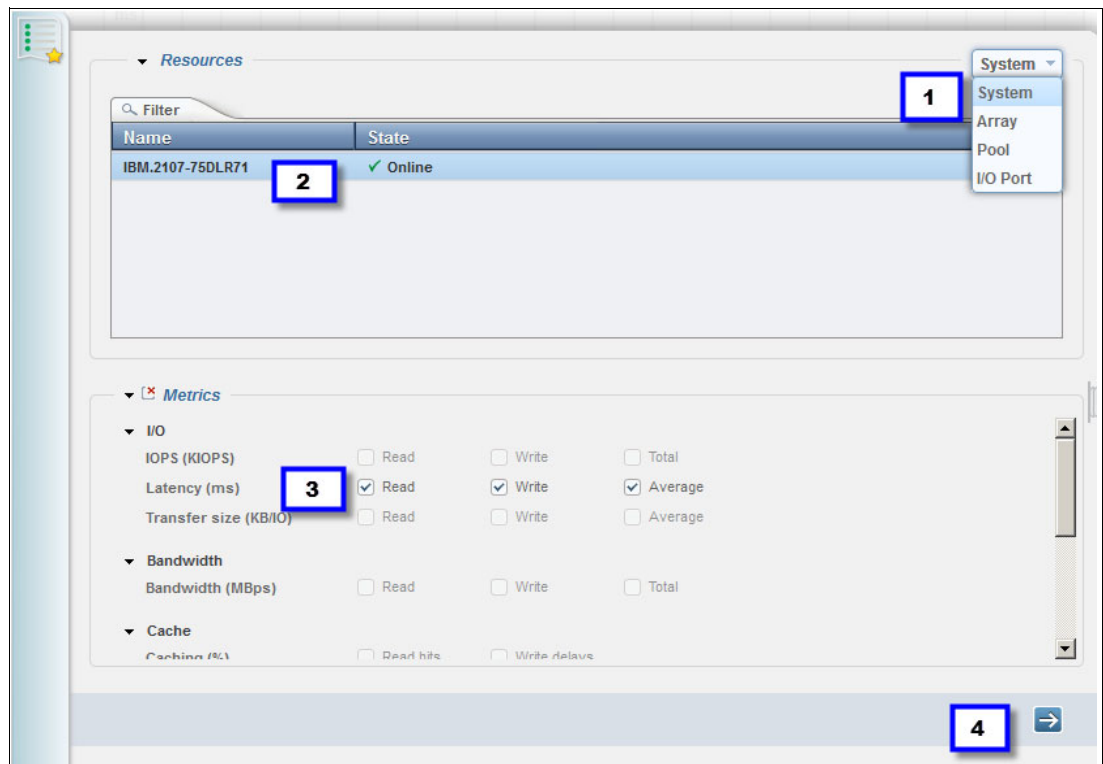
Figure 11-111   System graph definition window

4. Draw the graph, which is shown in Figure 11-112.

Figure 11-112   System performance graph

Creating I/O port performance graphs

To create a graph from the System window, click the **Monitoring** icon. Click **Performance** to open the Performance Graphs window. Complete the sequence that is indicated by the numbers that are shown in Figure 11-113.
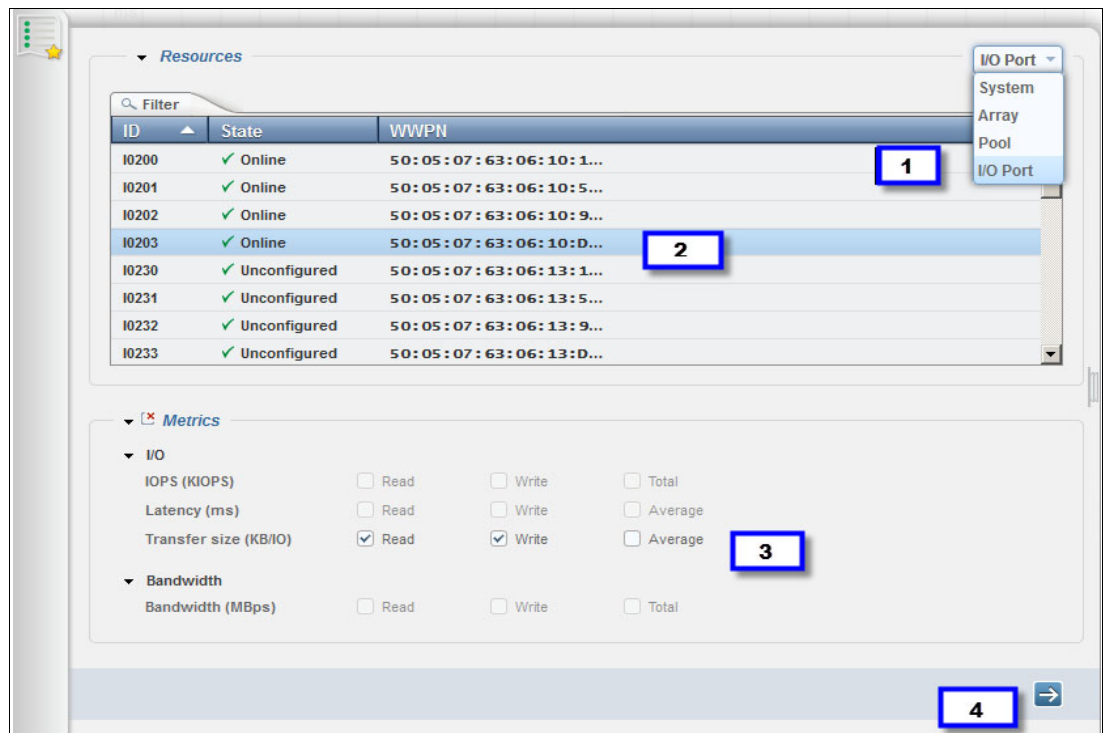
Figure 11-113   I/O port graph definition window

Figure 11-114 shows the completed graph.

Figure 11-114   I/O port performance graph

Adding a graph to the Favorites menu

You can add a graph to the Favorites menu by completing the sequence that is indicated by the numbers that are shown in Figure 11-115.
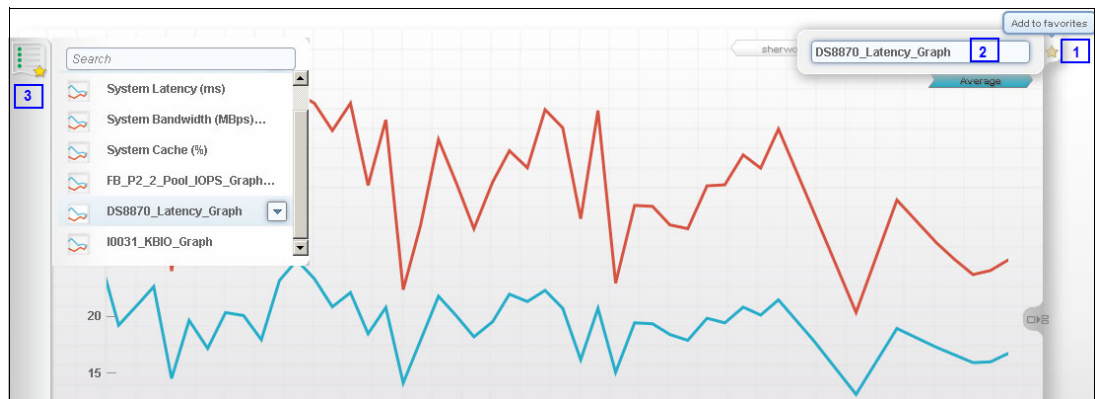
Figure 11-115   Add a graph to the Favorites menu

Pinning a graph to the toolbar

Graphs that are added to the Favorites menu can be pinned to the toolbar to enable faster access by completing the sequence that is indicated by the numbers that are shown in Figure 11-116.
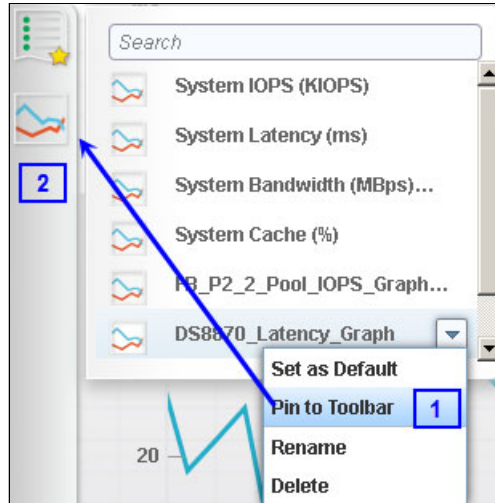
Figure 11-116   Pin a graph to the toolbar

11.9 Fibre Channel error rate statistics

Errors that occurred during data transmission through Fibre Channel ports over the past 24 hours can now be displayed in the Error Rates for I/O Port window in the Network window, as shown in Figure 11-117. This feature, which was implemented in the DS8870 GUI, is useful for a proactive check of SAN-related issues from a DS8880 perspective.
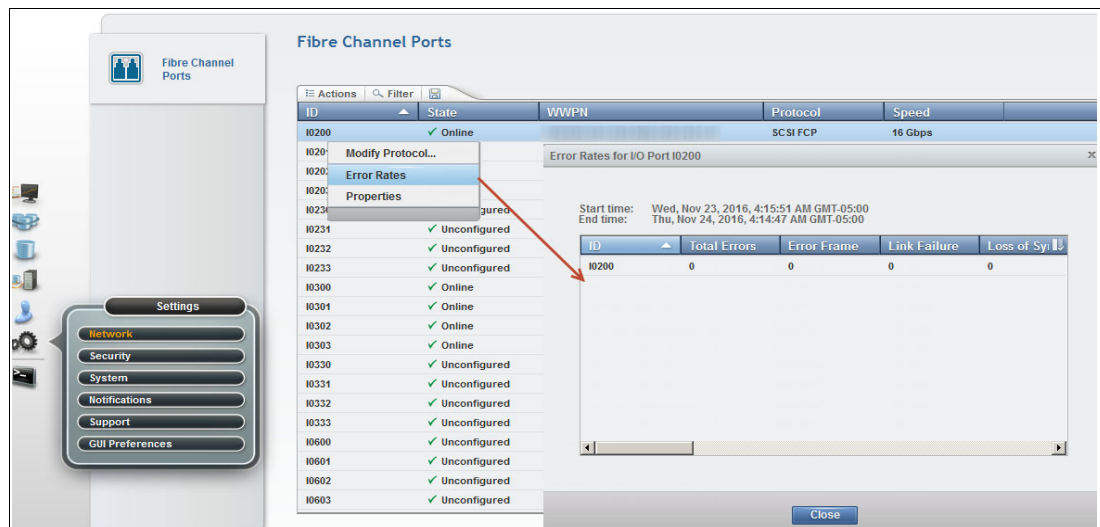
Figure 11-117   Fibre Channel Ports: Error Rates for I/O Port

The following list shows all of the statistics that are available for port error checking:

► Total errors. The total number of errors that were detected on the Fibre Channel port.

► Error frame: A Fibre Channel frame was received that was not consistent with the Fibre Channel protocol.

- ► Link failure: Fibre Channel connectivity with the port was broken. This type of error can occur when the system that is connected to the port is restarted, replaced, or serviced, and the Fibre Channel cable that is connected to the port is temporarily disconnected. It can also indicate a faulty connector or cable. Link failures result in degraded performance of the Fibre Channel port until the failure is fixed.

- ► Loss of sync: A synchronization loss error was detected on the Fibre Channel link. This type of error can occur when the system that is connected to the port is restarted, replaced, or serviced, and the Fibre Channel cable that is connected to the port is temporarily disconnected. It also can indicate a faulty connector or cable. If a synchronization loss error persists, it can result in a link failure error.

- ► Loss of signal: A loss of signal was detected on the Fibre Channel link. This type of error can occur when the system that is connected to the port is replaced or serviced, and the Fibre Channel cable that is connected to the port is temporarily disconnected. It also can indicate a faulty connector or cable. If a loss of signal error persists, it can result in a link failure error.

- ► CRC error: A Fibre Channel frame was received with cyclic redundancy check (CRC) errors. This type of error is often fixed when the frame is retransmitted. This type of error is often recoverable and it does not degrade system performance, unless the error persists and the data cannot be relayed after retransmission.

- ► Primitive sequence protocol error: A primitive sequence protocol error was detected. A primitive sequence is an ordered set that is transmitted and repeated continuously to indicate specific conditions within the port. The set also might indicate conditions that are encountered by the receiver logic of the port. This type of error occurs when an unexpected primitive sequence is received.

- ► Transmission word count: A bit error was detected. A transmission word is the smallest transmission unit that is defined in Fibre Channel. This unit consists of four transmission characters, 4 x 10, or 40 bits. This type of error can include code violations, invalid special code alignment, and disparity errors.

- ► Link reset transmitted: The state of the Fibre Channel port changed from active (AC) to link recovery (LR1).

- ► Link reset received: The state of the Fibre Channel port changed from active (AC) to link recovery (LR2) state.

- ► Out of order data: A missing frame was detected. The frame was either missing from a data sequence or it was received beyond the Fibre Channel port's sequence reassembly threshold.

- ► Out of order ACK: An out of order acknowledgment (ACK) frame was detected. ACK frames signify that the transmission was received. The frame was either missing from a data sequence or it was received beyond the Fibre Channel port's sequence reassembly threshold.

- ► Duplicate frame: A frame that was detected as previously processed was received.

- ► Invalid relative offset: A frame with an invalid relative offset parameter in the frame header was received.

- ► Sequence timeout: The Fibre Channel port detected a timeout condition when a sequence initiator was received.

- ► Uncorrectable bad blocks: A data block with errors was unable to be fixed by Forward Error Correction (FEC).

- ► Correctable bad blocks: A data block with errors was fixed by FEC.

- ► Transport mode write retries: A transport mode write operation retry was requested. The buffer was not large enough to receive unsolicited data.

11.10 Accessing the previous DS GUI

For certain special functions, access to the previous version of the DS GUI might be needed. This possibility still exists with Release 8.2 of DS8880 microcode. Directions are provided to the previous DS GUI, which can provide the specific functions that might be required.

For details, see the online help by selecting **Help Contents** from the menu that is shown on the left side of Figure 11-118.

To access the previous DS GUI, from the DS8880 Storage Management system page, click the **Help** icon (question mark in the upper-right corner). From the drop-down menu, select **Previous GUI**. This option opens a separate window to the previous GUI software as shown in Figure 11-118. Log on with the administrator role user name and password.

Figure 11-118   Access the previous GUI

12

Configuration with the data storage command-line interface

This chapter describes how to configure storage on the IBM DS8880 by using the data storage command-line interface (DS CLI).

This chapter covers the following topics:

- ► DS command-line interface overview
- ► I/O port configuration
- ► DS8880 storage configuration for fixed-block volumes
- ► DS8880 storage configuration for the CKD volumes
- ► Metrics with DS CLI
- ► Private network security commands
- ► Copy Services commands
- ► Further information

12.1 DS command-line interface overview

The DS CLI provides a full-function command set with which you can check your storage unit configuration and perform specific application functions. For more information about DS CLI use and setup, see the *IBM DS8000 Series Version 8 Release 2 Command-Line Interface User's Guide,* SC27-8526. available at:

http://www.ibm.com/support/docview.wss?uid=ssg1S7005236

The following list highlights a few of the functions that you can perform with the DS CLI:

► Create user IDs that can be used with the graphical user interface (GUI) and the DS CLI.

► Manage user ID passwords.

► Install activation keys for licensed features.

► Manage storage complexes and units.

► Configure and manage Storage Facility Images.

► Create and delete Redundant Array of Independent Disks (RAID) arrays, ranks, and extent pools.

► Create and delete logical volumes.

► Manage the host access to volumes.

► Check the current Copy Services configuration that is used by the storage unit.

► Create, modify, or delete Copy Services configuration settings.

► Integrate Lightweight Directory Access Protocol (LDAP) policy usage and configuration.

► Implement encryption functions.

> **Single installation:** In almost all cases, you can use a single installation of the latest version of the DS CLI for all of your system needs. However, it is not possible to test every version of DS CLI with every licensed machine code (LMC) level. Therefore, an occasional problem might occur despite every effort to maintain that level of compatibility. If you suspect a version incompatibility problem, install the DS CLI version that corresponds to the LMC level that is installed on your system. You can have more than one version of DS CLI installed on your system, each in its own directory.

12.1.1 Supported operating systems for the DS CLI

The DS CLI can be installed on many operating systems:
► AIX
► HP-UX
► Red Hat Linux
► SUSE Linux
► IBM i
► Oracle Solaris
► HP OpenVMS
► VMware ESX
► Microsoft Windows
► IBM z/OS
► Embedded DS CLI on Storage Management GUI

> **Important:** For information about the currently supported operating systems, specific preinstallation concerns, and installation file locations, see the IBM System Storage DS8000 IBM Knowledge Center at this website:
>
> https://www.ibm.com/support/knowledgecenter/ST5GLJ_8.2.0/com.ibm.storage.ssic.h
> elp.doc/f2c_ichomepage_v8.20.html

Before you can install the DS CLI, ensure that at least Java version 1.6 or later is installed. A suitable level of Java might already be installed on many hosts. The installation program checks for this requirement during the installation process and does not install the DS CLI if a suitable version of Java is not installed.

The installation process can be performed through a shell, such as the bash or Korn shell, or the Windows command prompt, or through a GUI. If the installation process is installed by using a shell, the installation can be a silent installation by using a profile file. The installation process also installs software that allows the DS CLI to be uninstalled when the DS CLI is no longer required.

For more information about Embedded DS CLI on DS GUI, see Chapter 11, "DS8880 Storage Management GUI" on page 251.

12.1.2 Installation hints and tips

In some cases and often on Windows systems with a non-English Windows version, you might get an error when trying to install the DS CLI. Figure 12-1 shows an error that is issued when trying to install the DS CLI on the German version of Windows.
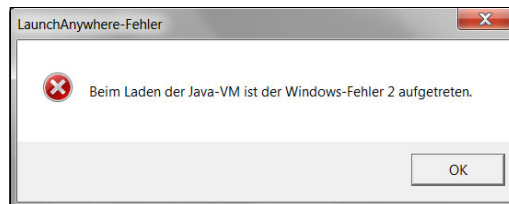
Figure 12-1   DSCLI installation error when java.exe is not found

It says that when loading the Java-VM, there was a Windows error 2. This error means that the `java.exe` file could not be found. For some reason, the DS CLI cannot find the Java code despite Java being available and path variables set.

If this error occurs, you must explicitly specify where to find the `java.exe` file. Open a CMD window and go to the subdirectory, where the DS CLI `dsclisetup.exe` file is located.

In this example, the path is `E:\IMAGES\HMC\Disk1\InstData\Windows\NoVM>`.

When you enter the **dsclisetup.exe** command, you must add the **LAX_VM** parameter and specify the full path to the `java.exe` file.

Here you might encounter another problem with non-English Windows. When you locate the `java.exe` file in your Browser, you might see a path as shown in Figure 12-2.
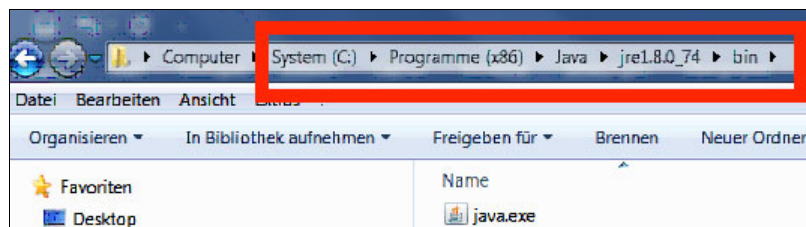
Figure 12-2   Location of the java.exe

You might think that `C:\Programme (x86)\Java\jre1.8.0.74\bin\` is the path to your `java.exe` file, but it is not. The directory `C:\Programme (x86)` does not exist as shown in Example 12-1.

Example 12-1   System cannot find directory

```
C:\>cd "Programme (x86)"
<<Das System kann den angegebenen Pfad nicht finden.
```

Although Windows shows a national translation in the Browser for the path, it internally uses the English version of the path. Therefore, the command shown in Example 12-2 works.

Example 12-2   Using the Windows internal path name

```
C:\>cd "Program Files (x86)"
C:\Program Files (x86)>
```

To get the correct path to use in the command line, click at the paths in the Browser as shown in Figure 12-3.
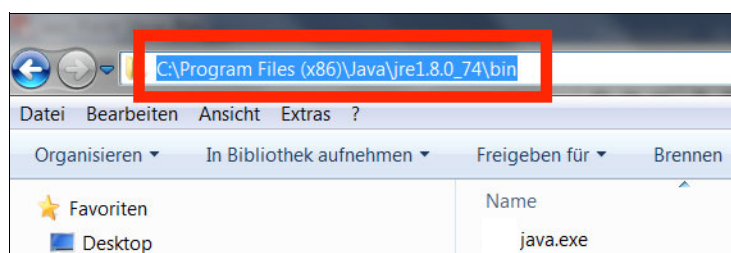
Figure 12-3   Windows internal path to java.exe

Copy this path and add `java.exe` to it, and run the command shown in Example 12-3 in the Windows command window.

Example 12-3   DSCLI installation with path to java.exe

```
E:\IMAGES\HMC\Disk1\InstData\Windows\NoVM>dsclisetup.exe LAX_VM "C:\Program Files
(x86)\Java\jre1.8.0_74\bin\java.exe"
```

12.1.3 Installing the DS CLI on a Windows 10 system

To install the DS CLI on a Windows 10 system, you must specify that you want to install it in Windows 7 or Windows 8 compatibility mode. On an English Windows, you can just right-click and select **Troubleshoot compatibility** and then the compatibility mode.

On a non-English Windows, you have to specify the path to your `java.exe` file. One way to do both is to first create a shortcut to `dsclisetup.exe` on your desktop (or anywhere else).

On the desktop, right-click the icon and select **Properties** (or **Eigenschaften**, in this example on the German version of Windows) and add the LAX_VM parameter with the full path to `java.exe` as shown in Figure 12-4.
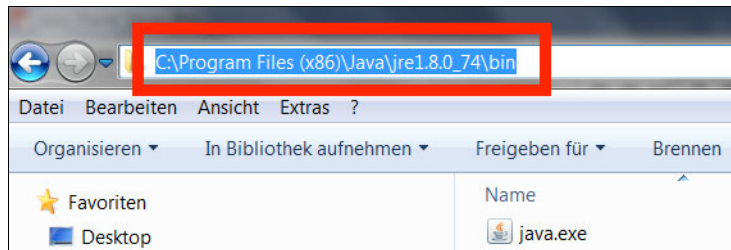
Figure 12-4   Modify shortcut and add LAX_VM parameter with path to java.exe

Now you can right-click the desktop shortcut to `dsclisetup.exe` *with the correct path* and select the **Troubleshoot compatibility** equivalent, which on the German Windows system is **Behandeln von Compatibilitaetsproblemen**, and select the Windows 7 or Windows 8 compatibility mode.

12.1.4 Installing the DS CLI on z/OS System

Starting with Release 8.2, you can also install the DS CLI along with Copy Services Manager on z/OS system. It is a regular SMP/E installation.

The DS CLI runs under UNIX Services for z/OS, and has a separate *FMID HIWN61K*. You can also install the DS CLI separately from Copy Services Manager.

For more information, see the *IBM DS CLI on z/OS Program Directory*. You can use the order number (GI13-3563) to search for it on the IBM Publications Center website:

`https://www.ibm.com/e-business/linkweb/publications/servlet/pbi.wss#`

After the installation is done, the first thing to do is to access your UNIX Services for z/OS. This process can vary from installation to installation. Ask your z/OS system programmer how to access it.

> **Tip:** Set your TSO REGION SIZE to 512 MB to allow the DS CLI to run.

In our test system, we logged on TSO and, using option 6, ISPF Command Shell, we issued the command OMVS to start the z/OS UNIX Shell. See Figure 12-5.

```
  Menu  List  Mode  Functions  Utilities  Help
  ------------------------------------------------------------------------------
  MCECEBC                          ISPF Command Shell
  Enter TSO or Workstation commands below:


  ===> omvs


 Place cursor on choice and press enter to Retrieve command

   => omvs
```

Figure 12-5   OMVS command to start z/OS UNIX Shell

When the initial panel is displayed, change your working directory by issuing the command **cd /opt/IBM/DSCLI**. See Figure 12-6.

```
IBM
Licensed Material - Property of IBM
...
GSA ADP Schedule Contract with IBM Corp.

IBM is a registered trademark of the IBM Corp.


-------------------------------------------------------------------------
 Welcome to the MCECEBC System in Mainz.

 Business Notice :
  IBM's internal systems must only be used for conducting IBM's
  business or for purposes authorized by IBM management.
-------------------------------------------------------------------------

$


 ===> cd /opt/IBM/CSMDSCLI

INPUT
ESC=¢      1=Help      2=SubCmd     3=HlpRetrn  4=Top       5=Bottom     6=TSO        7=BackScr
8=Scroll    9=NextSess 10=Refresh 11=FwdRetr  12=Retrieve
```

Figure 12-6   Command cd /opt/IBM/CSMDSCLI

You can choose between issuing the following commands to start DS CLI on z/OS: **./dscli**, or **./dscli -cfg ./profile/aca91.profile** if you decide to use a profile.

See (2) in Figure 12-7.

```
IBM
Licensed Material - Property of IBM
...
GSA ADP Schedule Contract with IBM Corp.

IBM is a registered trademark of the IBM Corp.


--------------------------------------------------------------------------
 Welcome to the MCECEBC System in Mainz.

 Business Notice :
   IBM's internal systems must only be used for conducting IBM's
   business or for purposes authorized by IBM management.
--------------------------------------------------------------------------


$ cd /opt/IBM/CSMDSCLI
$
 ===> ./dscli


INPUT
ESC=¢      1=Help      2=SubCmd     3=HlpRetrn  4=Top       5=Bottom    6=TSO       7=BackScr
8=Scroll     9=NextSess 10=Refresh 11=FwdRetr  12=Retrieve
```

Figure 12-7   The ./DSCLI command

If you change your mind and decide to quit here, instead of typing **.**/dscli, use **F2** from your keyboard to activate the *SubCmd* as shown in the Figure 12-7 (**2=SubCmd**). The OMVS Subcommand line is displayed, and you can issue a **quit** command.

As shown in Figure 12-8, the message *CEE5210S The signal SIGHUP was received* followed by *** *displays*. Press **Enter** to quit OMVS.

```
IBM
Licensed Material - Property of IBM
...
IBM is a registered trademark of the IBM Corp.
...
--------------------------------------------------------------------------
$ cd /opt/IBM/CSMDSCLI
$
 OMVS Subcommand ==> quit

SUBCOMMAND
ESC=       1=Help      2=SubCmd 3=Return    4=Top       5=Bottom    6=TSO       7=BackScr
8=Scroll     9=NextSess 10=Refresh 11=FwdRetr  12=Retrieve


----------------------------------------------------------
 CEE5210S The signal SIGHUP was received.
 ***
```

Figure 12-8   Sequence to leave DS CLI

Using the DS CLI on z/OS allows you to issue single commands, or use a script mode or go into batch mode, using z/OS JCL (Job Control Language). Figure 12-9 shows how to access the command interface.

```
 Business Notice :
   IBM's internal systems must only be used for conducting IBM's
   business or for purposes authorized by IBM management.
   -----------------------------------------------------------------------

 $ cd /opt/IBM/CSMDSCLI
 $ ./dscli
 Enter the primary management console IP address: <enter-your-machine-ip-address>
 Enter the secondary management console IP address:
 Enter your username: <enter-your-user-name-as-defined-on-the-machine>
 Enter your password: <enter-your-user-password-to-access-the-machine>
 dscli> ver -l
 ...
 dscli>
  ===>


 INPUT
 ESC=¢     1=Help      2=SubCmd    3=HlpRetrn  4=Top      5=Bottom   6=TSO      7=BackScr  8=Scroll
 9=NextSess 10=Refresh 11=FwdRetr  12=Retrieve
```

Figure 12-9   Accessing DS CLI on z/OS

The command that you run on DS CLI on z/OS has the same syntax as in other platforms. Some examples of those commands are shown in Figure 12-10.

```
 dscli> lssi
 Name          ID              Storage Unit    Model WWNN          State ESSNet
 ============================================================================================
 IBM.2107-75ACA91 IBM.2107-75ACA91 IBM.2107-75ACA90 980  5005076303FFD13E Online Enabled
 dscli> lsckdvol -lcu EF
 Name    ID  accstate datastate configstate deviceMTM voltype orgbvols extpool cap (cyl)
 ============================================================================================
 ITSO_EF00 EF00 Online  Normal    Normal     3390-A    CKD Base -    P1      262668
 ITSO_EF01 EF01 Online  Normal    Normal     3390-9    CKD Base -    P1       10017
 dscli> mkckdvol -dev IBM.2107-75ACA91 -cap 3339 -datatype 3390 -eam rotateexts -name ITSO_#h -extpool P1
 EF02-EF02
 CMUC00021I mkckdvol: CKD Volume EF02 successfully created.
 dscli> lsckdvol -lcu EF
 Name    ID  accstate datastate configstate deviceMTM voltype orgbvols extpool cap (cyl)
 ============================================================================================
 ITSO_EF00 EF00 Online  Normal    Normal     3390-A    CKD Base -    P1      262668
 ITSO_EF01 EF01 Online  Normal    Normal     3390-9    CKD Base -    P1       10017
 ITSO_EF02 EF02 Online  Normal    Normal     3390-3    CKD Base -    P1        3339
 dscli> rmckdvol EF02
 CMUC00023W rmckdvol: The alias volumes associated with a CKD base volume are automatically deleted
 before deletion of the CKD base volume. Are you sure you want to delete CKD volume EF02? Ỹy/n¨: y
 CMUC00024I rmckdvol: CKD volume EF02 successfully deleted.
 dscli>
  ===>
```

Figure 12-10   Common commands on DS CLI

Below are several illustrations of how to take advantage of using your own JCL:

► Example 12-4 shows a job to run a DS CLI script (multiple commands, stored in a UNIX file, in script mode).

Example 12-4   Run multiple commands from a UNIX file

```
//job-card-goes-here
//*****************************************************************/
//*                                                             */
//*  Run a DSCLI script in z/OS                                 */
//*                                                             */
//*  DSCLI on z/OS = CSMDSCLI                                   */
//*                                                             */
//*  All output (stdin and stderr) is directed to the job log   */
//*  The actual UNIX command follows the SH statement, it can span */
//*  more than one line                                         */
//*                                                             */
//*****************************************************************/
//*
//PBXBAT  EXEC PGM=BPXBATCH
//STDIN  DD DUMMY
//STDOUT DD SYSOUT=*
//STDERR DD SYSOUT=*
//SYSOUT DD SYSOUT=*
//STDENV DD DUMMY
//STDPARM DD *
SH
/opt/IBM/CSMDSCLI/dscli -cfg /opt/IBM/CSMDSCLI/profile/aca91.profile
    -script /u/itso/dscli_script.txt
/*
```

► Example 12-5 shows a JCL to run a single command by using a script.

Example 12-5   Single command by using a script

```
//job-card-goes-here
//*****************************************************************/
//*                                                             */
//*  Run a DSCLI script in z/OS                                 */
//*                                                             */
//*  DSCLI on z/OS = CSMDSCLI                                   */
//*                                                             */
//*  All output (stdin and stderr) is directed to the job log   */
//*  The actual UNIX command follows the SH statement, it can span */
//*  more than one line                                         */
//*                                                             */
//*****************************************************************/
//*
//PBXBAT  EXEC PGM=BPXBATCH
//STDIN  DD DUMMY
//STDOUT DD SYSOUT=*
//STDERR DD SYSOUT=*
//SYSOUT DD SYSOUT=*
//STDENV DD DUMMY
//STDPARM DD *
SH
```

```
/opt/IBM/CSMDSCLI/dscli -cfg /opt/IBM/CSMDSCLI/profile/aca91.profile
    -script /u/itso/dscli_script.txt
/*
```

► Example 12-6 shows a JCL to run several commands in a row, each in "single shot mode."

Example 12-6   Several "single shot mode" commands

```
//job-card-goes-here
//******************************************************************/
//*                                                              */
//*  Run several DSCLI command in z/OS                           */
//*                                                              */
//*  DSCLI on z/OS = CSMDSCLI                                    */
//*                                                              */
//*  All output (stdin and stderr) is directed to the job log   */
//*  The actual UNIX commands follow the SH statement, they can  */
//*  span more than one line                                     */
//*                                                              */
//*  Commands are separated by semicolon                        */
//*                                                              */
//******************************************************************/
//*
//PBXBAT  EXEC PGM=BPXBATCH
//STDIN  DD DUMMY
//STDOUT DD SYSOUT=*
//STDERR DD SYSOUT=*
//SYSOUT DD SYSOUT=*
//STDENV DD DUMMY
//STDPARM DD *
SH
echo "Command 1:";
/opt/IBM/CSMDSCLI/dscli -cfg /opt/IBM/CSMDSCLI/profile/aca91.profile
    ver -l;
echo "Command 2:";
/opt/IBM/CSMDSCLI/dscli -cfg /opt/IBM/CSMDSCLI/profile/aca91.profile
    lsrank;
echo "Command 3:";
/opt/IBM/CSMDSCLI/dscli -cfg /opt/IBM/CSMDSCLI/profile/aca91.profile
    lsarray;
/*
```

12.1.5 DS CLI version

The **ver** command displays the version of the DS CLI client program, the Hardware Management Console (HMC) code level (*Storage Manager*), the HMC DS CLI version, the LMC version, and the code bundle version. The **ver** command uses the following parameters:

-s (optional): The **-s** parameter displays the version of the DS CLI client program. You cannot use the **-s** and **-l** parameters together.

-l (optional): The **-l** parameter displays the versions of the DS CLI client program, Storage Manager, HMC code level, licensed machine code, and the code bundle. You cannot use the **-l** and **-s** parameters together. See Example 12-7.

-cli (optional): The **-cli** parameter displays the version of the DS CLI client program. Version numbers are in this format: *version.release.modification.fixlevel*.

-stgmgr (optional): The **-stgmgr** parameter displays the version of the Storage Manager. This ID is not the Storage Manager GUI. This ID relates to the HMC code level information.

-lmc (optional): The **-lmc** parameter displays the version of the LMC.

Example 12-7 shows an example of the **ver** command.

Example 12-7   DS CLI version command

```
dscli> ver -l
...
DSCLI          7.8.20.186
StorageManager 8.8.4.0.20160818.1
HMC DSCLI      7.8.20.186
===============Version=================
Storage Image   LMC          Bundle Version
========================================
IBM.2107-75ACA91 7.8.20.186 88.20.71.0
```

The following DS CLI commands are added with Release 8.2 to check the CSM version and install new software level of CSM:

► The **lssoftware -l** command displays the CSM version that is installed on both HMCs. See Example 12-8.

Example 12-8   Display CSM Version

```
dscli> lssoftware -l

IBM DSCLI Version: 7.8.20.261 DS: IBM.2107-75DLR71
Type Version                  Status  HMC
=======================================
 CSM V6.1.3-a20160721-2100 Running 1
 CSM V6.1.3-a20160721-2100 Running 2
```

► The **installsoftware** command is used to install a new version of CSM software. See Example 12-9.

Example 12-9   The installsoftware command output

```
dscli> installsoftware -loc
/home/hscroot/wht/csm-setup-6.1.1.2-linux-x86_64.bin -certloc
/home/hscroot/wht/csm-setup-6.1.1.2-linux-x86_64.bin.crt -type CSM -hmc 1
```

```
IBM DSCLI Version: 0.0.0.0 DS: IBM.2107-75DMC41
CMUC00294I installsoftware: Upload file successfully.
CMUC00294I installsoftware: Software CSM is successfully installed on HMC 1.
```

12.1.6 User accounts

DS CLI communicates with the DS8880 through the HMC. The primary or secondary HMC can be used. The DS CLI access is authenticated by using IBM Enterprise Storage Server Network Interface (ESSNI), which is also referred to as the DS Network Interface (DSNI) on the HMC. The same user ID is used for DS CLI and DS GUI access. For more information about user accounts, see 8.5, "Management Console user management" on page 214. The default user ID is admin and the password is admin. The system forces you to change the password at the first login. If you forget the admin password, a reset can be performed that resets the admin password to the default value.

12.1.7 User management by using the DS CLI

Apart from the administration user, you might want to define other users, maybe with different authorities.

The following commands are used to manage user IDs by using the DS CLI:

► **mkuser**

A user account that can be used with the DS CLI and the DS GUI is created by using this command. Example 12-10 shows the creation of a user that is called JohnDoe, which is in the op_storage group. The temporary password of the user is passw0rd. The user must use the **chpass** command when they log in for the first time.

Example 12-10   Use the mkuser command to create a user

```
dscli> mkuser -pw passw0rd -group op_storage JohnDoe
CMUC00133I mkuser: User JohnDoe successfully created.
```

► **rmuser**

An existing user ID is removed by using this command. Example 12-11 shows the removal of a user called JaneSmith.

Example 12-11   Remove a user

```
dscli> rmuser JaneSmith
CMUC00135W rmuser: Are you sure you want to delete user JaneSmith? [y/n]:y
CMUC00136I rmuser: User JaneSmith successfully deleted.
```

► **chuser**

Use this command to change the password or group (or both) of an existing user ID. It also can be used to unlock a user ID that was locked by exceeding the allowable login retry count. The administrator can also use this command to lock a user ID. In Example 12-12, unlock the user, change the password, and change the group membership for a user that is called JohnDoe. The user must use the **chpass** command the next time that they log in.

Example 12-12   Change a user with chuser

```
dscli> chuser -unlock -pw time2change -group op_storage JohnDoe
CMUC00134I chuser: User JohnDoe successfully modified.
```

► **lsuser**

By using this command, a list of all user IDs can be generated. Example 12-13 shows a list of three users, including the administrator account.

Example 12-13   Use the lsuser command to list users

```
dscli> lsuser
Name           Group                      State
==============================================
JohnDoe        op_storage                 active
secadmin       admin                      active
admin          admin                      active
```

► **showuser**

The account details of a user ID can be displayed by using this command. Example 12-14 lists the details of the user JohnDoe.

Example 12-14   Use the showuser command to list user information

```
dscli> showuser JohnDoe
Name        JohnDoe
Group       op_storage
State       active
FailedLogin 0
DaysToExpire 365
Scope       PUBLIC
```

► **managepwfile**

An encrypted password file that is placed onto the local machine is created or added by using this command. This file can be referred to in a DS CLI profile. You can run scripts without specifying a DS CLI user password in clear text. If you are manually starting the DS CLI, you can also refer to a password file with the **-pwfile** parameter. By default, the file is in the following directories:

– Windows: C:\Users*<User>*\dscli\security.dat
– Other than Windows: HOME/dscli/security.dat

Example 12-15 shows managing the password file by adding the user ID JohnDoe. The password is now saved in an encrypted file that is called security.dat.

Example 12-15   Use the managepwfile command

```
dscli> managepwfile -action add -name JohnDoe -pw passw0rd
CMUC00206I managepwfile: Record 10.0.0.1/JohnDoe successfully added to password
file C:\Users\Administrator\dscli\security.dat.
```

► **chpass**

By using this command, you can change two password policies: Password expiration (in days) and the number of failed logins that are allowed. Example 12-16 shows changing the expiration to 365 days and five failed login attempts.

Example 12-16   Change rules by using the chpass command

```
dscli> chpass -expire 365 -fail 5
CMUC00195I chpass: Security properties successfully set.
```

► **showpass**

The properties for passwords (Password Expiration days and Failed Logins Allowed) are listed by using this command. Example 12-17 shows that passwords are set to expire in 90 days and that four login attempts are allowed before a user ID is locked.

Example 12-17   Use the showpass command

```
dscli> showpass
Password Expiration   365 days
Failed Logins Allowed 5
Password Age          0 days
Minimum Length        6
Password History      4
```

12.1.8 DS CLI profile

To access the DS8880 with the DS CLI, you must provide certain information with the `dscli` command. At a minimum, the IP address or host name of the DS8880 HMC, a user name, and a password are required. You can also provide other information, such as the output format for list commands, the number of rows for each page in the command-line output, and whether a banner is included with the command-line output.

If you create one or more profiles to contain your preferred settings, you do not need to specify this information each time that you use the DS CLI. When you start the DS CLI, you can specify a profile name by using the `dscli` command. You can override the values of the profile by specifying a different parameter value with the `dscli` command.

When you install the command-line interface software, a default profile is installed in the profile directory with the software. The file name is `dscli.profile`. For example, use **`c:\Program Files (x86)\IBM\dscli`** for Windows 7 (and later) and **`/opt/ibm/dscli/profile/dscli.profile`** for AIX (UNIX) and Linux platforms.

The following options are available for using profile files:

► You can modify the system default profile: `dscli.profile`.

► You can create a personal default profile by copying the system default profile as *<user_home>*`/dscli/profile/dscli.profile`. The default home directory *<user_home>* is in the following directories:

  – Windows system: `%USERPROFILE%` usually `C:\Users\Administrator`
  – UNIX/Linux system: `$HOME`

► You can create specific profiles for different storage units and operations. Save the profile in the user profile directory, for example:

  – `%USERPROFILE%\IBM\DSCLI\profile\operation_name1`
  – `%USERPROFILE%\IBM\DSCLI\profile\operation_name2`

Default profile file: The default profile file that you created when you installed the DS CLI might be replaced every time that you install a new version of the DS CLI. It is a good practice to open the default profile and then save it as under a new file. You can then create multiple profiles and reference the relevant profile file by using the **`-cfg`** parameter. The following example uses a different profile when it starts the DS CLI: **`dscli -cfg newprofile.profile`** (or whatever name you gave to the new profile).

These profile files can be specified by using the DS CLI command parameter **-cfg**
<profile_name>. If the **-cfg** file is not specified, the default profile of the user is used. If a
profile of a user does not exist, the system default profile is used.

> **Two default profiles:** If two default profiles are called `dscli.profile`, one profile in the
> default system's directory and one profile in your personal directory, your personal profile is
> loaded.

Profile change illustration

Complete the following steps to edit the profile. This sequence assumes that your
`%userprofile%` is `C:\Users\Administrator.`

1. Use Windows Explorer to copy the profile folder from `C:\Program Files (x86)\IBM\dscli`
   to `C:\Users\Administrator\dscli`.

2. From the Windows desktop, double-click the DS CLI icon.

3. In the command window that opens, enter the following command:

   `cd C:\Users\Administrator\dscli\`

4. In the profile directory, enter the **notepad dscli.profile** command, as shown in
   Example 12-18.

Example 12-18   Command prompt operation

```
C:\Users\Administrator\dscli>cd profile
C:\Users\Administrator\dscli\profile>notepad dscli.profile
```

5. The notepad opens and includes the DS CLI profile. Consider adding four lines. Examples
   of these lines are shown in bold in Example 12-19.

> **Default newline delimiter:** The default newline delimiter is a UNIX delimiter, which can
> render text in the notepad as one long line. Use a text editor that correctly interprets
> UNIX line endings.

Example 12-19   DSCLI profile example

```
# DS CLI Profile
#
# Management Console/Node IP Address(es)
#   hmc1 and hmc2 are equivalent to -hmc1 and -hmc2 command options.
#hmc1:127.0.0.1
#hmc2:127.0.0.1

# Default target Storage Image ID
#    "devid" and "remotedevid" are equivalent to
#    "-dev storage_image_ID" and "-remotedev storeage_image_ID" command
options, respectively.
#devid:  IBM.2107-AZ12341
#remotedevid:IBM.2107-AZ12341

devid:    IBM.2107-75ABCD1
hmc1:     10.0.0.250
username: admin
password: passw0rd
```

Adding the serial number by using the **devid** parameter and adding the HMC IP address by using the **hmc1** parameter are suggested. Not only does this addition help you to avoid mistakes when you use more profiles, but you do not need to specify this parameter for certain `dscli` commands that require it. Additionally, if you specify the `dscli` profile for Copy Services usage, the use of the **remotedevid** parameter is suggested for the same reasons. To determine the ID of a storage system, use the **lssi** CLI command.

Although adding the **username** and **password** parameters simplifies the DS CLI start, generally do not add them because they are an undocumented feature that might not be supported in the future. Also, the password is saved in clear text in the profile file.

Instead, it is better to create an encrypted password file with the **managepwfile** command. A password file that is generated by using the **managepwfile** command is placed in the `user_home_directory/dscli/profile/security/security.dat` directory.

> **Important:** Be careful if you add multiple **devid** and HMC entries. Uncomment (or literally, unhash) only one entry at a time. If multiple **hmc1** or **devid** entries exist, the DS CLI uses the entry that is closest to the bottom of the profile.

The following customization parameters also affect **dscli** output:

- **banner**: Date and time with the **dscli** version are printed for each command.
- **header**: Column names are printed.
- **format**: The output format, which is specified as **default**, **xml**, **delim**, or **stanza**.
- **paging**: For interactive mode, this parameter breaks output after several rows (24, by default).

6. After you save your changes, use Windows Explorer to copy the updated profile from `C:\Users\Administrator\dscli\profile` to `C:\Program Files (x86)\IBM\dscli\profile`.

12.1.9 Configuring the DS CLI to use a second HMC

The second HMC can be specified on the command line or in the profile file that is used by the DS CLI. To specify the second HMC in a command, use the **-hmc2** parameter, as shown in Example 12-20.

Example 12-20 Use the -hmc2 parameter

```
C:\Program Files (x86)\IBM\dscli>dscli -hmc1 10.0.0.1 -hmc2 10.0.0.5
Enter your user name: JohnDoe
Enter your password: xxxxx
IBM.2107-75DMC11
dscli>
```

Alternatively, you can modify the following lines in the `dscli.profile` (or any profile) file:

```
# Management Console/Node IP Address(es)
# hmc1 and hmc2 are equivalent to -hmc1 and -hmc2 command options.
hmc1:10.0.0.1
hmc2:10.0.0.5
```

After these changes are made and the profile is saved, the DS CLI automatically communicates through HMC2 if HMC1 becomes unreachable. By using this change, you can perform configuration and execute Copy Services commands with full redundancy.

> **Two HMCs:** If you use two HMCs and you specify only one of them in a DS CLI command (or profile), any changes that you make to users are still replicated onto the other HMC.

12.1.10 Command structure

This section describes the components and structure of a DS CLI command. A DS CLI command consists of 1 - 4 types of components that are arranged in the following order:

1. The command name: Specifies the task that the DS CLI will perform.

2. Flags: Flags are used to modify the command. They provide more information that directs the DS CLI to perform the command task in a specific way.

3. Flags parameter: Provides information that is required to implement the command modification that is specified by a flag.

4. Command parameters: Provide basic information that is necessary to perform the command task. When a command parameter is required, it is always the last component of the command, and is not preceded by a flag.

12.1.11 Using the DS CLI application

To issue commands to the DS8880, you must first log in to the DS8880 through the DS CLI with one of the following command modes of execution:

► Single-shot command mode
► Interactive command mode
► Script command mode

Single-shot command mode

Use the DS CLI single-shot command mode if you want to issue an occasional command from the operating system (OS) shell prompt where you need special handling, such as redirecting the DS CLI output to a file. You also use this mode if you are embedding the command into an OS shell script.

You must supply the login information and the command that you want to process at the same time. Complete the following steps to use the single-shot mode:

1. At the OS shell prompt, enter the following command:

   `dscli -hmc1 <hostname or ip address> -user <adm user> -passwd <pwd> <command>`

   or

   `dscli -cfg <dscli profile> -pwfile <security file> <command>`

   > **Important:** Avoid embedding the user name and password into the profile. Instead, use the **-pwfile** command.

2. Wait for the command to process and display the results.

Example 12-21 shows the use of the single-shot command mode.

Example 12-21   Single-shot command mode

```
C:\Program Files (x86)\IBM\dscli>dscli -hmc1 10.10.10.1 -user admin -passwd <pwd>
lsuser
Name          Group                         State
```

```
=================================================
AlphaAdmin    admin                      locked
AlphaOper     op_copy_services           active
BetaOper      op_copy_services           active
admin         admin                      active
[ exit status of dscli = 0 ]
```

> **Important:** When you are typing the command, you can use the host name or the IP address of the HMC. It is important to understand that when a command is run in single-shot mode, the user must be authenticated. The authentication process can take a considerable amount of time.

Interactive command mode

Use the DS CLI interactive command mode when you want to issue a few infrequent commands without needing to log on to the DS8880 for each command.

The interactive command mode provides a history function that simplifies repeating or checking earlier command usage.

Complete the following steps to use the interactive command mode:

1. Log on to the DS CLI application at the directory where it is installed.

2. Provide the information that is requested by the information prompts. The information prompts might not appear if you provided this information in your profile file. The command prompt switches to a **dscli** command prompt.

3. Use the DS CLI commands and parameters. You are not required to begin each command with **dscli** because this prefix is provided by the **dscli** command prompt.

4. Use the **quit** or **exit** command to end interactive mode.

> **Interactive mode:** In interactive mode for long outputs, the message `Press Enter To Continue` appears. The number of rows can be specified in the profile file. Optionally, you can turn off the paging feature in the profile file by using the **paging:off** parameter.

Example 12-22 shows the use of interactive command mode by using the profile DS8880.profile.

Example 12-22   Interactive command mode

```
C:\Program Files (x86)\IBM\dscli>dscli -cfg DS8880.profile

dscli> lsarraysite -l

...
arsite DA Pair dkcap (10^9B) diskrpm State       Array diskclass encrypt
==========================================================================
S1     2         1200.0      10000 Assigned    A20   ENT       supported
S2     2         1200.0      10000 Assigned    A18   ENT       supported
S3     2         1200.0      10000 Unassigned  -     ENT       supported
S4     2         1200.0      10000 Assigned    A5    ENT       supported
S5     2         1200.0      10000 Assigned    A3    ENT       supported
S6     2         1200.0      10000 Assigned    A1    ENT       supported
S7     2          600.0      15000 Assigned    A13   ENT       supported
S8     2          600.0      15000 Assigned    A10   ENT       supported
S9     2          600.0      15000 Assigned    A17   ENT       supported
S10    2          600.0      15000 Assigned    A19   ENT       supported
```

```
S11    2                  600.0   15000 Assigned   A23   ENT       supported
S12    2                  600.0   15000 Assigned   A21   ENT       supported
S13    2                  600.0   15000 Assigned   A14   ENT       supported
dscli> lssi
...
Name              ID                Storage Unit      Model WWNN             State ESSNet
===========================================================================================
IBM.2107-75ACA91 IBM.2107-75ACA91 IBM.2107-75ACA90 980    5005076303FFD13E Online Enabled
dscli>
```

Script command mode

Use the DS CLI script command mode if you want to use a sequence of DS CLI commands. If you want to run a script that contains only DS CLI commands, you can start the DS CLI in script mode. The script that DS CLI runs can contain only DS CLI commands.

Example 12-23 shows the contents of a DS CLI script file. The file contains only DS CLI commands, although comments can be placed in the file by using a number sign (#). Empty lines are also allowed. One advantage of using this method is that scripts that are written in this format can be used by the DS CLI on any operating system that you can install the DS CLI on. Only one authentication process is needed to run all of the script commands.

Example 12-23   Example of a DS CLI script file

```
# Sample dscli script file
# Comments can appear if hashed
lsarraysite -l
lsarray -l
lsrank -l
```

For script command mode, you can turn off the banner and header for easier output parsing. Also, you can specify an output format that might be easier to parse by your script.

Example 12-24 shows starting the DS CLI by using the **-script** parameter and specifying a profile and the name of the script that contains the commands from Example 12-23.

Example 12-24   Run a DS CLI file

```
C:\Program Files (x86)\IBM\dscli>dscli -cfg DS8880.profile -script c:\ds8000.script
arsite DA Pair dkcap (10^9B) diskrpm State    Array diskclass encrypt
======================================================================
S13    2           4000.0   7200 Assigned A2    NL      supported
S14    2           4000.0   7200 Assigned A3    NL      supported
S15    2           4000.0   7200 Assigned A4    NL      supported
S16    2            400.0  65000 Assigned A0    SSD     supported
S17    2            400.0  65000 Assigned A1    SSD     supported
S24    10           400.0  65000 Assigned A23   Flash   supported
S25    10           400.0  65000 Assigned A24   Flash   supported
S26    10           400.0  65000 Assigned A25   Flash   supported
S27    10           400.0  65000 Assigned A26   Flash   supported
Array State    Data    RAIDtype    arsite Rank DA Pair DDMcap (10^9B) diskclass encrypt
========================================================================================
A0    Assigned Normal 5 (6+P)     S25   R0   18           400.0 Flash     supported
A1    Assigned Normal 5 (7+P)     S6    R1   2           1200.0 ENT       supported
A2    Assigned Normal 5 (6+P)     S24   R2   18           400.0 Flash     supported
A3    Assigned Normal 5 (7+P)     S5    R3   2           1200.0 ENT       supported
A4    Assigned Normal 5 (6+P+S)   S23   R4   18           400.0 Flash     supported
A5    Assigned Normal 5 (6+P+S)   S4    R5   2           1200.0 ENT       supported
A6    Assigned Normal 5 (6+P+S)   S22   R6   18           400.0 Flash     supported
A8    Assigned Normal 6 (5+P+Q+S) S21   R8   2           4000.0 NL        supported
```

```
ID   Group State  datastate Array RAIDtype extpoolID extpoolnam  stgtype exts     usedexts encryptgrp marray extsize (cap)
==============================================================================================================================
R0   0 Normal Normal    A0    5 P0         CKD_L       ckd     2392         31        - MA25   1113cyl
R1   0 Normal Normal    A1    5 P0         CKD_L       ckd     8313        814        - MA6    1113cyl
R2   1 Normal Normal    A2    5 P1         CKD_L       ckd     2392         31        - MA24   1113cyl
R3   1 Normal Normal    A3    5 P1         CKD_L       ckd     8313        800        - MA5    1113cyl
R4   0 Normal Normal    A4    5 P2         FB_Gold_L   fb      2132        208        - MA23   1GiB
R5   0 Normal Normal    A5    5 P2         FB_Gold_L   fb      6348       1572        - MA4    1GiB
R6   1 Normal Normal    A6    5 P3         FB_Gold_L   fb      2132       1818        - MA22   1GiB
R8   0 Normal Normal    A8    6 P4         FB_Silver_L fb     17261        181        - MA21   1GiB
R9   0 Normal Normal    A9    5 P4         FB_Silver_L fb      3688         20        - MA18   1GiB
R10  0 Normal Normal    A10   5 P4         FB_Silver_L fb      3156         16        - MA8    1GiB
R11  1 Normal Normal    A11   6 P5         FB_Silver_L fb     17261         87        - MA20   1GiB
R12  1 Normal Normal    A12   5 P5         FB_Silver_L fb      3688         11        - MA17   1GiB
R13  1 Normal Normal    A13   5 P5         FB_Silver_L fb      3156         13        - MA7    1GiB
R14  0 Normal Normal    A14   5 P6         FB_Bronze_S fb    236053       2918        - MA13   16MiB
R15  1 Normal Normal    A15   5 P7         FB_Bronze_S fb    236053      51278        - MA16   16MiB
R16  0 Normal Normal    A16   6 P6         FB_Bronze_S fb   1104704     103832        - MA19   16MiB
R17  0 Normal Normal    A17   5 P8         CKD_S       ckd   219245      23672        - MA9    21cyl
R18  0 Normal Normal    A18   5 P8         CKD_S       ckd   377388      23672        - MA2    21cyl
```

> **Important:** The DS CLI script can contain only DS CLI commands. The use of shell commands results in process failure.

12.1.12 Return codes

When the DS CLI exits, the exit status code is provided, which effectively is a return code. If DS CLI commands are issued as separate commands (rather than by using script mode), a return code is presented for every command. If a DS CLI command fails (for example, because of a syntax error or the use of an incorrect password), a failure reason and return code are shown. Standard techniques to collect and analyze return codes can be used.

The return codes that are used by the DS CLI are listed in the *Command-Line Interface User's Guide*, SC27-8526.

12.1.13 User assistance

The DS CLI is designed to include several forms of user assistance. The main form of user assistance is through the IBM DS8000 IBM Knowledge Center, which is available at this website:

https://www.ibm.com/support/knowledgecenter/ST5GLJ/ds8000_kcwelcome.html

Look under the Command-line interface tab. You can also get user assistance when using the DS CLI program through the **help** command. The following examples of usage are included:

► **help** lists all of the available DS CLI commands.
► **help -s** lists all of the DS CLI commands with brief descriptions of each command.
► **help -l** lists all of the DS CLI commands with their syntax information.

To obtain information about a specific DS CLI command, enter the command name as a parameter of the **help** command. The following examples of usage are included:

► **help** *<command name>* provides a detailed description of the specified command.
► **help -s** *<command name>* provides a brief description of the specified command.
► **help -l** *<command name>* provides syntax information about the specified command.

Man pages

A *man page* is available for every DS CLI command. Man pages are most commonly seen in UNIX operating systems, and provide information about command capabilities. This information can be displayed by issuing the relevant command followed by the `-h`, `-help`, or `-?` flags.

12.2 I/O port configuration

Set the I/O ports to the topology that you want. Example 12-25 lists the I/O ports by using the `lsioport` command. `I0000 - I0003` are on one card, and `I0100 - I0103` are on another card.

Example 12-25   List the I/O ports

```
dscli> lsioport -dev IBM.2107-75DMC11
ID    WWPN             State Type            topo    portgrp
================================================================
I0000 500507630300008F Online Fibre Channel-SW SCSI-FCP 0
I0001 500507630300408F Online Fibre Channel-SW SCSI-FCP 0
I0002 500507630300808F Online Fibre Channel-SW SCSI-FCP 0
I0003 500507630300C08F Online Fibre Channel-SW SCSI-FCP 0
I0100 500507630308008F Online Fibre Channel-LW FICON    0
I0101 500507630308408F Online Fibre Channel-LW SCSI-FCP 0
I0102 500507630308808F Online Fibre Channel-LW FICON    0
I0103 500507630308C08F Online Fibre Channel-LW FICON    0
```

The following possible topologies for each I/O port are available:

► Small Computer System Interface - Fibre Channel Protocol (SCSI-FCP): Fibre Channel-switched fabric, which is also called *switched point-to-point*. This port type is also used for mirroring.

► Fibre Channel Arbitrated Loop (FC-AL): This port type is for direct attachment without a storage area network (SAN) switch.

► Fibre Channel connection (FICON): This port type is for z Systems hosts only.

If added to the `setioport` command, the `-force` parameter will allows a topology change to an online I/O port even if a current topology is set. Example 12-26 shows setting I/O ports without and with `-force` option to the FICON topology, and then checking the results.

Example 12-26   Change the topology by using setioport

```
dscli> lsioport i0231
ID    WWPN             State Type            topo  portgrp
================================================================
I0231 500507630313513E Online Fibre Channel-SW SCSI-FCP 0
dscli> setioport -topology ficon i0231
CMUN04008E setioport: I0231: The change topology request failed because the
current topology is already set and force option is not provided.
dscli> setioport -topology ficon -force i0231
CMUC00011I setioport: I/O Port I0231 successfully configured.
dscli> lsioport i0231
ID    WWPN             State Type            topo  portgrp
============================================================
I0231 500507630313513E Online Fibre Channel-SW FICON 0
```

12.3 DS8880 storage configuration for fixed-block volumes

This section reviews examples of a typical DS8880 storage configuration when the DS8880 storage is attached to open system hosts. You can perform the DS8880 storage configuration by completing the following steps:

1. Create the arrays.
2. Create the ranks.
3. Create the extent pools.
4. Optional: Create the repositories for space-efficient volumes (not included).
5. Create the volumes.
6. Create the volume groups.
7. Create the host connections.

12.3.1 Disk classes

Arrays and array sites are associated with the following disk classes (`diskclass`), which represent a classification of the different drive types that are available for the DS8880:

► Nearline (NL disk class) specifies 3.5-inch large form factor (LFF) serial-attached SCSI (SAS) nearline drives with a 7.2K revolutions per minute (RPM) drive speed.

► Enterprise (ENT disk class) specifies 2.5-inch small form factor (SFF) SAS enterprise drives with 10K or 15K RPM drive speed.

► Solid-state drive (SSD disk class) specifies 2.5-inch SFF SAS Flash drives, which are also known as SSDs.

► Flash (Flash disk class) specifies 2.5-inch High Performance flash cards. They are part of the high-performance flash enclosure (HPFE).

Both disk classes SSD and Flash are part of the SSD storage tier in terms of the IBM Easy Tier feature.

Important: For information about the current drive choices and RAID capacities, see the IBM Knowledge Center at:

http://www.ibm.com/support/knowledgecenter/ST5GLJ_8.2.0/com.ibm.storage.ssic.help.doc/f2c_diskdrives_lfg4me.html

and

http://www.ibm.com/support/knowledgecenter/ST5GLJ_8.2.0/com.ibm.storage.ssic.help.doc/f2c_effectivecapa_ds8880.html

12.3.2 Creating the arrays

This step creates the arrays. Before the arrays are created, list the array sites. Use the **lsarraysite** command to list the array sites, as shown in Example 12-27. *Array sites* are groups of eight drives that are predefined in the DS8880.

> **Important:** An *array* for a DS8880 contains only one array site, and a DS8880 *array site* contains eight drives.

Example 12-27   List array sites

```
dscli> lsarraysite -l
arsite DA Pair dkcap (10^9B) diskrpm State       Array diskclass encrypt
=============================================================================
S1     11            400.0    65000 Assigned   A0    Flash     supported
S2     11            400.0    65000 Assigned   A5    Flash     supported
S13    2             600.0    10000 Assigned   A12   ENT       supported
S14    2             600.0    10000 Assigned   A13   ENT       supported
S25    0            4000.0     7200 Assigned   A24   NL        supported
S26    0            4000.0     7200 Assigned   A25   NL        supported
S27    0            4000.0     7200 Assigned   A26   NL        supported
S32    3             400.0    65000 Assigned   A31   SSD       supported
S33    3             400.0    65000 Assigned   A32   SSD       supported
S38    1             600.0    10000 Unassigned -     ENT       supported
S39    1             600.0    10000 Unassigned -     ENT       supported
```

In Example 12-27, you can see two unassigned array sites. Therefore, you can create two arrays. The **-l** option reports the `diskclass` information.

You can issue the **mkarray** command to create arrays, as shown in Example 12-28. The example uses one array site (in the first array, S1) to create a single RAID 5 array. If you want to create a RAID 10 array, change the **-raidtype** parameter to 10. If you want to create a RAID 6 array, change the **-raidtype** parameter to 6 (instead of 5).

Example 12-28   Create arrays with mkarray

```
dscli> mkarray -raidtype 5 -arsite S38
CMUC00004I mkarray: Array A34 successfully created.
dscli> mkarray -raidtype 5 -arsite S39
CMUC00004I mkarray: Array A35 successfully created.
```

You can now see the arrays that were created by using the **lsarray** command, as shown in Example 12-29.

Example 12-29   Listing the arrays with lsarray

```
dscli> lsarray -l
Array State      Data   RAIDtype     arsite Rank DA Pair DDMcap (10^9B) diskclass encrypt
===========================================================================================
A0    Assigned   Normal 5 (6+P+S)    S1     R0   11            400.0 Flash     supported
A5    Assigned   Normal 5 (6+P+S)    S2     R9   11            400.0 Flash     supported
A12   Assigned   Normal 5 (6+P+S)    S13    R10  2             600.0 ENT       supported
A13   Assigned   Normal 5 (6+P+S)    S14    R11  2             600.0 ENT       supported
A25   Assigned   Normal 6 (5+P+Q+S)  S26    R20  0            4000.0 NL        supported
A26   Assigned   Normal 6 (5+P+Q+S)  S27    R24  0            4000.0 NL        supported
A31   Unassigned Normal 5 (6+P+S)    S32    -    3             400.0 SSD       supported
```

```
A32    Unassigned Normal 5 (6+P+S)    S33    -    3    400.0 SSD    supported
A34    Unassigned Normal 5 (7+P)      S38    -    1    600.0 ENT    supported
A35    Unassigned Normal 5 (7+P)      S39    -    1    600.0 ENT    supported
```

You can see the following information in this example:

► Type of RAID array (RAID 5)

► Number of drives that are allocated to the array (7+P, which means that the usable space of the array is seven times the drive size)

► Capacity of the drives that are being used (600 GB)

► Array sites (S38 and S39) that were used to create the arrays

► Diskclass (ENT for Enterprise).

12.3.3 Creating the ranks

After you create all of the required arrays, create the ranks by using the **mkrank** command. The format of the command is **mkrank -array Ax -stgtype** *xxx*, where *xxx* is fixed block (**FB**) or count key data (**CKD**), depending on whether you are configuring for open systems hosts or z Systems hosts.

After all of the ranks are created, the **lsrank** command is run. This command displays this information:

► All of the ranks that were created
► The server to which the rank is attached (attached to none, in the example up to now)
► The RAID type
► The format of the rank (fb or ckd)

Example 12-30 shows the **mkrank** command and the result of a successful **lsrank** command.

Example 12-30   Create and list ranks with mkrank and lsrank commands

```
dscli> mkrank  -array A34 -stgtype fb
CMUC00007I mkrank: Rank R25 successfully created.
dscli> lsrank
ID   Group State       datastate Array RAIDtype extpoolID stgtype
==================================================================
R0       0 Normal      Normal    A0           5 P0        fb
R9       1 Normal      Normal    A5           5 P3        ckd
R10      0 Normal      Normal    A12          5 P2        ckd
R11      1 Normal      Normal    A13          5 P7        ckd
R20      0 Normal      Normal    A25          6 P8        fb
R24      - Unassigned  Normal    A26          6 -         ckd
R25      - Unassigned  Normal    A34          5 -         fb
```

When defining a rank, you can also specify the extent size. Starting with DS8880 R8.1 firmware, you can have ranks and extent pools with large 1 GiB FB extents or small 16 MiB FB extents. The extent unit is specified with the **-extunit** parameter of the **mkrank** command. The first rank added to an extent pool determines the extent size of the extent pool.

Default RAID policy

With the new models of DS8880 (984, 985, 986 and 988), the default RAID is 6. You will get an alert message if you try to use RAID 5 as shown in Example 12-31.

Example 12-31   Alert message about using RAID 5

```
dscli> mkarray -dev IBM.2107-75DMC81 -raidtype 5 -arsite S10
CMUC00005W mkarray: IBM Service Delivery organization highly recommends the use of RAID 6 over
RAID 5 for increased reliability.  Please acknowledge you understand the risks associated with
RAID 5.
Are you sure you want to accept the disclaimer above? [y/n]: n
CMMCI9002I Command mkarray halted.
---------------------------------------------------------------------
dscli> mkarray -dev IBM.2107-75DMC81 -raidtype 5 -arsite S10 -force
CMUC00005I mkarray: You have accepted the following disclaimer:
IBM Service Delivery organization highly recommends the use of RAID 6 over RAID 5 for increased
reliability.  Please acknowledge you understand the risks associated with RAID 5.
CMUC00004I mkarray: Array A10 successfully created.
```

12.3.4 Creating the extent pools

The next step is to create the extent pools. Remember the following points when you create the extent pools:

► Each extent pool includes an associated rank group that is specified by the **-rankgrp** parameter, which defines the extent pool's server affinity (0 for server0 or 1 for server1).

► The extent pool type is FB or CKD, and is specified by the **-stgtype** parameter.

► The number of extent pools can range from one to the number of existing ranks. However, to associate ranks with both servers, you need at least two extent pools.

For easier management, create empty extent pools that relate to the type of storage that is in the pool. For example, create an extent pool for high capacity disk, create another extent pool for high performance, and, if needed, create extent pools for the CKD environment.

When an extent pool is created, the system automatically assigns it an extent pool ID, which is a decimal number that starts from 0, preceded by the letter P. The ID that was assigned to an extent pool is shown in the CMUC00000I message, which is displayed in response to a successful **mkextpool** command. Extent pools that are associated with rank group 0 receive an even ID number. Extent pools that are associated with rank group 1 receive an odd ID number. The extent pool ID is used when you refer to the extent pool in subsequent DS CLI commands. Therefore, it is good practice to make a note of the ID.

Example 12-32 shows one example of extent pools that you can define on your system. This setup requires a system with at least six ranks.

Example 12-32   An extent pool layout plan

```
FB Extent Pool high capacity 1200GB disks assigned to server 0 (FB_LOW_0)
FB Extent Pool high capacity 1200GB disks assigned to server 1 (FB_LOW_1)
FB Extent Pool high performance 300GB disks assigned to server 0 (FB_High_0)
FB Extent Pool high performance 300GB disks assigned to server 0 (FB_High_1)
CKD Extent Pool High performance 300GB disks assigned to server 0 (CKD_High_0)
CKD Extent Pool High performance 300GB disks assigned to server 1 (CKD_High_1)
```

The **mkextpool** command forces you to name the extent pools. To do so, complete these steps:

1. Create empty extent pools by using the **mkextpool** command as shown in Example 12-33.

2. List the extent pools to obtain their IDs.

3. Attach a rank to an empty extent pool by using the **chrank** command.

4. List the extent pools again by using **lsextpool** and note the change in the capacity of the extent pool.

Example 12-33   Create an extent pool by using mkextpool, lsextpool, and chrank

```
dscli> mkextpool -rankgrp 0 -stgtype fb FB_high_0
CMUC00000I mkextpool: Extent Pool P0 successfully created.
dscli> mkextpool -rankgrp 1 -stgtype fb FB_high_1
CMUC00000I mkextpool: Extent Pool P1 successfully created.
dscli> lsextpool
Name       ID stgtype rankgrp status availstor (2^30B) %allocated available reserved numvols
==============================================================================================
FB_high_0 P0 fb            0  below              0         0         0        0       0
FB_high_1 P1 fb            1  below              0         0         0        0       0
dscli> chrank -extpool P0 R0
CMUC00008I chrank: Rank R0 successfully modified.
dscli> chrank -extpool P1 R1
CMUC00008I chrank: Rank R1 successfully modified.
dscli> lsextpool
Name       ID stgtype rankgrp status availstor (2^30B) %allocated available reserved numvols
==============================================================================================
FB_high_0 P0 fb            0  below           1578         0      1578        0       0
FB_high_1 P1 fb            1  below           1578         0      1578        0       0
```

After a rank is assigned to an extent pool, you can see this change when you display the ranks.

In Example 12-34, you can see that rank R0 is assigned to extpool P0.

Example 12-34   Display the ranks after a rank is assigned to an extent pool

```
dscli> lsrank
ID   Group State  datastate Array RAIDtype extpoolID stgtype
=========================================================
R0        0 Normal Normal    A0        5 P0         ckd
R1        0 Normal Normal    A1        5 P0         ckd
R2        1 Normal Normal    A2        5 P1         ckd
R3        1 Normal Normal    A3        5 P1         ckd
R4        0 Normal Normal    A4        5 P2         fb
R5        0 Normal Normal    A5        5 P2         fb
R8        0 Normal Normal    A8        6 P4         fb
R11       1 Normal Normal    A11       6 P5         fb
R16       0 Normal Normal    A16       6 P6         fb
```

Example 12-35 shows the extent size in the query.

Example 12-35   Command showing the extent size

```
dscli> lsrank -l
ID   Group State  datastate Array RAIDtype extpoolID extpoolnam stgtype exts    usedexts encryptgrp marray extsize (cap)
=============================================================================================================================
R0    0 Normal Normal    A0        5 P0       CKD_L      ckd     2392         31          - MA25  1113cyl
R1    0 Normal Normal    A1        5 P0       CKD_L      ckd     8313        814          - MA6   1113cyl
R2    1 Normal Normal    A2        5 P1       CKD_L      ckd     2392         31          - MA24  1113cyl
R3    1 Normal Normal    A3        5 P1       CKD_L      ckd     8313        800          - MA5   1113cyl
R4    0 Normal Normal    A4        5 P2       FB_Gold_L  fb      2132        208          - MA23  1GiB
R5    0 Normal Normal    A5        5 P2       FB_Gold_L  fb      6348       1572          - MA4   1GiB
R6    1 Normal Normal    A6        5 P3       FB_Gold_L  fb      2132       1818          - MA22  1GiB
R8    0 Normal Normal    A8        6 P4       FB_Silver_L fb     17261       181          - MA21  1GiB
R9    0 Normal Normal    A9        5 P4       FB_Silver_L fb      3688       279          - MA18  1GiB
R10   0 Normal Normal    A10       5 P4       FB_Silver_L fb      3156       277          - MA8   1GiB
R11   1 Normal Normal    A11       6 P5       FB_Silver_L fb     17261        87          - MA20  1GiB
R12   1 Normal Normal    A12       5 P5       FB_Silver_L fb      3688        11          - MA17  1GiB
R13   1 Normal Normal    A13       5 P5       FB_Silver_L fb      3156        13          - MA7   1GiB
R14   0 Normal Normal    A14       5 P6       FB_Bronze_S fb    236053     19498          - MA13  16MiB
R15   1 Normal Normal    A15       5 P7       FB_Bronze_S fb    236053     51278          - MA16  16MiB
R16   0 Normal Normal    A16       6 P6       FB_Bronze_S fb   1104704    103832          - MA19  16MiB
R17   0 Normal Normal    A17       5 P8       CKD_S      ckd   219245     23672          - MA9   21cyl
R18   0 Normal Normal    A18       5 P8       CKD_S      ckd   377388     23672          - MA2   21cyl
R19   1 Normal Normal    A19       5 P9       CKD_S      ckd   219245     23664          - MA10  21cyl
R20   1 Normal Normal    A20       5 P9       CKD_S      ckd   377388     23664          - MA1   21cyl
R21   0 Normal Normal    A21       5 P6       FB_Bronze_S fb    236053     19471          - MA12  16MiB
R22   1 Normal Normal    A22       5 P7       FB_Bronze_S fb    236053     51274          - MA14  16MiB
R29   0 Normal Normal    A23      10 P10      Fabian_pool fb      2093         0          - MA11  1GiB
```

12.3.5 Creating the FB volumes

Now, you can create volumes and volume groups. When you create the volumes or groups, try to distribute them evenly across the two rank groups in the storage unit.

Although an FB-type volume can be created as standard (thick) and thin (extent space efficient (ESE))-type volumes, this section describes the creation of the standard type only.

Creating the standard volumes

Use the following command format when creating a volume:

mkfbvol -extpool pX -cap xx -name high_fb_0#h XXXX-XXXX

The last parameter is the *volume_ID*, which can be a range or single entry. The four-digit entry is based on LL and VV. LL (00 - FE) equals the logical subsystem (LSS) that the volume belongs to, and VV (00 - FF) equals the volume number on the LSS. Therefore, the DS8880 can support 255 LSSs, and each LSS can support a maximum of 256 volumes.

Example 12-36 shows the creation of eight volumes, each with a capacity of 10 GiB. The first four volumes are assigned to rank group 0, and are assigned to LSS 10 with volume numbers of 00 - 03. The second four volumes are assigned to rank group 1, and are assigned to LSS 11 with volume numbers of 00 - 03.

Example 12-36   Create fixed-block volumes by using mkfbvol

```
dscli> lsextpool
Name       ID stgtype rankgrp status availstor (2^30B) %allocated available reserved numvols
=============================================================================================
FB_high_0 P0 fb            0  below             1578            0      1578        0        0
FB_high_1 P1 fb            1  below             1578            0      1578        0        0
dscli> mkfbvol -extpool p0 -cap 10 -name high_fb_0_#h 1000-1003
CMUC00025I mkfbvol: FB volume 1000 successfully created.
CMUC00025I mkfbvol: FB volume 1001 successfully created.
CMUC00025I mkfbvol: FB volume 1002 successfully created.
```

```
CMUC00025I mkfbvol: FB volume 1003 successfully created.
dscli> mkfbvol -extpool p1 -cap 10 -name high_fb_1_#h 1100-1103
CMUC00025I mkfbvol: FB volume 1100 successfully created.
CMUC00025I mkfbvol: FB volume 1101 successfully created.
CMUC00025I mkfbvol: FB volume 1102 successfully created.
CMUC00025I mkfbvol: FB volume 1103 successfully created.
```

Looking closely at the **mkfbvol** command that is used in Example 12-36, you see that volumes 1000 - 1003 are in extpool P0. That extent pool is attached to rank group 0, which means server 0. Rank group 0 can contain only even-numbered LSSs, which means that volumes in that extent pool must belong to an even-numbered LSS. The first 2 digits of the volume serial number are the LSS number. So, in this case, volumes 1000 - 1003 are in LSS 10.

For volumes 1100 - 1003 in Example 12-36, the first 2 digits of the volume serial number are 11 (an odd number), which signifies that they belong to rank group 1. The **-cap** parameter determines the size. However, because the **-type** parameter was not used, the default size is a binary size. Therefore, these volumes are 10 GB binary, which equates to 10,737,418,240 bytes. If you used the **-type ess** parameter, the volumes are decimally sized and they are a minimum of 10,000,000,000 bytes in size.

Example 12-36 on page 373 named the volumes by using the naming scheme high_fb_0_#h, where #h means that you are using the hexadecimal volume number as part of the volume name. This naming convention can be seen in Example 12-37, where you list the volumes that you created by using the **lsfbvol** command. You then list the extent pools to see how much space is left after the volume is created.

Example 12-37   Check the machine after the volumes are created by using lsfbvol and lsextpool

```
dscli> lsfbvol
Name          ID   accstate datastate configstate deviceMTM datatype extpool cap (2^30B)
==========================================================================================
high_fb_0_1000 1000 Online   Normal    Normal      2107-900  FB 512   P0            10.0
high_fb_0_1001 1001 Online   Normal    Normal      2107-900  FB 512   P0            10.0
high_fb_0_1002 1002 Online   Normal    Normal      2107-900  FB 512   P0            10.0
high_fb_0_1003 1003 Online   Normal    Normal      2107-900  FB 512   P0            10.0
high_fb_1_1100 1100 Online   Normal    Normal      2107-900  FB 512   P1            10.0
high_fb_1_1101 1101 Online   Normal    Normal      2107-900  FB 512   P1            10.0
high_fb_1_1102 1102 Online   Normal    Normal      2107-900  FB 512   P1            10.0
high_fb_1_1103 1103 Online   Normal    Normal      2107-900  FB 512   P1            10.0
dscli> lsextpool
Name      ID stgtype rankgrp status availstor (2^30B) %allocated available reserved numvols
==========================================================================================
FB_high_0 P0 fb           0  below             1538          2      1538        0       4
FB_high_1 P1 fb           1  below             1538          2      1538        0       4
```

Important considerations:

► For the DS8880, the LSSs can be ID 00 to ID FE. The LSSs are in address groups. Address group 0 is LSS 00 to 0F, address group 1 is LSS 10 to 1F, and so on, except group F, which is F0 - FE. When you create an FB volume in an address group, that entire address group can be used only for FB volumes. Be aware of this fact when you plan your volume layout in a mixed FB and CKD DS8880. The LSS is automatically created when the first volume is assigned to it.

► You can configure a volume to belong to a certain Performance I/O Priority Manager by using the `-perfgrp` *`<perf_group_ID>`* flag in the `mkfbvol` command. For more information, see *DS8000 I/O Priority Manager*, REDP-4760.

Resource group: You can configure a volume to belong to a certain resource group by using the `-resgrp` *`<RG_ID>`* flag in the `mkfbvol` command. For more information, see *IBM System Storage DS8000 Copy Services Scope Management and Resource Groups*, REDP-4758.

T10 Data Integrity Field volumes

A standard for end-to-end error checking from the application to the disk drives is emerging that is called SCSI T10 Data Integrity Field (DIF). T10 DIF requires volumes to be formatted in 520-byte sectors with cyclic redundancy check (CRC) bytes that are added to the data. If you want to use this technique, you must create volumes that are formatted for T10 DIF usage. You configure T10 DIF by adding the `-t10dif` parameter to the `mkfbvol` command. It is possible to create T10 DIF volumes and use them as standard volumes, and then enable them later without configuration changes. For more information, see "T10 Data Integrity Field volumes" on page 375.

Storage pool striping

When a volume is created, you can choose how the volume is allocated in an extent pool with several ranks. The extents of a volume can be kept together in one rank (if enough free space exists on that rank). The next rank is used when the next volume is created. This allocation method is called *rotate volumes*.

You can also specify that you want the extents of the volume that you create to be evenly distributed across all ranks within the extent pool. This allocation method is called *rotate extents*. The storage pool striping spreads the I/O of a logical unit number (LUN) to multiple ranks, which improves performance and greatly reduces hot spots.

The extent allocation method is specified with the `-eam rotateexts` or `-eam rotatevols` option of the `mkfbvol` command, as shown in Example 12-38.

Default allocation policy: For DS8880, the default allocation policy is rotate extents.

Example 12-38   Create a volume with storage pool striping

```
dscli> mkfbvol -extpool p7 -cap 15 -name ITSO-XPSTR -eam rotateexts 1720
CMUC00025I mkfbvol: FB volume 1720 successfully created.
```

The **showfbvol** command with the **-rank** option (Example 12-39) shows that the volume that you created is distributed across three ranks. It also shows how many extents on each rank were allocated for this volume.

Example 12-39   Use showfbvol for information about a striped volume

```
dscli> showfbvol -rank f000
Name                AKB_F000
ID                  F000
accstate            Online
datastate           Normal
configstate         Normal
deviceMTM           2107-900
datatype            FB 512
addrgrp             F
extpool             P6
exts                320
cap (MiB)           5120
captype             DS
cap (2^30B)         5.0
cap (10^9B)         -
cap (blocks)        10485760
volgrp              -
ranks               2
dbexts              0
sam                 Standard
repcapalloc         -
eam                 managed
reqcap (blocks)     10485760
realextents         320
virtualextents      2
realcap (MiB)       5120
migrating           0
migratingcap (MiB)  0
perfgrp             PG0
migratingfrom       -
resgrp              RG0
tierassignstatus    -
tierassignerror     -
tierassignorder     -
tierassigntarget    -
%tierassigned       0
etmonpauseremain    -
etmonitorreset      unknown
GUID                6005076303FFD13E000000000000F000
==============Rank extents==============
rank extents capacity (MiB/cyl)
===============================
R14      160 2560
R21      160 2560
```

Dynamic volume expansion

A volume can be expanded without removing the data within the volume. You can specify a new capacity by using the **chfbvol** command, as shown in Example 12-40.

Example 12-40   Expand a striped volume

```
dscli> chfbvol -cap 6 f000
CMUC00332W chfbvol: Some host operating systems do not support changing the volume
size. Are you sure that you want to resize the volume? [y/n]: y
CMUC00026I chfbvol: FB volume F000 successfully modified.
```

The largest LUN size is now 16 TiB. Copy Services are not supported for LUN sizes larger than 2 TiB.

> **New capacity:** The new capacity must be larger than the previous capacity. You cannot shrink the volume.

Because the original volume included the rotateexts attribute, the other extents are also striped, as shown in Example 12-41. See (1) in both examples and check the difference.

Example 12-41   Check the status of an expanded volume

```
ddscli> showfbvol -rank f000
Name               AKB_F000
ID                 F000
accstate           Online
datastate          Normal
configstate        Normal
deviceMTM          2107-900
datatype           FB 512
addrgrp            F
extpool            P6
exts               384
cap (MiB)          6144
captype            DS
cap (2^30B)        6.0
cap (10^9B)        -
cap (blocks)       12582912
volgrp             -
ranks              2
dbexts             0
sam                Standard
repcapalloc        -
eam                managed
reqcap (blocks)    12582912
realextents        384
virtualextents     2
realcap (MiB)      6144
migrating          0
migratingcap (MiB) 0
perfgrp            PG0
migratingfrom      -
resgrp             RG0
tierassignstatus   -
tierassignerror    -
```

```
tierassignorder    -
tierassigntarget   -
%tierassigned      0
etmonpauseremain   -
etmonitorreset     unknown
GUID               6005076303FFD13E000000000000F000
==============Rank extents==============
rank extents capacity (MiB/cyl)
===============================
R14      192 3072
R21      192 3072
```

> **Important:** Before you can expand a volume, you must delete all Copy Services relationships for that volume.

Defining thin-provisioned ESE volumes

The DS8880 supports thin-provisioned volumes. A thin-provisioned volume is created by the DS CLI specifying the storage allocation method (SAM) of ESE for a volume. This specification is done with the **-sam ese** option of the **mkfbvol** command as shown in Example 12-42.

Example 12-42   Creating a thin-provisioned ESE volume

```
dscli> mkfbvol -extpool p4 -cap 1000 -sam ese 2000
CMUC00025I mkfbvol: FB volume 2000 successfully created.
```

The extent size is defined by the extent pools where the volume is created.

More DS CLI commands are available to control and protect the space in an extent pool for thin-provisioned volumes. One of these commands is the **mksestg** command to reserve space for thin-provisioned volumes. For more information about thin-provisioning, see *IBM DS8880 Thin Provisioning*, REDP-5343.

Deleting volumes

FB volumes can be deleted by using the **rmfbvol** command.

The command includes options to prevent the accidental deletion of volumes that are in use. An FB volume is considered to be in use if it is participating in a Copy Services relationship or if the volume received any I/O operation in the previous five minutes.

Volume deletion is controlled by the **-safe** and **-force** parameters (they cannot be specified at the same time) in the following manner:

► If none of the parameters are specified, the system performs checks to see whether the specified volumes are in use. Volumes that are not in use are deleted and the volumes that are in use are not deleted.

► If the **-safe** parameter is specified and if any of the specified volumes are assigned to a user-defined volume group, the command fails without deleting any volumes.

► The **-force** parameter deletes the specified volumes without checking to see whether they are in use.

Example 12-43 shows the creation of volumes 2100 and 2101, then the assignment of volume 2100 to a volume group. You then try to delete both volumes with the **-safe** option, but the attempt fails without deleting either of the volumes. You can delete volume 2101 with the **-safe** option because it is not assigned to a volume group. Volume 2100 is not in use, so you can delete it by not specifying either parameter.

Example 12-43   Delete an FB volume

```
dscli> mkfbvol -extpool p1 -cap 12 -eam rotateexts 2100-2101
CMUC00025I mkfbvol: FB volume 2100 successfully created.
CMUC00025I mkfbvol: FB volume 2101 successfully created.
dscli> chvolgrp -action add -volume 2100 v0
CMUC00031I chvolgrp: Volume group V0 successfully modified.
dscli> rmfbvol -quiet -safe 2100-2101
CMUC00253E rmfbvol: Volume IBM.2107-75NA901/2100 is assigned to a user-defined
volume group. No volumes were deleted.
dscli> rmfbvol -quiet -safe 2101
CMUC00028I rmfbvol: FB volume 2101 successfully deleted.
dscli> rmfbvol 2100
CMUC00027W rmfbvol: Are you sure you want to delete FB volume 2100? [y/n]: y
CMUC00028I rmfbvol: FB volume 2100 successfully deleted.
```

12.3.6 Creating the volume groups

FB volumes are assigned to open system hosts by using volume groups. Do not confuse them with the *volume groups* term, which is used in AIX. A fixed-block volume can be a member of multiple volume groups. Volumes can be added or removed from volume groups as required. Each volume group must be SCSI MAP256 or SCSI MASK, depending on the SCSI LUN address discovery method that is used by the operating system to which the volume group is attached.

Determining whether an open system host is SCSI MAP256 or SCSI MASK

First, determine the type of SCSI host with which you are working. Then, use the **lshosttype** command with the **-type** parameter of scsimask and then scsimap256.

Example 12-44 shows the results of each command.

Example 12-44   List host types with the lshosttype command

```
dscli> lshosttype -type scsimask
HostType        Profile                                   AddrDiscovery LBS
==========================================================================
Hp              HP - HP/UX                                reportLUN     512
SVC             San Volume Controller                     reportLUN     512
SanFsAIX        IBM pSeries - AIX/SanFS                   reportLUN     512
pSeries         IBM pSeries - AIX                         reportLUN     512
pSeriesPowerswap IBM pSeries - AIX with Powerswap support reportLUN     512
zLinux          IBM zSeries - zLinux                      reportLUN     512
dscli> lshosttype -type scsimap256
HostType     Profile             AddrDiscovery LBS
===================================================
AMDLinuxRHEL AMD - Linux RHEL    LUNPolling    512
AMDLinuxSuse AMD - Linux Suse    LUNPolling    512
AppleOSX     Apple - OSX         LUNPolling    512
```

```
Fujitsu      Fujitsu - Solaris    LUNPolling   512
HpTru64      HP - Tru64           LUNPolling   512
HpVms        HP - Open VMS        LUNPolling   512
LinuxDT      Intel - Linux Desktop LUNPolling  512
LinuxRF      Intel - Linux Red Flag LUNPolling 512
LinuxRHEL    Intel - Linux RHEL   LUNPolling   512
LinuxSuse    Intel - Linux Suse   LUNPolling   512
Novell       Novell               LUNPolling   512
SGI          SGI - IRIX           LUNPolling   512
SanFsLinux   - Linux/SanFS        LUNPolling   512
Sun          SUN - Solaris        LUNPolling   512
VMWare       VMWare               LUNPolling   512
Win2000      Intel - Windows 2000 LUNPolling   512
Win2003      Intel - Windows 2003 LUNPolling   512
Win2008      Intel - Windows 2008 LUNPolling   512
Win2012      Intel - Windows 2012 LUNPolling   512
iLinux       IBM iSeries - iLinux LUNPolling   512
nSeries      IBM N series Gateway LUNPolling   512
pLinux       IBM pSeries - pLinux LUNPolling   512
```

Creating a volume group

After you determine the host type, create a volume group. In Example 12-45, the example host type is AIX. In Example 12-44 on page 379, you can see that the address discovery method for AIX is scsimask.

Example 12-45   Create a volume group with mkvolgrp and display it by using lsvolgrp

```
dscli> mkvolgrp -type scsimask -volume 1000-1002,1100-1102 AIX_VG_01
CMUC00030I mkvolgrp: Volume group V11 successfully created.
dscli> lsvolgrp
Name                 ID  Type
=====================================
ALL CKD              V10 FICON/ESCON All
AIX_VG_01            V11 SCSI Mask
ALL Fixed Block-512  V20 SCSI All
ALL Fixed Block-520  V30 OS400 All
dscli> showvolgrp V11
Name AIX_VG_01
ID   V11
Type SCSI Mask
Vols 1000 1001 1002 1100 1101 1102
```

Adding or deleting volumes in a volume group

In this example, you add volumes 1000 - 1002 and 1100 - 1102 to the new volume group to evenly spread the workload across the two rank groups. You then list all available volume groups by using the `lsvolgrp` command. Finally, list the contents of volume group V11 because you created this volume group.

You might also want to add or remove volumes to this volume group later. To add or remove volumes, use the `chvolgrp` command with the `-action` parameter.

Example 12-46 shows adding volume 1003 to volume group V11, displaying the results, and then removing the volume.

Example 12-46   Change a volume group with chvolgrp

```
dscli> chvolgrp -action add -volume 1003 V11
CMUC00031I chvolgrp: Volume group V11 successfully modified.
dscli> showvolgrp V11
Name AIX_VG_01
ID   V11
Type SCSI Mask
Vols 1000 1001 1002 1003 1100 1101 1102
dscli> chvolgrp -action remove -volume 1003 V11
CMUC00031I chvolgrp: Volume group V11 successfully modified.
dscli> showvolgrp V11
Name AIX_VG_01
ID   V11
Type SCSI Mask
Vols 1000 1001 1002 1100 1101 1102
```

> **Important:** Not all operating systems can manage the removal of a volume. See your operating system documentation to determine the safest way to remove a volume from a host.

12.3.7 Creating host connections

The final step in the logical configuration process is to create host connections for your attached hosts. You must assign volume groups to those connections. Each host's host bus adapter (HBA) can be defined only once. Each host connection (hostconnect) can only have one volume group assigned to it. A volume can be assigned to multiple volume groups.

Starting with Release 8.2, the **mkhostconnect** command returns an error if the "host_name" on the command does not match an existing one. You must define it first using the new command **mkhost** as shown in Example 12-47. Four new commands were also added: **chhost**, **rmhost**, **lshost**, and **showhost**.

Example 12-47   Create host name to a new connection

```
dscli> mkhost -type generic -hostport 210100E08B295F86 test_1
```

Example 12-48 shows the creation of a single host connection that represents one HBA in the example AIX host. Use the **-hosttype** parameter to include the host type that you used in Example 12-44 on page 379. Allocate it to volume group V11. If the SAN zoning is correct, the host will be able to see the LUNs in volume group V11.

Example 12-48   Create host connections by using mkhostconnect and lshostconnect

```
dscli> mkhostconnect -wwname 100000C912345678 -hosttype pSeries -volgrp V11 AIX_Server_01
CMUC00012I mkhostconnect: Host connection 0000 successfully created.
dscli> lshostconnect
Name          ID   WWPN             HostType Profile           portgrp volgrpID ESSIOport
========================================================================================
AIX_Server_01 0000 100000C912345678 pSeries  IBM pSeries - AIX       0 V11      all
```

You can also use **-profile** instead of **-hosttype**. However, this method is not a preferred practice. The use of the **-hosttype** parameter starts both parameters (**-profile** and **-hosttype**). In contrast, the use of **-profile** leaves the **-hosttype** column unpopulated.

The option in the `mkhostconnect` command to restrict access only to certain I/O ports is also available by using the **-ioport** parameter. Restricting access in this way is unnecessary. If you want to restrict access for certain hosts to certain I/O ports on the DS8880, perform zoning on your SAN switch.

Managing hosts with multiple HBAs

If your host features multiple HBAs, consider the following points:

► For the DS GUI to consider multiple host connects to be used by the same server, the host connects must have the same name. In Example 12-49, host connects 0010 and 0011 appear in the DS GUI as a single server with two HBAs. However, host connects 000E and 000F appear as two separate hosts even though they are used by the same server. If you do not plan to use the DS GUI to manage host connections, this consideration is not important. The use of more verbose host connection (hostconnect) naming might make management easier.

► If you want to use a single command to change the assigned volume group of several host connections (hostconnects) at the same time, assign the hostconnects to a unique port group. Then use the `managehostconnect` command. This command changes the assigned volume group for all hostconnects that are assigned to a particular port group.

When hosts are created, you can specify the **-portgrp** parameter. By using a unique port group number for each attached server, you can detect servers with multiple HBAs.

Example 12-49 shows six host connections. By using the port group number, you see that three separate hosts exist, each with two HBAs. Port group 0 is used for all hosts that do not have a port group number set.

Example 12-49   Use the portgrp number to separate attached hosts

```
dscli> lshostconnect
Name            ID   WWPN              HostType  Profile                 portgrp volgrpID
===============================================================================================
Name ID   WWPN            HostType Profile              portgrp volgrpID ESSIOport
=================================================================================
-    0001 21000024FF41D1CE Win2012  Intel - Windows 2012      1 V1        all
-    0002 21000024FF41D1CF Win2012  Intel - Windows 2012      1 V1        all
-    0004 10000000C970113A pSeries  IBM pSeries - AIX         2 V2        all
-    0005 10000000C970113B pSeries  IBM pSeries - AIX         2 V2        all
-    0007 10000090FA1F7A82 pSeries  IBM pSeries - AIX         3 V3        all
-    0008 10000090FA1F7A83 pSeries  IBM pSeries - AIX         3 V3        all
-    000A 2100000E1E30B962 VMWare   VMWare                    4 V4        all
-    000B 2100000E1E30B963 VMWare   VMWare                    4 V4        all
-    000D 10000090FA02F560 iSeries  IBM iSeries - OS/400      5 V5        all
-    0010 10000000C987323D iSeries  IBM iSeries - OS/400      6 V6        all
-    0011 10000000C98C8DB2 iSeries  IBM iSeries - OS/400      6 V6        all
-    0013 21000024FF41CBE2 VMWare   VMWare                    7 V7        all
```

Changing host connections

If you want to change a host connection, use the `chhostconnect` command. This command can be used to change nearly all parameters of the host connection, except for the worldwide port name (WWPN).

If you must change the WWPN, first create a host connection. To change the assigned volume group, use the **chhostconnect** command to change one host connection (hostconnect) at a time, or use the **managehostconnect** command to simultaneously reassign all of the hostconnects in one port group.

12.3.8 Mapping open system host disks to storage unit volumes

When you assign volumes to an open system host and install the DS CLI on a host, you can run the **lshostvol** DS CLI command on that host. This command maps assigned LUNs to open system host volume names.

This section provides examples for several operating systems. Each example assigns several logical volumes to an open system host and then installs DS CLI on that host. Then, you log on to this host and start DS CLI. It does not matter which HMC you connect to with the DS CLI. Then, issue the **lshostvol** command.

> **Important:** The **lshostvol** command communicates only with the operating system of the host on which the DS CLI is installed. You cannot run this command on one host to see the attached disks of another host.

AIX: Mapping disks when Multipath I/O is used

Example 12-50 shows an AIX server that uses Multipath I/O (MPIO). Two volumes, 1800 and 1801, are assigned to this host. Because MPIO is used, you do not see the number of paths.

In fact, from this display, it is not possible to tell whether MPIO is even installed. You must run the **pcmpath query device** command to confirm the path count.

Example 12-50   Running the lshostvol command on an AIX host by using MPIO

```
dscli> lshostvol
Disk Name Volume Id          Vpath Name
==========================================
hdisk3    IBM.2107-1300819/1800 ---
hdisk4    IBM.2107-1300819/1801 ---
```

> **Open HyperSwap:** If you use Open HyperSwap on a host, the **lshostvol** command might fail to show any devices.

AIX: Mapping disks when Subsystem Device Driver is used

Example 12-51 shows an AIX server that uses Subsystem Device Driver (SDD). Two volumes, 1000 and 1100, are assigned to this host. Each volume has four paths.

Example 12-51   Running the lshostvol command on an AIX host by using SDD

```
dscli> lshostvol
Disk Name                   Volume Id          Vpath Name
============================================================
hdisk1,hdisk3,hdisk5,hdisk7 IBM.2107-1300247/1000 vpath0
hdisk2,hdisk4,hdisk6,hdisk8 IBM.2107-1300247/1100 vpath1
```

Windows: Mapping disks

Example 12-52 shows the use of the **datapath query device** command on a Microsoft Windows host with IBM Subsystem Device Driver Device Specific Module (SDDDSM) multi-pathing software installed. The disks are listed by Windows Disk number. If you want to know which disk is associated with which drive letter, you must look at the Windows Disk Manager.

Example 12-52   IBM SDDDSM datapath query device that is run on a Windows host

```
C:\Program Files\IBM\SDDDSM>datapath query device

Total Devices : 3


DEV#:   0  DEVICE NAME: Disk2 Part0           TYPE: 2107900    POLICY: LEAST I/O AND WEIGHT
SERIAL: 75ZA5714670    Reserved: No    LUN SIZE: 10.0GB
HOST INTERFACE: FC
==========================================================================================
Path#     Adapter/Hard Disk                State       Mode      Select      Errors
   0      Scsi Port5 Bus0/Disk2 Part0       OPEN       NORMAL    7671581          0
   1      Scsi Port6 Bus0/Disk2 Part0       OPEN       NORMAL    4809566          0

DEV#:   1  DEVICE NAME: Disk3 Part0          TYPE: 2107900    POLICY: LEAST I/O AND WEIGHT
SERIAL: 75ZA57146A0    Reserved: No    LUN SIZE: 32.0GB
HOST INTERFACE: FC
==========================================================================================
Path#     Adapter/Hard Disk                State       Mode      Select      Errors
   0      Scsi Port5 Bus0/Disk3 Part0       OPEN       NORMAL    8289379          0
   1      Scsi Port6 Bus0/Disk3 Part0       OPEN       NORMAL    5367333          0

DEV#:   2  DEVICE NAME: Disk4 Part0          TYPE: 2107900    POLICY: LEAST I/O AND WEIGHT
SERIAL: 75ZA57147A0    Reserved: No    LUN SIZE: 32.0GB
HOST INTERFACE: FC
==========================================================================================
Path#     Adapter/Hard Disk                State       Mode      Select      Errors
   0      Scsi Port5 Bus0/Disk4 Part0       OPEN       NORMAL    18202701         0
   1      Scsi Port6 Bus0/Disk4 Part0       OPEN       NORMAL    12000914         0

DEV#:   3  DEVICE NAME: Disk5 Part0          TYPE: 2107900    POLICY: LEAST I/O AND WEIGHT
SERIAL: 75ZA5714671    Reserved: No    LUN SIZE: 50.0GB
HOST INTERFACE: FC
==========================================================================================
Path#     Adapter/Hard Disk                State       Mode      Select      Errors
   0      Scsi Port5 Bus0/Disk5 Part0       OPEN       NORMAL    7068919          0
   1      Scsi Port6 Bus0/Disk5 Part0       OPEN       NORMAL    4188012          0
```

12.4 DS8880 storage configuration for the CKD volumes

This list contains the steps to configure CKD storage in the DS8880:

1. Create the arrays.
2. Create the CKD ranks.
3. Create the CKD extent pools.
4. Create the LCUs.
5. Create the CKD volumes.

You do not need to create volume groups or host connects for CKD volumes. If I/O ports in FICON mode exist, access to CKD volumes by FICON hosts is granted automatically, following the specifications in the input/output definition file (IODF).

12.4.1 Creating the arrays

Array creation for CKD volumes is the same as for FB volumes. For more information, see 12.3.2, "Creating the arrays" on page 369.

12.4.2 Creating the ranks and extent pools

When ranks and extent pools are created, you must specify `-stgtype ckd`. Then you can create the extent pool as shown in Example 12-53.

Example 12-53   Rank and extent pool creation for CKD volumes

```
dscli> mkrank -array A0 -stgtype ckd
CMUC00007I mkrank: Rank R0 successfully created.
dscli> lsrank
ID Group State      datastate Array RAIDtype extpoolID stgtype
==============================================================
R0    - Unassigned Normal    A0             6 -       ckd
dscli> mkextpool -rankgrp 0 -stgtype ckd CKD_High_0
CMUC00000I mkextpool: Extent Pool P0 successfully created.
dscli> chrank -extpool P0 R0
CMUC00008I chrank: Rank R0 successfully modified.
dscli> lsextpool
Name       ID stgtype rankgrp status availstor (2^30B) %allocated available reserved numvol
==========================================================================================
CKD_High_0 0 ckd          0  below              252           0       287        0      0
```

When defining a rank, you can also specify the extent size. Starting with DS8880 R8.1 firmware, you can have ranks and extent pools with large 1113 cylinder CKD extents, or small 21 cylinder CKD extents. The extent unit is specified with the `-extsize` parameter of the `mkrank` command. The first rank added to an extent pool determines the extent size of the extent pool. Example 12-54 shows ranks with different extent sizes.

Example 12-54   CKD ranks with different extent sizes

```
dscli> lsrank -l
...
ID  Group State   datastate Array RAIDtype extpoolID extpoolnam stgtype exts    usedexts encryptgrp marray extsize (cap)
=======================================================================================================================
R4   0 Normal  Normal   A4       5 P2        Small      ckd     187636   9904          -   MA19  21cyl
R5   0 Normal  Normal   A5       5 P2        Small      ckd     219245   9919          -   MA24  21cyl
R9   0 Normal  Normal   A9       5 P0        CKD_       ckd       4113   2406          -   MA6   1113cyl
R10  0 Normal  Normal   A10      5 P0        CKD_       ckd       3520   2399          -   MA3   1113cyl
```

12.4.3 Logical control unit creation

When volumes for a CKD environment are created, you must create logical control units (LCUs) before the volumes are created. In Example 12-55, you can see what happens if you try to create a CKD volume without creating an LCU first.

Example 12-55   Trying to create CKD volumes without creating an LCU first

```
dscli> mkckdvol -extpool p2 -cap 262668 -name ITSO_EAV1_#h C200
CMUN02282E mkckdvol: C200: Unable to create CKD logical volume: CKD volumes require a CKD
logical subsystem.
```

First, use the `mklcu` command. The command uses the following format:

```
mklcu -qty XX -id XX -ss XXXX
```

To display the LCUs that you created, use the `lslcu` command.

Example 12-56 shows the creation of two LCUs by using the `mklcu` command and then listing the created LCUs by using the `lslcu` command. By default, the LCUs that were created are the 3990-6 type.

Example 12-56   Create a logical control unit with mklcu

```
dscli> mklcu -qty 2 -id BC -ss BC00
CMUC00017I mklcu: LCU BC successfully created.
CMUC00017I mklcu: LCU BD successfully created.
dscli> lslcu
ID Group addrgrp confgvols subsys conbasetype
============================================
BC    0 C            0 0xBC00 3990-6
BD    1 C            0 0xBC01 3990-6
```

Because two LCUs were created (by using the parameter `-qty 2`), the first LCU, which is ID BC (an even number), is in address group 0, which equates to rank group 0. The second LCU, which is ID BD (an odd number), is in address group 1, which equates to rank group 1. By placing the LCUs into both address groups, performance is maximized by spreading the workload across both servers in the DS8880.

> **Important:** For the DS8880, the CKD LCUs can be ID 00 to ID FE. The LCUs fit into one of 16 address groups. Address group 0 is LCUs 00 to 0F, address group 1 is LCUs 10 to 1F, and so on, except group F is F0 - FE. If you create a CKD LCU in an address group, that address group cannot be used for FB volumes. Likewise, if, for example, FB volumes were in LSS 40 to 4F (address group 4), that address group cannot be used for CKD. Be aware of this limitation when you plan the volume layout in a mixed FB/CKD DS8880. Each LCU can manage a maximum of 256 volumes, including alias volumes for the parallel access volume (PAV) feature.

12.4.4 Creating the CKD volumes

Now that an LCU was created, the CKD volumes can be created by using the `mkckdvol` command. The `mkckdvol` command uses the following format:

```
mkckdvol -extpool P2 -cap 262668 -datatype 3390-A -eam rotatevols -name
ITSO_EAV1_#h BC06
```

The greatest difference with CKD volumes is that the capacity is expressed in cylinders or as mod1 (Model 1) extents (1113 cylinders). To not waste space, use volume capacities that are a multiple of 1113 cylinders. The support for extended address volumes (EAVs) was enhanced. The DS8880 supports EAV volumes up to 1,182,006 cylinders. The EAV device type is called *3390 Model A*. You need z/OS V1.12 or later to use these volumes.

> **Important:** For 3390-A volumes, the size can be specified as 1 - 65,520 in increments of 1, and from 65,667, which is the next multiple of 1113, to 1,182,006 in increments of 1113.

The last parameter in the command is the `volume_ID`. This value determines the LCU that the volume belongs to and the unit address for the volume. Both of these values must be matched to a control unit and device definition in the input/output configuration data set (IOCDS) that a z Systems server uses to access the volume.

The volume_ID has a format of LLVV. LL (00 - FE) equals the LCU that the volume belongs to, and VV (00 - FF) equals the offset for the volume. Only one volume of an LCU can use a unique VV of 00 - FF.

Example 12-57 creates a single 3390-A volume with a capacity of 262,668 cylinders, and it is assigned to LCU BC with an offset of 06.

Example 12-57   Create CKD volumes by using mkckdvol

```
dscli> mkckdvol -extpool P2 -cap 262668 -datatype 3390-A -eam rotatevols -name ITSO_EAV1_#h BC06
CMUC00021I mkckdvol: CKD Volume BC06 successfully created.
dscli> lsckdvol
Name           ID    accstate datastate configstate deviceMTM voltype   orgbvols extpool cap (cyl)
===================================================================================================
ITSO_BC00      BC00  Online   Normal    Normal      3390-9    CKD Base  -        P2          10017
ITSO_BC01      BC01  Online   Normal    Normal      3390-9    CKD Base  -        P2          10017
ITSO_BC02      BC02  Online   Normal    Normal      3390-9    CKD Base  -        P2          10017
ITSO_BC03      BC03  Online   Normal    Normal      3390-9    CKD Base  -        P2          10017
ITSO_BC04      BC04  Online   Normal    Normal      3390-9    CKD Base  -        P2          10017
ITSO_BC05      BC05  Online   Normal    Normal      3390-9    CKD Base  -        P2          10017
ITSO_EAV1_BC06 BC06  Online   Normal    Normal      3390-A    CKD Base  -        P2         262668
ITSO_BD00      BD00  Online   Normal    Normal      3390-9    CKD Base  -        P3          10017
ITSO_BD01      BD01  Online   Normal    Normal      3390-9    CKD Base  -        P3          10017
ITSO_BD02      BD02  Online   Normal    Normal      3390-9    CKD Base  -        P3          10017
ITSO_BD03      BD03  Online   Normal    Normal      3390-9    CKD Base  -        P3          10017
ITSO_BD04      BD04  Online   Normal    Normal      3390-9    CKD Base  -        P3          10017
ITSO_BD05      BD05  Online   Normal    Normal      3390-9    CKD Base  -        P3          10017
```

Remember, you can create only CKD volumes in LCUs that you already created.

Be aware that volumes in even-numbered LCUs must be created from an extent pool that belongs to rank group 0. Volumes in odd-numbered LCUs must be created from an extent pool in rank group 1.

> **Important:** You can configure a volume to belong to a certain resource group by using the `-resgrp <RG_ID>` flag in the `mkckdvol` command. For more information, see *IBM System Storage DS8000 Copy Services Scope Management and Resource Groups*, REDP-4758.

Defining thin-provisioned extent space efficient CKD volumes

The DS8880 supports thin-provisioned volumes. A thin-provisioned volume is created by the DS CLI by specifying the SAM of ESE for a volume. This process is done with the `-sam ese` option of the `mkckdvol` command as shown in Example 12-58.

Example 12-58   Creating a thin-provisioned CKD volume

```
dscli> mkckdvol -extpool p2 -cap 1000 -sam ese 1010
CMUC00021I mkckdvol: CKD Volume 1010 successfully created.
```

The extent size is defined by the extent pools where the volume is created.

Space release of thin-provisioned ESE CKD volumes

With Release 8.2, it is possible to release space at an extent level for CKD volumes. This feature is enabled by the z/OS utility DFSMSdss, which performs the release operation. DFSMSdss includes a new parameter, SPACERel, with its options. A new IBM RACF® FACILITY Class profile provides the protection as well. It is available on z/OS V2.1 and z/OS V2.2 with PTF OA50675.

More DS CLI commands are available to control and protect the space in an extent pool for thin-provisioned volumes. One of these commands is the `mksestg` command to reserve space for thin-provisioned volumes. For more information about thin-provisioning, see *IBM DS8880 Thin Provisioning*, REDP-5343.

Storage pool striping

When a volume is created, you can choose how the volume is allocated in an extent pool with several ranks. The extents of a volume can be kept together in one rank (if enough free space exists on that rank). The next rank is used when the next volume is created. This allocation method is called *rotate volumes*.

You can also specify that you want the extents of the volume to be evenly distributed across all ranks within the extent pool. This allocation method is called *rotate extents*.

> **Rotate extents:** For the DS8880, the default allocation policy is rotate extents.

The extent allocation method is specified with the `-eam rotateexts` or the `-eam rotatevols` option of the `mkckdvol` command (Example 12-59).

Example 12-59   Create a CKD volume with extent pool striping

```
dscli> mkckdvol -extpool p4 -cap 10017 -name ITSO-CKD-STRP -eam rotateexts 0080
CMUC00021I mkckdvol: CKD Volume 0080 successfully created.
```

The `showckdvol` command with the `-rank` option (Example 12-60) shows that the volume that was created is distributed across two ranks. It also displays how many extents on each rank were allocated for this volume.

Example 12-60   Obtain information about a striped CKD volume

```
dscli> showckdvol -rank EF00
...
Name            ITSO_EF00
ID              EF00
accstate        Online
datastate       Normal
```

```
configstate           Normal
deviceMTM             3390-3
volser                -
datatype              3390
voltype               CKD Base
orgbvols              -
addrgrp               E
extpool               P1
exts                  3
cap (cyl)             3339
cap (10^9B)           2.8
cap (2^30B)           2.6
ranks                 1
sam                   Standard
repcapalloc           -
eam                   managed
reqcap (cyl)          3339
cap (Mod1)            3.0
realextents           3
virtualextents        2
realcap (cyl)         3339
migrating             0
migratingcap (cyl)    0
perfgrp               PG0
migratingfrom         -
resgrp                RG0
tierassignstatus      -
tierassignerror       -
tierassignorder       -
tierassigntarget      -
%tierassigned         0
etmonpauseremain      -
etmonitorreset        unknown
==============Rank extents==============
rank extents capacity (MiB/cyl)
===============================
R3         3 3339
```

Dynamic volume expansion

A volume can be expanded without removing the data within the volume. You can specify a new capacity by using the **chckdvol** command, as shown in Example 12-61. The new capacity must be larger than the previous capacity. *You cannot shrink the volume*.

Example 12-61   Expand a striped CKD volume

```
dscli> chckdvol -cap 30051 EF00
CMUC00332W chckdvol: Some host operating systems do not support changing the
volume size. Are you sure that you want to resize the volume? [y/n]: y
CMUC00022I chckdvol: CKD Volume EF00 successfully modified.
```

Because the original volume used the `rotateexts` attribute, the additional extents are also striped, as shown in Example 12-62.

Example 12-62   Check the status of an expanded CKD volume

```
dscli> showckdvol -rank EF00
Name                   ITSO_EF00
ID                     EF00
accstate               Online
datastate              Normal
configstate            Normal
deviceMTM              3390-9
volser                 -
datatype               3390
voltype                CKD Base
orgbvols               -
addrgrp                E
extpool                P1
exts                   27
cap (cyl)              30051
cap (10^9B)            25.5
cap (2^30B)            23.8
ranks                  1
sam                    Standard
repcapalloc            -
eam                    managed
reqcap (cyl)           30051
cap (Mod1)             27.0
realextents            27
virtualextents         4
realcap (cyl)          30051
migrating              0
migratingcap (cyl)     0
perfgrp                PG0
migratingfrom          -
resgrp                 RG0
tierassignstatus       -
tierassignerror        -
tierassignorder        -
tierassigntarget       -
%tierassigned          0
etmonpauseremain       -
etmonitorreset         unknown
==============Rank extents==============
rank extents capacity (MiB/cyl)
===============================
R3         27 30051
```

Important: Before you can expand a volume, you first must delete all Copy Services relationships for that volume. Also, you cannot specify both **-cap** and **-datatype** in the same **chckdvol** command.

It is possible to expand a 3390 Model 9 volume to a 3390 Model A. Expand the volume by specifying a new capacity for an existing Model 9 volume. When you increase the size of a 3390-9 volume beyond 65,520 cylinders, its device type automatically changes to 3390-A.

Important: A 3390 Model A can be used only in z/OS V1.10 (depending on the size of the volume) and later, as shown in Example 12-63.

Example 12-63   Expand a 3390 to a 3390-A

```
*** Command to show CKD volume definition before expansion:

dscli> showckdvol -rank EF00
Name              ITSO_EF00
ID                EF00
accstate          Online
datastate         Normal
configstate       Normal
deviceMTM         3390-3
volser            -
datatype          3390
voltype           CKD Base
orgbvols          -
addrgrp           E
extpool           P1
exts              3
cap (cyl)         3339
cap (10^9B)       2.8
cap (2^30B)       2.6
ranks             1
sam               Standard
repcapalloc       -
eam               managed
reqcap (cyl)      3339
cap (Mod1)        3.0
realextents       3
virtualextents    2
realcap (cyl)     3339
migrating         0
migratingcap (cyl) 0
perfgrp           PG0
migratingfrom     -
resgrp            RG0
tierassignstatus  -
tierassignerror   -
tierassignorder   -
tierassigntarget  -
%tierassigned     0
etmonpauseremain  -
etmonitorreset    unknown
==============Rank extents==============
rank extents capacity (MiB/cyl)
================================
R3         3 3339

*** Command to expand CKD volume from 3390-9 to 3390-A:

dscli> chckdvol -cap 262668 EF00
```

```
CMUC00332W chckdvol: Some host operating systems do not support changing the
volume size. Are you sure that you want to resize the volume? [y/n]: y
CMUC00022I chckdvol: CKD Volume EF00 successfully modified.

*** Command to show CKD volume definition after expansion:

dscli> showckdvol -rank EF00
Name               ITSO_EF00
ID                 EF00
accstate           Online
datastate          Normal
configstate        Normal
deviceMTM          3390-A
volser             -
datatype           3390-A
voltype            CKD Base
orgbvols           -
addrgrp            E
extpool            P1
exts               236
cap (cyl)          262668
cap (10^9B)        223.3
cap (2^30B)        207.9
ranks              1
sam                Standard
repcapalloc        -
eam                managed
reqcap (cyl)       262668
cap (Mod1)         236.0
realextents        236
virtualextents     25
realcap (cyl)      262668
migrating          0
migratingcap (cyl) 0
perfgrp            PG0
migratingfrom      -
resgrp             RG0
tierassignstatus   -
tierassignerror    -
tierassignorder    -
tierassigntarget   -
%tierassigned      0
etmonpauseremain   -
etmonitorreset     unknown
==============Rank extents==============
rank extents capacity (MiB/cyl)
===============================
R3      236 262668
```

You cannot reduce the size of a volume. If you try to reduce the size, an error message is displayed, as shown in Example 12-64.

Example 12-64   Attempt to reduce a volume size

```
dscli> chckdvol -cap 10017 EF00
CMUC00332W chckdvol: Some host operating systems do not support changing the
volume size. Are you sure that you want to resize the volume? [y/n]: y
CMUN02541E chckdvol: EF00: The expand logical volume task was not initiated
because the logical volume capacity that you have requested is less than the
current logical volume capacity.
```

Deleting volumes

CKD volumes can be deleted by using the **rmckdvol** command. FB volumes can be deleted by using the **rmfbvol** command.

The command includes a capability to prevent the accidental deletion of volumes that are in use. A CKD volume is considered *in use* if it participates in a Copy Services relationship, or if the IBM z Systems path mask indicates that the volume is in a grouped state or online to any host system.

If the **-force** parameter is not specified with the command, volumes that are in use are not deleted. If multiple volumes are specified and several volumes are in use and several volumes are not, the volumes that are not in use are deleted.

If the **-force** parameter is specified on the command, the volumes are deleted without checking to see whether they are in use.

Example 12-65 shows trying to delete two volumes, 0900 and 0901. Volume 0900 is online to a host. Volume 0901 is not online to any host and not in a Copy Services relationship. The **rmckdvol 0900-0901** command deletes only volume 0901, which is offline. To delete volume 0900, use the **-force** parameter.

Example 12-65   Delete CKD volumes

```
dscli> lsckdvol 0900-0901
Name   ID   accstate datastate configstate deviceMTM voltype   orgbvols extpool cap (cyl)
==========================================================================================
ITSO_J 0900 Online   Normal    Normal      3390-9    CKD Base -         P1        10017
ITSO_J 0901 Online   Normal    Normal      3390-9    CKD Base -         P1        10017

dscli> rmckdvol -quiet 0900-0901
CMUN02948E rmckdvol: 0900: The Delete logical volume task cannot be initiated because the
Allow Host Pre-check Control Switch is set to true and the volume that you have specified
is online to a host.
CMUC00024I rmckdvol: CKD volume 0901 successfully deleted.

dscli> lsckdvol 0900-0901
Name   ID   accstate datastate configstate deviceMTM voltype   orgbvols extpool cap (cyl)
==========================================================================================
ITSO_J 0900 Online   Normal    Normal      3390-9    CKD Base -         P1        10017

dscli> rmckdvol -force 0900
CMUC00023W rmckdvol: Are you sure you want to delete CKD volume 0900? [y/n]: y
CMUC00024I rmckdvol: CKD volume 0900 successfully deleted.

dscli> lsckdvol 0900-0901
CMUC00234I lsckdvol: No CKD Volume found.
```

12.4.5 Resource groups

The resource group feature is designed for multi-tenancy environments. The resources are volumes, LCUs, and LSSs, and they are used for access control for Copy Services functions only.

For more information about resource groups, see *IBM System Storage DS8000 Copy Services Scope Management and Resource Groups*, REDP-4758.

12.4.6 Performance I/O Priority Manager

By using Performance I/O Priority Manager, you can control quality of service (QoS). Sixteen performance group policies, PG16 - PG31, exist for z/OS.

For more information about I/O Priority Manager, see *DS8000 I/O Priority Manager,* REDP-4760.

12.4.7 IBM Easy Tier

IBM Easy Tier is designed to automate data placement throughout the storage pool. It enables the system, without disrupting applications, to relocate data (at the extent level) across up to three storage tiers. The process is fully automated. Easy Tier also automatically rebalances extents among ranks within the same tier, removing the workload skew between ranks, even within homogeneous and single-tier extent pools.

Easy Tier Heat Map Transfer allows the transfer of Easy Tier heat maps from primary to auxiliary storage sites.

Easy Tier is covered in more detail in 6.7, "IBM Easy Tier" on page 173. Also, see the following publications:

- ► *IBM DS8000 EasyTier*, REDP-4667
- ► *IBM DS8870 Easy Tier Heat Map Transfer*, REDP-5015

12.5 Metrics with DS CLI

This section describes several command examples from the DS CLI that analyze the performance metrics from different levels in a storage unit. The DS GUI also provides new capabilities for performance monitoring, as described in 11.8, "Performance monitoring" on page 332.

Important: The `help` command shows specific information about each of the metrics.

Performance metrics: All performance metrics are an accumulation since the most recent counter-wrap or counter-reset. The performance counters are reset on the following occurrences:

- ► When the storage unit is turned on
- ► When a server fails and the failover and fallback sequence is run

Example 12-66 shows an example of the **showfbvol** command. This command displays the detailed properties for an individual volume and includes a **-metrics** parameter that returns the performance counter-values for a specific volume ID.

Example 12-66   Metrics for a specific fixed-block volume

```
dscli> showfbvol -metrics f000
ID                         F000
normrdrqts                 2814071
normrdhits                 2629266
normwritereq               2698231
normwritehits              2698231
seqreadreqs                1231604
seqreadhits                1230113
seqwritereq                1611765
seqwritehits               1611765
cachfwrreqs                0
cachfwrhits                0
cachefwreqs                0
cachfwhits                 0
inbcachload                0
bypasscach                 0
DASDtrans                  440816
seqDASDtrans               564977
cachetrans                 2042523
NVSspadel                  110897
normwriteops               0
seqwriteops                0
reccachemis                79186
qwriteprots                0
CKDirtrkac                 0
CKDirtrkhits               0
cachspdelay                0
timelowifact               0
phread                     1005781
phwrite                    868125
phbyteread                 470310
phbytewrite                729096
recmoreads                 232661
sfiletrkreads              0
contamwrts                 0
PPRCtrks                   5480215
NVSspallo                  4201098
timephread                 1319861
timephwrite                1133527
byteread                   478521
bytewrit                   633745
timeread                   158019
timewrite                  851671
zHPFRead                   -
zHPFWrite                  -
zHPFPrefetchReq            0
zHPFPrefetchHit            0
GMCollisionsSidefileCount 0
GMCollisionsSendSyncCount 0
```

Example 12-67 show an example of the **showckdvol** command. This command displays the detailed properties for an individual volume and includes a **-metrics** parameter that returns the performance counter-values for a specific volume ID.

Example 12-67   Metrics for a specific CKD volume

```
dscli> showckdvol -metrics 7b3d
ID                      7B3D
normrdrqts              9
normrdhits              9
normwritereq            0
normwritehits           0
seqreadreqs             0
seqreadhits             0
seqwritereq             0
seqwritehits            0
cachfwrreqs             0
cachfwrhits             0
cachefwreqs             0
cachfwhits              0
inbcachload             0
bypasscach              0
DASDtrans               201
seqDASDtrans            0
cachetrans              1
NVSspadel               0
normwriteops            0
seqwriteops             0
reccachemis             0
qwriteprots             0
CKDirtrkac              9
CKDirtrkhits            9
cachspdelay             0
timelowifact            0
phread                  201
phwrite                 1
phbyteread              49
phbytewrite             0
recmoreads              0
sfiletrkreads           0
contamwrts              0
PPRCtrks                0
NVSspallo               0
timephread              90
timephwrite             0
byteread                0
bytewrit                0
timeread                0
timewrite               0
zHPFRead                0
zHPFWrite               0
zHPFPrefetchReq         0
zHPFPrefetchHit         0
GMCollisionsSidefileCount 0
GMCollisionsSendSyncCount 0
```

Example 12-68 shows an example of the output of the `showrank` command. This command generates two types of reports. One report displays the detailed properties of a specified rank, and the other report displays the performance metrics of a specified rank by using the `-metrics` parameter.

Example 12-68   Metrics for a specific rank

```
dscli> showrank -metrics R29
ID             R29
byteread       4721
bytewrit       8466240
Reads          67160
Writes         2099084
timeread       2121
timewrite      1416473
dataencrypted  no
timeload       1935317
```

Example 12-69 shows an example of the `showioport` command. This command shows the properties of a specified I/O port and the performance metrics by using the `-metrics` parameter. Monitoring the I/O ports is one of the most important tasks of the system administrator. The I/O port is where the HBAs, SAN, and DS8880 exchange information. If one of these components has problems because of hardware or configuration issues, all of the other components are affected.

Example 12-69   Metrics for a specific 16 Gbps I/O port

```
dscli> showioport -metrics I0332
ID                      I0332
byteread (FICON/ESCON)  0
bytewrit (FICON/ESCON)  0
Reads (FICON/ESCON)     0
Writes (FICON/ESCON)    0
timeread (FICON/ESCON)  0
timewrite (FICON/ESCON) 0
CmdRetries (FICON)      0
TransferReady (FICON)   0
bytewrit (PPRC)         0
byteread (PPRC)         0
Writes (PPRC)           0
Reads (PPRC)            0
timewrite (PPRC)        0
timeread (PPRC)         0
byteread (SCSI)         80812254
bytewrit (SCSI)         32831993
Reads (SCSI)            119218224
Writes (SCSI)           19247885
timeread (SCSI)         1121646
timewrite (SCSI)        875189
LinkFailErr (FC)        0
LossSyncErr (FC)        0
LossSigErr (FC)         0
PrimSeqErr (FC)         0
InvTxWordErr (FC)       0
CRCErr (FC)             0
LRSent (FC)             0
```

```
LRRec (FC)                0
IllegalFrame (FC)         0
OutOrdData (FC)           0
OutOrdACK (FC)            0
DupFrame (FC)             0
InvRelOffset (FC)         0
SeqTimeout (FC)           0
BitErrRate (FC)           0
RcvBufZero (FC)           272
SndBufZero (FC)           529
RetQFullBusy (FC)         0
ExchOverrun (FC)          0
ExchCntHigh (FC)          0
ExchRemAbort (FC)         0
CurrentSpeed (FC)         16 Gb/s
%UtilizeCPU (FC)          1 Dedicated
TxPower(RDP)              -2.8 dBm(528.5 uW)
RxPower(RDP)              -2.5 dBm(559.6 uW)
TransceiverTemp(RDP)      52 C
SupplyVolt(RDP)           3353.0 mV
TxBiasCurrent(RDP)        7.582 mA
ConnectorType(RDP)        SFP+
TxType(RDP)               Laser-SW
FECStatus(RDP)            Inactive
UncorrectedBitErr(RDP)    -
CorrectedBitErr(RDP)      -
```

The output of the **showioport** command includes several metric counters. For example, the %UtilizeCPU metric for the CPU utilization of the HBA and the CurrentSpeed that the port uses might be useful information.

Example 12-69 on page 397 shows the many important metrics that are returned by the command. It provides the performance counters of the port and the FC link error counters. The FC link error counters are used to determine the health of the overall communication.

The following groups of errors point to specific problem areas:

► Any nonzero figure in the counters LinkFailErr, LossSyncErr, LossSigErr, and PrimSeqErr indicates that unstable HMCs might be attached to the SAN. These unstable HBAs log in to and log out of the SAN, creating name server congestion and performance degradation.

► If the InvTxWordErr counter increases by more than 100 each day, the port is receiving light from a source that is not a small form-factor pluggable (SFP). The cable that is connected to the port is not covered at the end or the I/O port is not covered by a cap.

► The CRCErr counter shows the errors that occur between the last sending SFP in the SAN and the receiving port of the DS8880. These errors do not appear in any other place in the data center. To resolve this issue, replace the cable that is connected to the port or the SFP in the SAN.

► The link reset counters LRSent and LRRec also suggest hardware defects in the SAN. These errors must be investigated.

► The counters IllegalFrame, OutOrdData, OutOrdACK, DupFrame, InvRelOffset, SeqTimeout, and BitErrRate point to congestion in the SAN. These counters can be influenced only by configuration changes in the SAN.

► The 16 Gbps host adapter implements the T11 Read Diagnostic Parameters (RDP) standard and provides extra details, such as optical signal strength, error counters, and other information that is crucial to determining the quality of the link. For example, the cable - connector path (including the cleanliness of the optical connectors) is diagnosed by calculating the RxPower (RDP)/TxPower (RDP) ratio. Receivers rarely fail, and the receiver sensitivity does not change. Therefore, if the receiver optical power is too low for good signal reception and the calculated RxPower (RDP)/TxPower (RDP) ratio is too low, you need to clean the connector. If this RX/TX ratio continues to be low, the cable might be broken.

Consideration: The 8-Gbps host adapter does not support RDP. The corresponding RDP area of the DS CLI command `showioport -metrics` contains no values. Example 12-70 shows a partial output for an 8-Gpbs HBA.

Example 12-70   Partial output for RDP metrics for a specific 8 Gbps I/O port

```
dscli> showioport -metrics IO137
:
CurrentSpeed (FC)      8 Gb/s
:
TxPower(RDP)           0
RxPower(RDP)           0
TransceiverTemp(RDP)   0
SupplyVolt(RDP)        0
TxBiasCurrent(RDP)     0
ConnectorType(RDP)     Unknown
TxType(RDP)            Unknown
FECStatus(RDP)         Inactive
UncorrectedBitErr(RDP) -
CorrectedBitErr(RDP)   -
```

Tip: If you are running on z/OS V2.1 or later on a z13 or z13s processor and have a DS8870 R7.5 or DS8880 R8.1 or later and 16 Gb/s HBA, be sure to check APAR OA49089 at http://www.ibm.com/support/docview.wss?uid=isg1OA49089 and the required documentation at http://publibz.boulder.ibm.com/zoslib/pdf/OA49089.pdf about this command:

`D M=DEV(devno,(chp)),LINKINFO={FIRST/LAST/REFRESH/COMPARE}`

Offload Performance Data and other new parameters

There is a new option to download performance data by using the DS CLI. The command is `offloadfile -perfdata`. It contains performance data that includes one week of data at 1-minute intervals for the system, arrays, pools, and I/O port.

Restriction: The `offloadfile` command is not supported by the embedded DS CLI.

Here is the list of new parameters of `offloadfile` command:

► `-config`: Offloads two files: One file that contains the advanced settings, and another file that contains the Install Corrective Service settings. Only one file set parameter must be specified.

► `-pepackage`: Generates a PE package that is sent to IBM, but only if you do not specify `-showstatus`. Specify only one file set.

- ► **-statesave**: Generates a file that is sent to IBM, but only if you do not specify **-showstatus** or **-list**. Specify only one file set.
- ► **-odd**: Generates a file that is sent to IBM, but only if you do not specify **-showstatus** or **-list**. Specify only one file set.
- ► **-perfdata**: Downloads the performance summary CSV file.

See Example 12-71.

Example 12-71   The offloadfile command issued from the Embedded DS CLI

```
dscli> offloadfile -perfdata c:\temp
Command offloadfile is not supported in Embedded DS CLI.
```

The result of the command at Example 12-71 is a *.csv file* with detailed information. The file is given a name, based on this pattern: StgSysPerfSummary_yyyymmdd.csv. See an illustration in Figure 12-11.

```
Storage System
Time,"Interval (seconds)","Name/ID","IOPS (IOPS) Read","IOPS (IOPS) Write","IOPS (IOPS) Total","Reads
"2016-11-28 10:15:28.799","60.2","DS8K-ESP-03","0.0","0.0","0.0","","0.0","0.0","0.0","0.0","0.0","0.
"2016-11-28 10:16:28.976","60.2","DS8K-ESP-03","0.0","0.0","0.0","","0.0","0.0","0.0","0.0","0.0","0.
"2016-11-28 10:17:29.132","60.2","DS8K-ESP-03","0.0","0.0","0.0","","0.0","0.0","0.0","0.0","0.0","0.
"2016-11-28 10:18:36.035","66.9","DS8K-ESP-03","0.0","0.0","0.0","","0.0","0.0","0.0","0.0","0.0","0.
"2016-11-28 10:19:36.577","60.5","DS8K-ESP-03","0.0","0.0","0.0","","0.0","0.0","0.0","0.0","0.0","0.
"2016-11-28 10:20:36.768","60.2","DS8K-ESP-03","0.0","0.0","0.0","","0.0","0.0","0.0","0.0","0.0","0.
"2016-11-28 10:21:36.95","60.2","DS8K-ESP-03","0.0","0.0","0.0","","0.0","0.0","0.0","0.0","0.0","0.0
"2016-11-28 10:22:37.134","60.2","DS8K-ESP-03","0.0","0.0","0.0","","0.0","0.0","0.0","0.0","0.0","0.
"2016-11-28 10:23:37.304","60.2","DS8K-ESP-03","0.0","0.0","0.0","","0.0","0.0","0.0","0.0","0.0","0.
"2016-11-28 10:24:37.48","60.2","DS8K-ESP-03","0.0","0.0","0.0","","0.0","0.0","0.0","0.0","0.0","0.0
"2016-11-28 10:25:37.722","60.2","DS8K-ESP-03","0.0","0.0","0.0","","0.0","0.0","0.0","0.0","0.0","0.
"2016-11-28 10:26:37.898","60.2","DS8K-ESP-03","0.0","0.0","0.0","","0.0","0.0","0.0","0.0","0.0","0.
"2016-11-28 10:27:38.074","60.2","DS8K-ESP-03","0.0","0.0","0.0","","0.0","0.0","0.0","0.0","0.0","0.
"2016-11-28 10:28:38.251","60.2","DS8K-ESP-03","0.0","0.0","0.0","","0.0","0.0","0.0","0.0","0.0","0.
"2016-11-28 10:29:38.424","60.2","DS8K-ESP-03","0.0","0.0","0.0","","0.0","0.0","0.0","0.0","0.0","0.
"2016-11-28 10:30:38.614","60.2","DS8K-ESP-03","0.0","0.0","0.0","","0.0","0.0","0.0","0.0","0.0","0.
"2016-11-28 10:31:38.779","60.2","DS8K-ESP-03","0.0","0.0","0.0","","0.0","0.0","0.0","0.0","0.0","0.
"2016-11-28 10:32:38.947","60.2","DS8K-ESP-03","0.0","0.0","0.0","","0.0","0.0","0.0","0.0","0.0","0.
"2016-11-28 10:33:39.126","60.2","DS8K-ESP-03","0.0","0.0","0.0","","0.0","0.0","0.0","0.0","0.0","0.
"2016-11-28 10:34:39.294","60.2","DS8K-ESP-03","0.0","0.0","0.0","","0.0","0.0","0.0","0.0","0.0","0.
"2016-11-28 10:35:39.463","60.2","DS8K-ESP-03","0.0","0.0","0.0","","0.0","0.0","0.0","0.0","0.0","0.
"2016-11-28 10:36:39.631","60.2","DS8K-ESP-03","0.0","0.0","0.0","","0.0","0.0","0.0","0.0","0.0","0.
"2016-11-28 10:37:39.8","60.2","DS8K-ESP-03","0.0","0.0","0.0","","0.0","0.0","0.0","0.0","0.0","0.0'
```

Figure 12-11   Snapshot from -perfdata csv file

12.6 Private network security commands

DS CLI commands are available that can be used to manage network security on the DS8880. With the introduction of support for National Institute of Standards and Technology (NIST) 800-131a compliance, new commands were introduced to enable compliance support. Network security includes both Internet Protocol Security (IPSec) to protect the data transmission and Transport Layer Security (TLS) to protect application access.

The following IPSec commands are available:

► `setipsec`

   Use the `setipsec` command to manage IPSec controls. You can enable and disable the IPSec server on the HMCs, either on the primary, secondary, or both.

   > **Note:** The server will not start without defined connections. The server starts automatically when the first connection is created, and stop when the last connection is deleted.

► `mkipsec`

   The `mkipsec` command creates an IPSec connection by importing an IPSec connection configuration file that contains a connection definition to the HMC.

► `rmipsec`

   The `rmipsec` command deletes an IPSec connection from the IPSec server.

► `chipsec`

   The `chipsec` command modifies an existing IPSec connection. Use it to enable and disable existing IPSec connections.

► `lsipsec`

   The `lsipsec` command displays a list of defined IPSec connection configurations.

► `mkipseccert`

   The `mkipseccert` command imports an IPSec certificate to the DS8880.

► `rmipseccert`

   The `rmipseccert` command deletes an IPSec certificate from the HMC.

► `lsipseccert`

   The `lsipseccert` command displays a list of IPSec certificates.

The following commands are available to manage TLS security settings:

► `manageaccess`

   The `manageaccess` command is used to manage the security protocol access settings of an HMC for all communications to and from the DS8880. It can also control port 1750 access to the Network Interface (NI) server for pre-Gen 2 certificate access.

   It is primarily used to manage server access in the HMC, which includes the Common Information Model (CIM) Storage Management Initiative Specification (SMI-S), DS GUI, web user interface (WUI), and NI servers.

   Each of these accesses can be set to an access level of either `Legacy` or `800131a`. When the access is set to the 800131a level, it is now NIST-800131a-compliant.

► `showaccess`

   This command displays the current setting for each access that is managed with the `manageaccess` command. It also displays the remote service access settings that are provided with the `lsaccess` command.

The following security commands are available to manage remote service access settings:

► **chaccess**

   Use the **chaccess** command to change the following settings of an HMC:

   – Enable and disable the command-line shell access to the HMC.
   – Enable and disable the WUI access on the HMC.
   – Enable and disable Assist On-site (AOS) access to the HMC.

   > **Important:**
   >
   > ► This command affects service access only and does not change access to the system by using the DS CLI or DS Storage Manager.
   >
   > ► Only users with administrator authority can access this command.

► **lsaccess**

   The **lsaccess** command displays the access settings of an HMC. If you add the **-l** parameter, the command also displays the AOS status. If AOS is enabled, an AOS connection is active. An AOS connection is only used for remote support purposes.

The following commands are added in Release 8.2 to enable TLS protocol for secure syslog traffic. TLS must be enabled before configuring all syslog servers. If you specify TLS, all Syslog servers configurations use the same protocol and certificates.

► **mksyslogserver**

   Example 12-72 shows the new DS CLI command, **mksyslogserver,** for configuring syslogserver with TLS enabled. The CA Certificate, HMC Certificate, and HMC private key locations are required when configuring the first syslogserver.

Example 12-72   Enable TLS for Secure Syslog Traffic

```
dscli> mksyslogserver -addr 9.100.10.7 -protocol tls -cacert /Users/heping/ut/syslog/ca.pem
-hmccert /Users/heping/ut/syslog/cert.pem -key /Users/heping/ut/syslog/key.pem test
CMUC00508I mksyslogserver: The syslog server machine.example.net has been created.
```

► **lssyslogserver -l**

   The **lssyslogserver -l** displays the list of all Syslog servers and their attributes. See Example 12-73.

Example 12-73   Display List Syslog Servers

```
dscli> lssyslogserver -l
name  IP address    port state   access protocol type            HMC


========================================================================

daisy 9.10.100.127  514 active online tls        audit,message,event 1
daisy 9.10.100.127  514 active online tls        audit,message,event 2
```

The following security commands were added with Release 8.2 to allow a user to set restrict which IP addresses can access each HMC:

> **Note:** Only users who have security administrator authority can configure and change existing IP address and port restrictions.

▶ `setfirewall`

The **setfirewall** command configures IP address and port restrictions for incoming and outgoing network traffic of the HMC firewall. Only users who have security administrator authority can configure and change existing IP address and port restrictions. See Example 12-74.

Example 12-74   The setfirewall command

```
dscli> setfirewall -replace -direction in -ip 9.0.0.0/8 -port 1750,1751,1755 -hmc
2
CMUC00520I setfirewall: The IP address and port restriction is configured
successfully on HMC 2, but the IP address and port restriction is not enabled.
```

In the Example 12-74, note that the restriction is configured but not enabled. To enable it, use the **setfirewall** subparameter **-restriction enable | disable**.

▶ `lsfirewall`

Displays the IP address and port restrictions of the HMC firewall as illustrated in Example 12-75.

Example 12-75   The lsfirewall output

```
dscli> lsfirewall
Interface     Restriction Direction IP          Port            HMC
====================================================================
9.10.100.165 disabled    in         9.0.0.0/8 [1750, 1751, 1755] 1
9.10.100.165 disabled    out        -          []                1
9.10.100.179 disabled    in         9.0.0.0/8 [1750, 1751, 1755] 2
9.10.100.179 disabled    out        -          []                2
```

▶ `lsfirewall -suspend`

Displays suspended information for HMCs as shown in Example 12-76.

Example 12-76   The lsfirewall -suspend command

```
dscli> lsfirewall -suspend
AllowSuspend IsSuspended HMC
============================
yes          no          1
yes          no          2
```

For more information, see Chapter 5 of the *Command-Line Interface User's Guide*, SC27-8526.

> **Important:** For more information about security issues and overall security management to implement NIST 800-131a compliance, see *IBM DS8870 and NIST SP 800-131a Compliance*, REDP-5069.

The following DS CLI commands were added with Release 8.2 for specifying a custom certificate for communication between the encryption key servers (typically SKLM) and the storage system:

▶ `managekeygrp -action -importcert`

Specifies the customer-generated certificate to import in Public-Key Cryptography Standards (PKCS) #12 format as shown in Example 12-77.

Example 12-77   The managekeygrp command with action to import a custom certificate

```
dscli> managekeygrp -action importcert -loc /home/hscroot/rashmic/da6.p12 -pw blah
1
CMUC00489I managekeygrp: The certificate for encryption group 1 has been imported.
```

▶ `managekeygrp -action updatecert`

When specified, the option `-certType CUSTOMER` will update an IBM supplied GEN1 or GEN2 certificate to a customer-generated certificate. as illustrated in Example 12-78. For more information about the encryption and DS8000 certificates, see *IBM DS8880 Data-at-rest Encryption*, REDP-4500.

Example 12-78   The managekeygrp command with action to update a custom certificate

```
dscli> managekeygrp -action updatecert -certType CUSTOMER -key data 1
CMUC00472I managekeygrp: The certificate for encryption group 1 has been updated.
```

For more information, see Chapter 5 of the *Command-Line Interface User's Guide*, SC27-8526.

12.7 Copy Services commands

Many more DS CLI commands are available. Many of these commands are used for the management of Copy Services, such as the FlashCopy, Metro Mirror, and Global Mirror commands.

These commands are not described in this chapter. For more information about these commands, see *DS8000 Copy Services*, SG24-8367.

12.8 Further information

For more information about Copy Services configuration, see *Command-Line Interface User's Guide*, SC27-8526.

For more information about DS CLI commands that relate to disk encryption, see*IBM DS8880 Data-at-rest Encryption*, REDP-4500.

For more information about DS CLI commands that relate to LDAP authentication, see *IBM DS8880 Integrated Copy Services Manager and LDAP Client on the HMC*, REDP-5356.

For more information about DS CLI commands that relate to resource groups, see *IBM System Storage DS8000 Copy Services Scope Management and Resource Groups*, REDP-4758.

For more information about DS CLI commands that relate to Performance I/O Priority Manager, see *DS8000 I/O Priority Manager,* REDP-4760.

For more information about Thin-provisioning, see *IBM DS8880 Thin Provisioning*, REDP-5343.

For more information about DS CLI commands that relate to Easy Tier, see the following publications:

► *IBM DS8000 EasyTier*, REDP-4667
► *DS8870 Easy Tier Application*, REDP-5014

Part 4

<div style="text-align: right">

Part 4

</div>

Maintenance and upgrades

This part of the book provides useful information about maintenance and upgrades.

This part contains the following chapters:

13

Licensed machine code

This chapter describes considerations that relate to the planning and installation of new licensed machine code (LMC) bundles on the IBM DS8880. The overall process for the DS8880 is the same as for previous models. However, several enhancements to power system firmware updates are described.

This chapter covers the following topics:

- ► How new Licensed Internal Code is released
- ► Bundle installation
- ► Code updates
- ► Host adapter firmware updates
- ► Loading the code bundle
- ► Fast path concurrent code load
- ► Postinstallation activities
- ► Summary

13.1 How new Licensed Internal Code is released

Most of the hardware components within the DS8880 system can be updated with new firmware when it is available:

► Device adapters (DAs)

► Host adapters (HAs)

► Power subsystems:
  – Direct current uninterruptible power supply (DC-UPS)
  – Rack power control (RPC)
  – Fibre Channel Interface Cards (FCICs)

In addition, the Licensed Internal Code (LIC) and internal operating system that run on the Hardware Management Consoles (HMCs) and each central processor complex (CPC) can be updated. As IBM continues to develop the DS8880, new features are released through new LMC levels.

When IBM releases new LIC for the DS8880, it is released in the form of a bundle. The term *bundle* is used because a new code release can include updates for various DS8880 components. These updates are tested together, and then the various code packages are bundled together into one unified release. Components within the bundle each include their own revision levels.

For more information about a DS8880 cross-reference table of code bundles, see this website:

http://www.ibm.com/support/docview.wss?uid=ssg1S1004456

The cross-reference table shows the levels of code for currently released bundles. The cross-reference information is updated as new code bundles are released. It is important to maintain a current version of the data storage command-line interface (DS CLI).

The DS8880 uses the following naming convention for bundles: `PR.MM.FFF.EEEE`, where the components are:

► P: Product (8 = DS8880)
► R: Release Major (X)
► MM: Release Minor (xx)
► FFF: Fix Level (xxx)
► EEEE: EFIX Level (0 is base, and 1.*n* is the interim fix build that is later than the base level)

The naming convention is shown in Example 13-1.

Example 13-1   BUNDLE level information

```
For BUNDLE 88.11.25.0 :
Product          DS8880
Release Major    8
Release Minor    11
Fix Level        25
EFIX level       0
```

A release major/minor naming convention, such as 8.11, which is shown in Example 13-1, stands for Release 8.1.1.

If DS CLI is used, you can obtain the CLI and LMC code level information by using the **ver** command as shown in Example 13-2. The **ver** command uses the following optional parameters and displays the versions of the CLI, Storage Manager, and licensed machine code:

► **-s** (Optional): The **-s** parameter displays the version of the CLI program. You cannot use the **-s** and **-l** parameters together.

► **-l** (Optional): The **-l** parameter displays the versions of the CLI, Storage Manager, and licensed machine code. You cannot use the **-l** and **-s** parameters together. See Example 13-2.

► **-cli** (Optional): Displays the version of the CLI program. Version numbers are in the format *version.release.modification.fixlevel*.

► **-stgmgr** (Optional): Displays the version of the Storage Manager.

This ID is not for the graphical user interface (GUI) (Storage Manager GUI). This ID relates to HMC code bundle information.

► **-lmc** (Optional): Displays the version of the LMC.

Example 13-2   Output of DS CLI ver -l command

```
dscli> ver -l
Date/Time: 21 November 2016 14:11:18 CET IBM DSCLI Version: 7.8.20.257 DS: -
DSCLI          7.8.20.257
StorageManager 8.8.4.0.20161006.1
HMC DSCLI      7.8.20.257
================Version=================
Storage Image    LMC         Bundle Version
=======================================
IBM.2107-75BDV31 7.8.20.257 88.20.102.0
```

The LMC level also can be retrieved from the DS Storage Manager by clicking **Actions** →
Properties. See Figure 13-1.
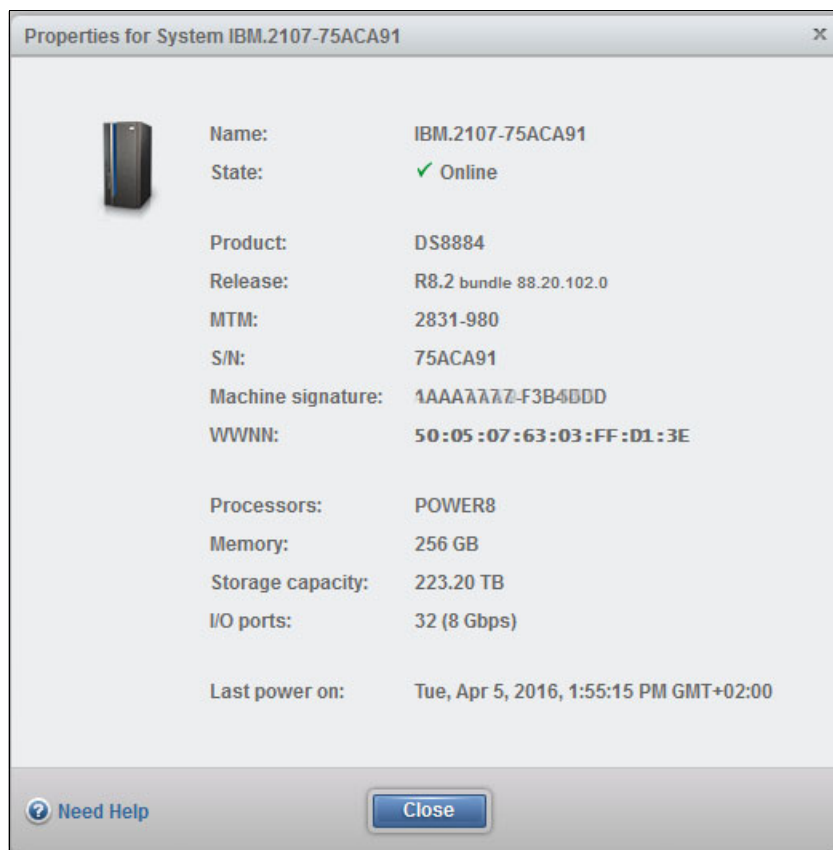
Figure 13-1   LMC level under DS Storage Manager

13.2 Bundle installation

> **Important:** The LMC is always provided by and installed by an IBM service support
> representative (IBM SSR). Installing a new LMC is not the client's task. Read the
> Prerequisites section or Attention Must Read section in the Licensed Internal Code update
> instructions. Your IBM SSR will inform you during the planning phase if any prerequisites
> need to be considered.

The bundle package contains the following new levels of updated code:

► HMC code levels:

   – HMC OS/managed system base

   – DS Storage Manager

   – Common Information Model (CIM) agent version

► Managed system code levels

► Program temporary fix (PTF) code levels

► Storage facility image (SFI) code levels

► HA code levels

- DA code level
- I/O enclosure code level
- Power code levels
- Fibre Channel interface card (FCIC) code levels
- Storage enclosure power supply unit code levels
- Disk drive module (DDM) firmware code level

The Code Distribution and Activation (CDA) software Preload is the current method that is used to run the concurrent code load distribution. By using the CDA software Preload, the IBM SSR performs every non-impacting concurrent code load step for loading code by inserting the physical media into the primary HMC or by running a network acquisition of the code level that is needed. The IBM SSRs can also download the bundle to their notebook and then load it on the HMC by using a service tool. After the CDA Preload software starts, the following steps are performed automatically:

1. Downloads or acquires the release bundle from either the DVD or network to the HMC hard disk drive (HDD).
2. Prepares the HMC with any code update-specific fixes.
3. Distributes the code updates to the logical partition (LPAR) and stages them on an alternate base operating system (BOS) repository.
4. Performs scheduled precheck scans until the distributed code is activated by the user for up to 11 days.
5. After 11 days without Activation, the code expires and is automatically removed from the alternate BOS.

Any time after the software preload completes, when the user logs in to the primary HMC, the user is guided automatically to correct any serviceable events that might be open, update the HMC, and activate the previously distributed code on the storage facility. The overall process is also known as concurrent code load (CCL).

The code installation process involves the following stages:

1. Update code on the primary HMC (HMC1).
2. If a dual HMC configuration is used, the code is acquired from the primary HMC (HMC1) and copied to the secondary HMC (HMC2).
3. Perform updates to the CPC operating system (currently AIX V7.*x*), and updates to the internal LMC, which are performed individually. To update each CPC, the logical subsystem fails over to the alternate CPC. This process updates the firmware that is running on each device adapter that is owned by that CPC.
4. Perform updates to the host adapters. For DS8880 host adapters, the impact of these updates on each adapter is less than 2.5 seconds and they do not affect connectivity. If an update takes longer, the multipathing software on the host or the control-unit initiated reconfiguration (CUIR) directs I/O to another host adapter. If a host is attached with only a single path, or through multipathing to different ports on the same host adapter, connectivity is lost. For more information about host attachments, see 3.3.2, "Host connections" on page 76.

5. New DC-UPS and RPC firmware is periodically released. New firmware can be loaded into each RPC card and DC-UPS directly from the HMC. The DS8880 includes the following enhancements about the power subsystem LIC update for DC-UPS and RPC cards. For more information, see 3.6, "RAS on the power subsystem" on page 90.

   – During the DC-UPS firmware update, the current power state is maintained so the DC-UPS remains operational during a LIC update.

   – During the RPC firmware update, RPC0 is first placed into Service Mode while updated and verified, then resumed and the alternate card RPC1 is completed.

6. New firmware for the hypervisor, service processor, system board, and I/O enclosure boards are periodically released. This firmware can be loaded into each device directly from the HMC. Activation of this firmware might require a shutdown and restart of each CPC individually. This process causes each CPC to fail over its logical subsystems to the alternative CPC. Certain updates do not require this step, or it might occur without processor restarts. For more information, see 3.2, "CPC failover and failback" on page 71.

7. It is important to maintain the latest DDM firmware because enhancements for efficiency, performance, and reliability are often delivered. DDM firmware update is a concurrent process in the DS8880 series family. Direction to update the DDM firmware by using Install Corrective Service (ICS) can be found in the Code Installation Instructions.

Although the microcode installation process might seem complex, it does not require significant user intervention. The IBM SSR normally starts the CDA process and then monitors its progress by using the HMC. From bundle 88.0.*x.x*, power subsystem firmware update activation (RPC cards and DC-UPSs) is included in the same general task that is started at CDA. In previous bundles, it was necessary to start a power update from an option in the HMC when the other elements were already updated. This option remains available when only a power subsystem update is required.

> **Important:** For the DS8880 models, DS CLI should be maintained at a current level. Matching the version to the storage facility is not required if the DS CLI version is at a higher level. The higher level can be used to support all other DS8000 models in the environment as well. For more information, see the Release Notes or consult your local SSR.

13.3 Code updates

The LIC currently running on the HMC is updated as part of a new code bundle installation. The HMC can hold up to six older versions of code in the Code Library. Each CPC can hold three versions of code (the previous version, the active version, and the next version). Most organizations plan for two code updates each year.

> **Preferred practice:** Many clients with multiple DS8000 systems follow the updating schedule that is detailed in this chapter. In this schedule, the HMC is updated a day or two before the rest of the bundle is applied. If a large gap exists between the present and destination level of bundles, certain DS CLI commands (especially DS CLI commands that relate to IBM Copy Services) might not be able to be run until the SFI is updated to the same level of the HMC. Your IBM SSR can assist you in this situation.

Before you update the CPC operating system and LIC, a pre-verification test is run to ensure that no conditions exist that prohibit a successful code load. The HMC code update installs the latest version of the pre-verification test. Then the newest test can be run.

If problems are detected, one or two days are available before the scheduled code installation window date to correct them. This procedure is shown in the following example:

► Thursday:

    a. Acquire the new code bundle to the HMCs.
    b. Update the HMCs to the new code bundle.
    c. Run the updated pre-verification test.
    d. Resolve any issues that were identified by the pre-verification test.

► Saturday:

Update (Activate) the SFI code.

The average code load time varies depending on the hardware that is installed, but 2.5 - 4 hours is normal. Always consult with your IBM SSR about proposed code load schedules. Code bundle preferences are listed on the following website:

http://www.ibm.com/support/docview.wss?uid=ssg1S1004456

Additionally, check multipathing drivers and storage area network (SAN) switch firmware levels for their current levels at regular intervals.

13.4 Host adapter firmware updates

One of the final steps in the concurrent code load process is updating the host adapters. Normally, every code bundle contains new host adapter firmware. For DS8880 Fibre Channel cards, regardless of whether they are used for open systems (FC) attachment or IBM z Systems (Fibre Channel connection (FICON)) attachment, the update process is concurrent to the attached hosts. The Fibre Channel cards use a technique that is known as *adapter fast-load*, which is the default option for updating. This technique allows the cards to switch to the new firmware in less than 2 seconds. This fast update means that single path hosts, hosts that boot from SAN, and hosts that do not have multipathing software do not need to be shut down during the update. They can keep operating during the host adapter update because the update is so fast. Also, no Subsystem Device Driver (SDD) path management is necessary.

Interactive HA also can be enabled if you want to control the host path manually. If so, before the HA cards are updated, a notification is sent and a confirmation is needed. You can then take the corresponding host paths offline and switch to other available paths.

13.4.1 Light-on fastload firmware update

Light-on fastload firmware update is a new DS8880 feature that allows a host adapter to perform a fastload firmware update without dropping light to hosts and SAN switches. It decreases generated registered state change notification (RSCN) messages in SAN fabrics. It also eliminates warning messages or problems that are generated because of the light-off action of a traditional fastload. In addition, this feature shortens the duration of host I/O interrupts during the host firmware update because the DS8880 host ports stay in login states without extra SAN fabric communication actions, such as a login back to the connected SAN switch.

This function is enabled, by default, for DS8880 16 Gbps host adapters. It can also be enabled on 8 Gbps host adapters if you want to use this new feature. Contact your IBM SSR for more information about how to enable this function.

13.4.2 Remote Mirror and Copy path considerations

No special considerations are required for Remote Mirror and Copy paths that use Fibre Channel ports. The ability to perform a fast load means that no interruption occurs to the Remote Mirror operations.

13.4.3 Control-unit initiated reconfiguration

*C*UIR prevents the loss of access to volumes in the IBM z Systems environments because of incorrect or wrong path handling. This function automates channel path management in z Systems environments in support of selected DS8880 service actions. CUIR is available for the DS8880 when the DS8880 is operated in the z/OS and z/VM environments. The CUIR function automates channel path vary on and vary off actions to minimize the manual operator intervention during selected DS8880 service actions.

CUIR allows the DS8880 to request that all attached system images set all paths that are required for a particular service action to the offline state. System images with the correct level of software support respond to these requests by varying off the affected paths. The image then notifies the DS8880 subsystem that the paths are offline, or that it cannot take the paths offline. CUIR reduces manual operator intervention and the possibility of human error during maintenance actions. CUIR also reduces the time that is required for the maintenance window. This feature is useful in environments in which many systems are attached to a DS8880.

13.5 Loading the code bundle

The DS8880 code bundle installation is performed by the IBM SSR. Contact your IBM SSR to review and arrange the required services.

13.6 Fast path concurrent code load

DS8880 supports concurrent code load (CCL). CCL in DS8880 is basically the same as previous generations of DS8000. CCL is referred to as traditional CCL. DS8000 development always strives to improve the code load function robustness and reduce activation durations. A faster, more fault-tolerant code load method is known as Fast Path CCL (FPCCL).

FPCCL is automatically set as the preferred code load function, assuming that the requirements of the bundle to be activated satisfy the requirements for FPCCL.

For Release 8, the FPCCL requirements are expanded to include the following features. The delta of the "coming from" level and the "go to" level consists of these elements:

► SFI code:
  – LPAR code
  – DA

► High-performance flash enclosure (HPFE):
  – SCSI Enclosure Services (SES) processor firmware
  – Power supply unit (PSU) firmware

► HA firmware

► AIX PTF

► Power firmware for DC-UPS and RPC cards

> **Important:** The code load function reverts to traditional CCL if any additional components, other than the components that are listed previously, are included in the update.

FPCCL includes an *autonomic recovery function*, which means that FPCCL is far more tolerant to temporary non-critical errors that might surface during the activation. During the FPCCL, if an error is posted, the LIC automatically analyzes the error and evaluates whether CCL can continue. If it cannot, the LIC suspends the CCL and calls for service. The DS8880 system can continue with the code update with tolerable errors. The DS8880 FPCCL update is more robust with a much shorter duration. After the code update completes, your IBM SSR works to resolve any of the problems that were generated during the code update at a convenient time, allowing DS8880 clients to schedule the code update in a controlled manner.

During an update, a system is under less redundant conditions because certain components are undergoing a firmware update. With FPCCL, firmware activation time is drastically reduced. Therefore, system redundancy is improved with less exposure to non-redundant durations. In addition, firmware distribution time is also minimized because fewer components are involved in the code update.

The CCL duration of the DS8000 family keeps improving with the introduction of new technology. With the latest DS8880 firmware, the LIC preload can be arranged before your code update service window achieves the code activation, distribution, and HMC update. The activation times of various components are greatly reduced.

13.7 Postinstallation activities

After a new code bundle is installed, you might need to complete the following tasks:

1. Upgrade the DS CLI of external workstations. For most of the new release code bundles, a corresponding new release of the DS CLI is available. The LMC version and the DS CLI version are usually identical. Ensure that you upgrade to the new version of the DS CLI to take advantage of any improvements.

   A current version of the DS CLI can be downloaded from Fix Central from the following websites:

   http://www.ibm.com/support/fixcentral
   https://www.ibm.com/support/entry/portal

2. Verify the connectivity from each DS CLI workstation to the DS8880.

3. Verify the DS Storage Manager connectivity by using a supported browser.

4. Verify the DS Storage Manager connectivity from IBM Spectrum Control to the DS8880.

5. Verify the DS Storage Manager connectivity from Copy Services Manager (CSM) to the DS8880.

6. Verify the connectivity from the DS8880 to all IBM Security Key Lifecycle Manager servers in use.

13.8 Summary

IBM might release changes to the DS8880 Licensed Machine Code. These changes might include code fixes and feature updates that relate to the DS8880.

These updates and the information about them are documented in the DS8880 Code cross-reference website. You can find this information for a specific bundle under the *Bundle Release Note information* section on the website.

14

Monitoring with Simple Network Management Protocol

This chapter provides information about Simple Network Management Protocol (SNMP) implementation and messages for the IBM DS8880 storage system.

This chapter covers the following topics:

► SNMP implementation on the DS8880
► SNMP notifications
► SNMP configuration

419

14.1 SNMP implementation on the DS8880

SNMP, as used by the DS8880, is designed so that the DS8880 sends traps only if a notification is necessary. The traps can be sent to a defined IP address.

SNMP alert traps provide information about problems that the storage unit detects. You or the service provider must correct the problems that the traps detect.

The DS8880 does not include an installed SNMP agent that can respond to SNMP polling. The default Community Name parameter is set to `public`.

The management server configured to receive the SNMP traps receives all of the generic trap 6 and specific trap 3 messages, which are sent in parallel with the call home to IBM.

Before SNMP is configured for the DS8880, you are required to get the destination address for the SNMP trap and the information about the port on which the Trap Daemon listens.

> **Standard port:** The standard port for SNMP traps is port 162.

14.1.1 Message Information Base file

The DS8880 storage system provides a Message Information Base (MIB) file that describes the SNMP trap objects. Load the file by using the software that is used for enterprise and SNMP monitoring.

The file is in the `snmp` subdirectory on the data storage command-line interface (DS CLI) installation CD, or the file available on the DS CLI installation CD image that is available from this FTP site:

https://ibm.biz/BdHUVL

14.1.2 Predefined SNMP trap requests

An SNMP agent can send SNMP trap requests to SNMP managers to inform them about the change of values or status on the IP host where the agent is running. Seven predefined types of SNMP trap requests exist, as shown in Table 14-1.

Table 14-1   SNMP trap request types

Trap type	Value	Description
coldStart	0	Restart after a crash.
warmStart	1	Planned restart.
linkDown	2	Communication link is down.
linkUp	3	Communication link is up.
authenticationFailure	4	Invalid SNMP community string was used.
egpNeighborLoss	5	Exterior Gateway Protocol (EGP) neighbor is down.
enterpriseSpecific	6	Vendor-specific event happened.

Each trap message contains an object identifier (OID) and a value, as shown in Table 14-1 on page 420, to notify you about the cause of the trap message. You can also use type 6, the *enterpriseSpecific* trap type, when you must send messages that do not fit the predefined trap types. For example, the DS8880 uses this type for notifications that are described in this chapter.

14.2 SNMP notifications

The Management Console, which is also known as the Hardware Management Console (HMC), of the DS8880 sends an SNMPv1 trap in the following cases:

▶ A serviceable event is reported to IBM by using call home.

▶ An event occurs in the Copy Services configuration or processing.

▶ When the Global Mirror operation pauses on the consistency group boundary.

▶ When the Global Mirror operation fails to unsuspend one or more Global Copy members.

▶ A space-efficient repository or an over-provisioned volume reaches a user-defined warning watermark.

▶ When the rank reaches I/O saturation.

▶ When Encryption Key Management issues an alert that communication between the control unit and one or more Encryption Key Manager (EKM) servers are lost or reconnected.

A serviceable event is posted as a generic trap 6, specific trap 3 message. The specific trap 3 is the only event that is sent for serviceable events and hardware service-related actions (data offload and remote secure connection). For reporting Copy Services events, generic trap 6 and specific traps 100, 101, 102, 200, 202, 210, 211, 212, 213, 214, 215, 216, 217, 218, 219, 220, 225, or 226 are sent.

> **Note:** Consistency group traps (200 and 201) must be prioritized above all other traps. They must be surfaced in less than 2 seconds from the real-time incident.

14.2.1 Serviceable event that uses specific trap 3

Example 14-1 shows the contents of generic trap 6, specific trap 3. The trap holds the following information:

▶ Serial number of the DS8880
▶ Event number that is associated with the manageable events from the HMC
▶ Reporting storage facility image (SFI)
▶ System reference code (SRC)
▶ Location code of the part that is logging the event

The SNMP trap is sent in parallel with a call home for service to IBM and email notification (if configured).

Example 14-1   SNMP special trap 3 of a DS8880

```
Manufacturer=IBM
ReportingMTMS=2107-981*1300960
ProbNm=3084
LparName=SF1300960ESS01
FailingEnclosureMTMS=2107-981*1300960
```

```
SRC=BE80CB13
EventText=Recovery error,the device hardware error.
FruLoc=Part Number 98Y4317  FRU CCIN U401
FruLoc=Serial Number 1731000A39FC
FruLoc=Location Code U2107.D03.G367012-P1-D2
```

For open events in the event log, a trap is sent every eight hours until the event is closed.

14.2.2 Copy Services event traps

For state changes in a remote Copy services environment, 13 traps are implemented. The
1*xx* traps are sent for a state change of a physical link connection. The 2*xx* traps are sent for
state changes in the logical Copy Services setup. For all of these events, no call home is
generated and IBM is not notified.

This chapter describes only the messages and the circumstances when traps are sent by the
DS8880. For more information about these functions and terms, see *IBM DS8870 Copy
Services for IBM z Systems*, SG24-6787, and *IBM DS8870 Copy Services for Open Systems*,
SG24-6788.

Physical connection events

Within the trap 1*xx* range, a state change of the physical links is reported. The trap is sent if
the physical remote copy link is interrupted. The Link trap is sent from the primary system.
The PLink and SLink columns are only used by the 2105 Enterprise Storage Server disk unit.

If one or several links (but not all links) are interrupted, a trap 100 (as shown in
Example 14-2), is posted. Trap 100 indicates that the redundancy is degraded. The reference
code (RC) column in the trap represents the return code for the interruption of the link.

Example 14-2   Trap 100: Remote Mirror and Copy links degraded

```
PPRC Links Degraded
UNIT: Mnf Type-Mod SerialNm LS
PRI:  IBM 2107-981 75-ZA571 12
SEC:  IBM 2107-981 75-CYK71 24
Path: Type  PP    PLink  SP    SLink  RC
1:    FIBRE 0143 XXXXXX 0010 XXXXXX 15
2:    FIBRE 0213 XXXXXX 0140 XXXXXX OK
```

If all of the links are interrupted, a trap 101 (as shown in Example 14-3) is posted. This event
indicates that no communication between the primary and the secondary system is possible.

Example 14-3   Trap 101: Remote Mirror and Copy links are inoperable

```
PPRC Links Down
UNIT: Mnf Type-Mod SerialNm LS
PRI:  IBM 2107-981 75-ZA571 10
SEC:  IBM 2107-981 75-CYK71 20
Path: Type  PP    PLink  SP    SLink  RC
1:    FIBRE 0143 XXXXXX 0010 XXXXXX 17
2:    FIBRE 0213 XXXXXX 0140 XXXXXX 17
```

After the DS8880 can communicate again by using any of the links, trap 102 (as shown in Example 14-4) is sent after one or more of the interrupted links are available again.

Example 14-4 Trap 102: Remote Mirror and Copy links are operational

```
PPRC Links Up
UNIT: Mnf Type-Mod SerialNm LS
PRI:  IBM 2107-981 75-ZA571 21
SEC:  IBM 2107-981 75-CYK71 11
Path: Type  PP    PLink  SP    SLink  RC
1:    FIBRE 0010  XXXXXX 0143  XXXXXX OK
2:    FIBRE 0140  XXXXXX 0213  XXXXXX OK
```

Remote Mirror and Copy events

If you configured consistency groups and a volume within this consistency group is suspended because of a write error to the secondary device, trap 200 is sent, as shown in Example 14-5. One trap for each logical subsystem (LSS) that is configured with the consistency group option is sent. This trap can be handled by automation software, such as Copy Services Manager, to freeze this consistency group. The SR column in the trap represents the suspension reason code, which explains the cause of the error that suspended the Remote Mirror (Peer-to-Peer Remote Copy (PPRC)) and Copy group. The suspension reason codes are listed in Table 14-2 on page 427.

Example 14-5 Trap 200: LSS pair consistency group Remote Mirror and Copy pair error

```
LSS-Pair Consistency Group PPRC-Pair Error
UNIT: Mnf Type-Mod SerialNm LS LD SR
PRI:  IBM 2107-981 75-ZA571 84 08
SEC:  IBM 2107-981 75-CYM31 54 84
```

Trap 202, as shown in Example 14-6, is sent if a Remote Copy pair goes into a suspend state. The trap contains the serial number (`SerialNm`) of the primary and secondary machine, the LSS (`LS`), and the logical device (`LD`). To avoid SNMP trap flooding, the number of SNMP traps for the LSS is throttled. The complete suspended pair information is represented in the summary. The last row of the trap represents the suspend state for all pairs in the reporting LSS. The suspended pair information contains a hexadecimal string of 64 characters. By converting this hex string into binary code, each bit represents a single device. If the bit is 1, the device is suspended. Otherwise, the device is still in full duplex mode.

Example 14-6 Trap 202: Primary Remote Mirror and Copy devices on LSS suspended due to error

```
Primary PPRC Devices on LSS Suspended Due to Error
UNIT: Mnf Type-Mod SerialNm LS LD SR
PRI:  IBM 2107-981 75-ZA571 28 00 01
SEC:  IBM 2107-981 75-CZM21 a8 00
Start: 2015/11/14 10:30:32 CST
PRI Dev Flags (1 bit/Dev, 1=Suspended):
C000000000000000000000000000000000000000000000000000000000000000
```

Trap 210, as shown in Example 14-7, is sent when a consistency group in a Global Mirror environment was successfully formed.

Example 14-7   Trap 210: Global Mirror initial consistency group successfully formed

```
Asynchronous PPRC Initial Consistency Group Successfully Formed
UNIT: Mnf Type-Mod SerialNm
IBM 2107-981 75-ZA571
Session ID: 4002
```

As shown in Example 14-8, trap 211 is sent if the Global Mirror setup is in a severe error state in which no attempts are made to form a consistency group.

Example 14-8   Trap 211: Global Mirror Session is in a fatal state

```
Asynchronous PPRC Session is in a Fatal State
UNIT: Mnf Type-Mod SerialNm
IBM 2107-981 75-CYM21
Session ID: 4002
```

Trap 212, as shown in Example 14-9, is sent when a consistency group cannot be created in a Global Mirror relationship for one of the following reasons:

► Volumes were taken out of a copy session.
► The Remote Copy link bandwidth might not be sufficient.
► The Fibre Channel (FC) link between the primary and secondary system is not available.

Example 14-9   Trap 212: Global Mirror consistency group failure: Retry is attempted

```
Asynchronous PPRC Consistency Group Failure - Retry will be attempted
UNIT: Mnf Type-Mod SerialNm
IBM 2107-981 75-ZA571
Session ID: 4002
```

Trap 213, as shown in Example 14-10, is sent when a consistency group in a Global Mirror environment can be formed after a previous consistency group formation failure.

Example 14-10   Trap 213: Global Mirror consistency group successful recovery

```
Asynchronous PPRC Consistency Group Successful Recovery
UNIT: Mnf Type-Mod SerialNm
IBM 2107-981 75-ZA571
Session ID: 4002
```

Trap 214, as shown in Example 14-11, is sent if a Global Mirror session is ended by using the DS CLI `rmgmir` command or the corresponding graphical user interface (GUI) function.

Example 14-11   Trap 214: Global Mirror master terminated

```
Asynchronous PPRC Master Terminated
UNIT: Mnf Type-Mod SerialNm
IBM 2107-981 75-ZA571
Session ID: 4002
```

As shown in Example 14-12, trap 215 is sent if, in the Global Mirror environment, the master detects a failure to complete the FlashCopy commit. The trap is sent after many commit retries fail.

Example 14-12   Trap 215: Global Mirror FlashCopy at remote site unsuccessful

```
Asynchronous PPRC FlashCopy at Remote Site Unsuccessful
A UNIT: Mnf Type-Mod SerialNm
IBM 2107-981 75-CZM21
Session ID: 4002
```

Trap 216, as shown in Example 14-13, is sent if a Global Mirror master cannot end the Global Copy relationship at one of its subordinates. This error might occur if the master is ended by using the `rmgmir` command but the master cannot end the copy relationship on the subordinate.

You might need to run a `rmgmir` command against the subordinate to prevent any interference with other Global Mirror sessions.

Example 14-13   Trap 216: Global Mirror subordinate termination unsuccessful

```
Asynchronous PPRC Slave Termination Unsuccessful
UNIT:   Mnf Type-Mod SerialNm
Master: IBM 2107-981 75-ZA571
Slave:  IBM 2107-981 75-CYM31
Session ID: 4002
```

Trap 217, as shown in Example 14-14, is sent if a Global Mirror environment is suspended by the DS CLI command `pausegmir` or the corresponding GUI function.

Example 14-14   Trap 217: Global Mirror paused

```
Asynchronous PPRC Paused
UNIT: Mnf Type-Mod SerialNm
IBM 2107-981 75-CYM31
Session ID: 4002
```

As shown in Example 14-15, trap 218 is sent if a Global Mirror exceeded the allowed threshold for failed consistency group formation attempts.

Example 14-15   Trap 218: Global Mirror number of consistency group failures exceeds threshold

```
Global Mirror number of consistency group failures exceed threshold
UNIT: Mnf Type-Mod SerialNm
IBM 2107-981 75-ZA571
Session ID: 4002
```

Trap 219, as shown in Example 14-16, is sent if a Global Mirror successfully formed a consistency group after one or more formation attempts previously failed.

Example 14-16   Trap 219: Global Mirror first successful consistency group after prior failures

```
Global Mirror first successful consistency group after prior failures
UNIT: Mnf Type-Mod SerialNm
IBM 2107-981 75-ZA571
Session ID: 4002
```

Trap 220, as shown in Example 14-17, is sent if a Global Mirror exceeded the allowed threshold of failed FlashCopy commit attempts.

Example 14-17   Trap 220: Global Mirror number of FlashCopy commit failures exceeds threshold

```
Global Mirror number of FlashCopy commit failures exceed threshold
UNIT: Mnf Type-Mod SerialNm
IBM 2107-981 75-ZA571
Session ID: 4002
```

Trap 225, as shown in Example 14-18, is sent when a Global Mirror operation paused on the consistency group boundary.

Example 14-18   Trap 225: Global Mirror paused on the consistency group boundary

```
Global Mirror operation has paused on the consistency group boundary
UNIT: Mnf Type-Mod SerialNm
IBM 2107-981 75-CYM31
Session ID: 4002
```

Trap 226, in Example 14-19, is sent when a Global Mirror operation failed to unsuspend one or more Global Copy members.

Example 14-19   Trap 226: Global Mirror unsuspend members failed

```
Global Mirror operation has failed to unsuspend one or more Global Copy members
UNIT: Mnf Type-Mod SerialNm
IBM 2107-981 75-CYM31
Session ID: 4002
```

Table 14-2 shows the Copy Services suspension reason (SR) codes.

Table 14-2   Copy Services suspension reason codes

Suspension reason code	Description
03	The host system sent a command to the primary volume of a Remote Mirror and Copy volume pair to suspend copy operations. The host system might specify an immediate suspension or a suspension after the copy completed and the volume pair reached a full duplex state.
04	The host system sent a command to suspend the copy operations on the secondary volume. During the suspension, the primary volume of the volume pair can still accept updates, but updates are not copied to the secondary volume. The out-of-sync tracks that are created between the volume pair are recorded in the change recording feature of the primary volume.
05	Copy operations between the Remote Mirror and Copy volume pair were suspended by a primary storage unit secondary device status command. This system resource code can be returned only by the secondary volume.
06	Copy operations between the Remote Mirror and Copy volume pair were suspended because of internal conditions in the storage unit. This system resource code can be returned by the control unit of the primary volume or the secondary volume.
07	Copy operations between the remote mirror and copy volume pair were suspended when the secondary storage unit notified the primary storage unit of a state change transition to the simplex state. The specified volume pair between the storage units is no longer in a copy relationship.
08	Copy operations were suspended because the secondary volume became suspended because of internal conditions or errors. This system resource code can be returned only by the primary storage unit.
09	The Remote Mirror and Copy volume pair was suspended when the primary or secondary storage unit was rebooted or when the power was restored. The paths to the secondary storage unit might not be disabled if the primary storage unit was turned off. If the secondary storage unit was turned off, the paths between the storage units are restored automatically, if possible. After the paths are restored, issue the `mkpprc` command to resynchronize the specified volume pairs. Depending on the state of the volume pairs, you might need to issue the `rmpprc` command to delete the volume pairs and reissue a `mkpprc` command to reestablish the volume pairs.
0A	The Remote Mirror and Copy pair was suspended because the host issued a command to freeze the Remote Mirror and Copy group. This system resource code can be returned only if a primary volume was queried.

14.2.3 I/O Priority Manager SNMP

When the I/O Priority Manager Control switch is set to `Monitor` or `Managed`, an SNMP trap alert message also can be enabled. The DS8880 Licensed Internal Code (LIC) monitors for *rank saturation*. If a rank is overdriven to the point of saturation (high usage), an SNMP trap alert message 224 is posted to the SNMP server.

The following SNMPs rules are followed:

► Up to eight SNMP traps for each SFI server in a 24-hour period (maximum: 16 for each 24 hours for each SFI).

► The rank enters the saturation state if it is in saturation for five consecutive 1-minute samples.

► The rank exits the saturation state if it is not in saturation for three of five consecutive 1-minute samples.

► The SNMP message 224 is reported when a rank enters saturation or every eight hours if the rank is in saturation. The message identifies the rank and SFI. See Example 14-20.

Example 14-20   Trap 224: Rank saturation status changed

```
Rank Saturated
UNIT: Mnf Type-Mod SerialNm
IBM 2107-981 75-ZA571
Rank ID: R21
Saturation Status: 0
```

Important: To receive traps from I/O Priority Manager (IOPM), set the IOPM to manage SNMP by issuing the following command:

```
chsi -iopmmode managesnmp <Storage_Image>
```

14.2.4 Thin provisioning SNMP

The DS8880 can trigger two specific SNMP trap alerts that relate to the thin provisioning feature. The trap is sent out when certain extent pool capacity thresholds are reached, which causes a change in the extent status attribute. A trap is sent under the following conditions:

► The extent status is not zero (available space is already below threshold) when the first extent space-efficient (ESE) volume is configured.

► ESE volumes are configured in the extent pool.

Example 14-21 shows an example of generated event trap 221.

Example 14-21   Trap 221: Space-efficient repository or over-provisioned volume reached a warning

```
Space Efficient Repository or Over-provisioned Volume has reached a warning
watermark
Unit: Mnf Type-Mod SerialNm
IBM 2107-981 75-ZA571
Volume Type: repository
Reason Code: 1
Extent Pool ID: f2
Percentage Full: 100%
```

Example 14-22 shows an example of generated event trap 223.

Example 14-22   Trap 223: Extent pool capacity reached a warning threshold

```
Extent Pool Capacity Threshold Reached
UNIT: Mnf Type-Mod SerialNm
IBM 2107-981 75-ZA571
Extent Pool ID: P1
Limit: 95%
Threshold: 95%Status: 0
```

14.3 SNMP configuration

The SNMP for the DS8880 is designed to send traps as notifications. The DS8880 does not include an installed SNMP agent that can respond to SNMP polling. Also, the SNMP community name for Copy Services-related traps is fixed and set to `public`.

14.3.1 SNMP preparation

During the planning for the installation (8.3.4, "Monitoring DS8880 with the Management Console" on page 212), the IP addresses of the management system are provided for IBM service personnel. This information must be applied by the IBM service support representative (SSR) during the installation. Also, the IBM SSR can configure the HMC to send a notification for every serviceable event or for only those events that call home to IBM.

The network management server that is configured on the HMC receives all of the generic trap 6, specific trap 3 messages, which are sent in parallel with any events that call home to IBM.

The SNMP alerts can contain a combination of a generic and a specific alert trap. The Traps list outlines the explanations for each of the possible combinations of generic and specific alert traps. The format of the SNMP traps, the list, and the errors that are reported by SNMP are available in the "Generic and specific alert traps" section of the Troubleshooting section of the IBM Knowledge Center for the DS8880 at the following site:

https://ibm.biz/BdH5zf

SNMP alert traps provide information about problems that the storage unit detects. You or the IBM SSR must perform corrective action for the related problems.

14.3.2 SNMP configuration with the HMC

Clients can configure the SNMP alerting by logging in to the DS8880 Service web user interface (WUI). The Service WUI can be started from the DS8000 Storage Management Console (HMC) (`https://<HMC_ip_address>/service`) remotely through a web browser. Access the Service Management Console and log in with the following client credentials:

▶   User ID: `customer`
▶   Password: `cust0mer` (default Password)

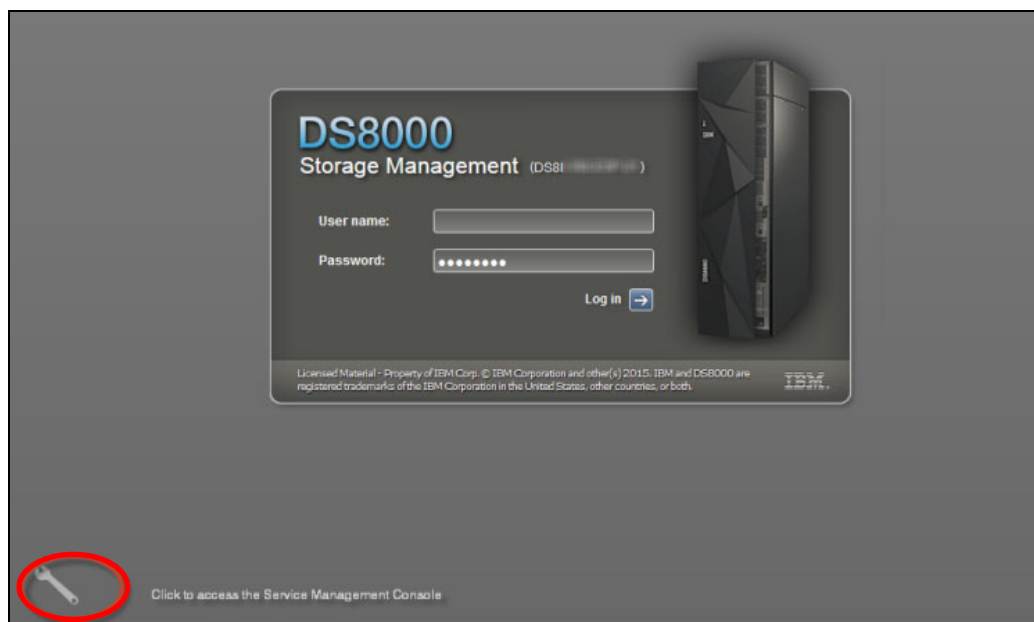Figure 14-1 shows the DS8000 Service Management Console login.

Figure 14-1   Remote access to Service Management Console

Complete the following steps to configure SNMP at the HMC:

1. Log in to the Service Management section on the HMC, as shown in Figure 14-2.
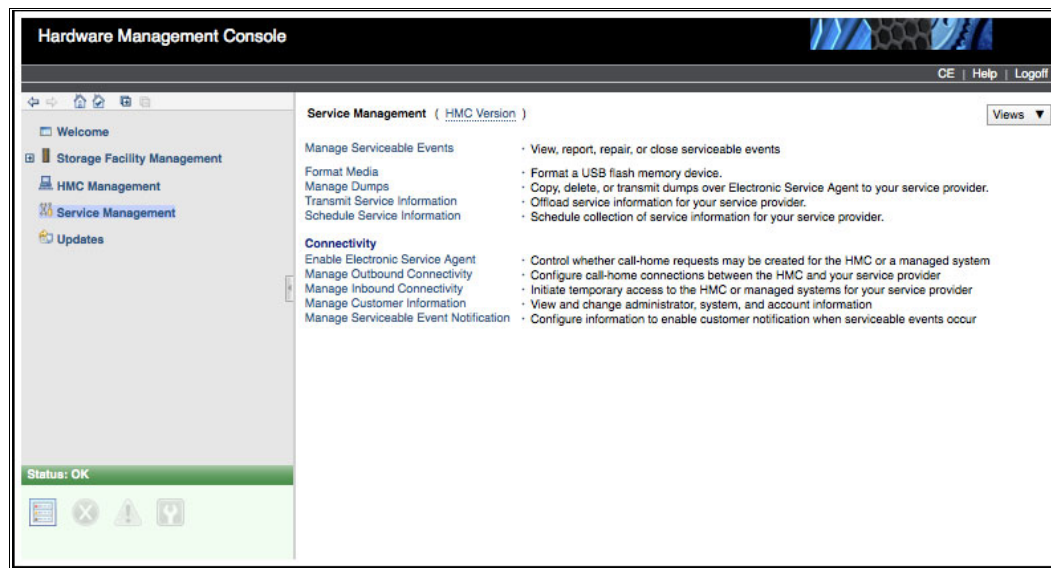

Figure 14-2   HMC Service Management

2. Select **Manage Serviceable Event Notification**, as shown in Figure 14-3, and enter the TCP/IP information of the SNMP server in the SNMP Trap Configuration folder.
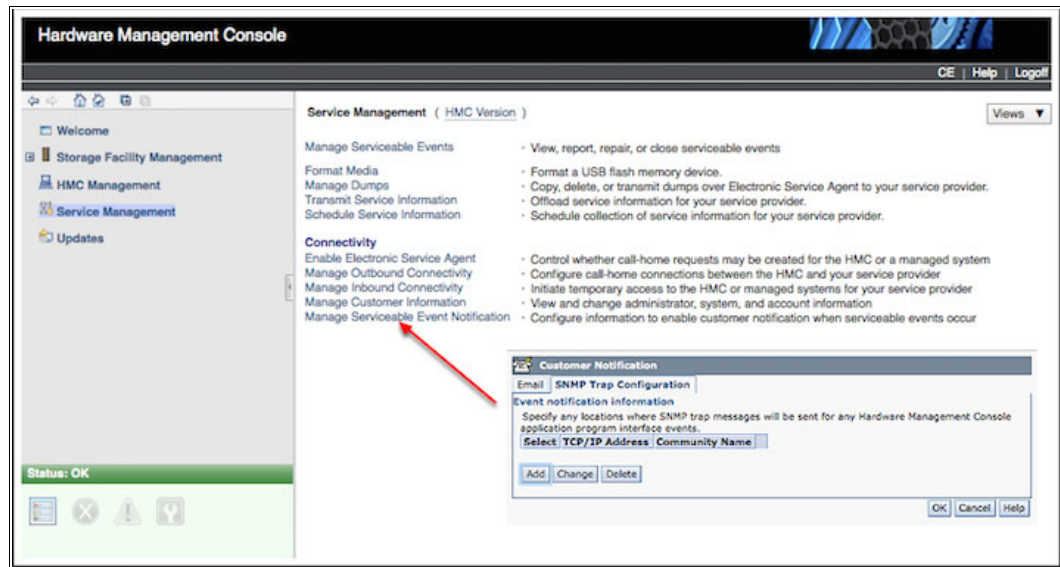

Figure 14-3   HMC Manage Serviceable Event Notification

3. To verify the successful setup of your environment, create a Test Event on your DS8880 Management Console. Select **Storage Facility Management** → **Services Utilities** → **Test Problem Notification (PMH, SNMP, Email)**, as shown in Figure 14-4.
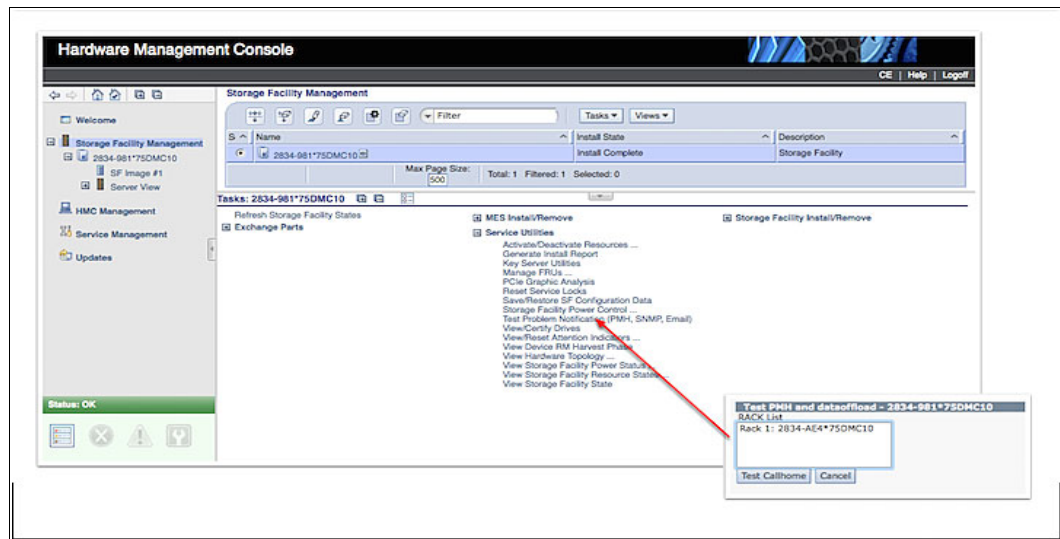

Figure 14-4   HMC test SNMP trap

4. The test generates the Service Reference Code BEB20010 and the SNMP server receives the SNMP trap notification, as shown in Figure 14-5.
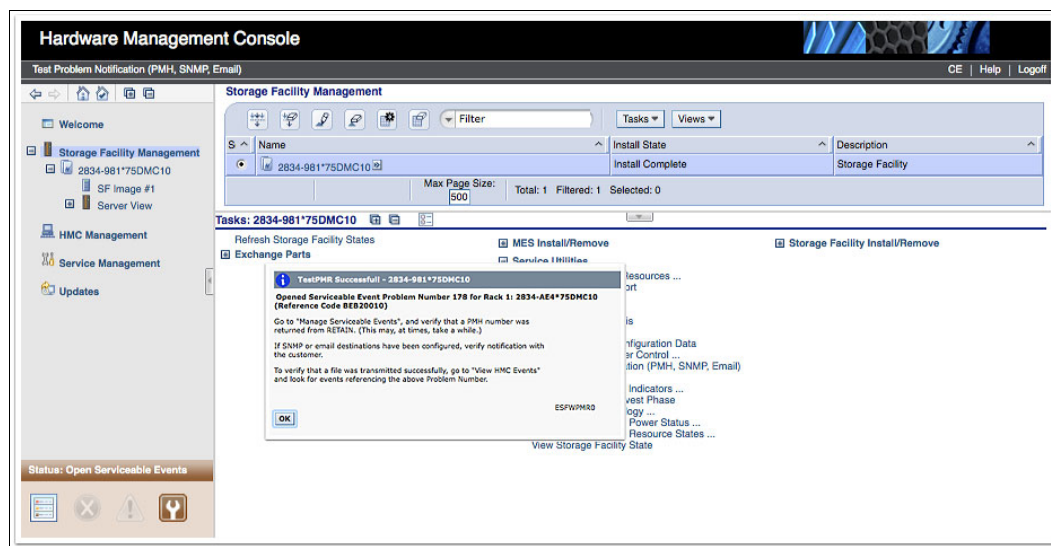

Figure 14-5 HMC SNMP trap test

14.3.3 SNMP configuration with the DS CLI

Perform the configuration process for receiving the operation-related traps, such as for Copy Services, thin provisioning, Encryption, or I/O Priority Manager, by using the DS CLI. Example 14-23 shows how SNMP is enabled by using the `chsp` command.

Example 14-23 Configuring the SNMP by using dscli

```
dscli> chsp -snmp on -snmpaddr 10.10.10.1,10.10.10.2
CMUC00040I chsp: Storage complex IbmStoragePlex successfully modified.

dscli> showsp
Name           IbmStoragePlex
desc           -
acct           -
SNMP           Enabled
SNMPadd        10.10.10.1,10.10.10.2
emailnotify    Disabled
emailaddr      -
emailrelay     Disabled
emailrelayaddr -
emailrelayhost -
numkssupported 4
```

SNMP preparation for the management software

To enable the trap-receiving software to display the correctly decoded message in a human-readable format, load the DS8880 specific MIB file.

The MIB file that is delivered with the latest DS8880 DS CLI CD is compatible with all previous levels of DS8880 LIC. Therefore, ensure that you loaded the latest MIB file that is available.

15

Remote support

This chapter describes the outbound (call home and support data offload) and inbound (code download and remote support) communications for the IBM System Storage DS8000 family.

The DS8880 maintains the same functions as in the previous generation.

This chapter covers the following topics:

- ► Introduction to remote support
- ► IBM policies for remote support
- ► Remote support advantages
- ► Remote support call home
- ► Remote Support Access (inbound)
- ► Audit logging

15.1 Introduction to remote support

IBM provides remote support capabilities for the DS8880. The remote support enables the storage to communicate with IBM, and allows IBM support to remotely connect to the system when authorized by the client.

The benefits of the remote support are that IBM Support can respond quickly to events reported by you or the system.

The following features can be enabled in the DS8880 for remote support:

► Call home support (outbound remote support):

  – Reporting problems to IBM
  – Sending heartbeat information
  – Offloading data

► Remote service (inbound) remote support

  IBM Service support access the DS8880 HMC through a network-based connection.

During the installation and planning phase, complete the remote support worksheets and supply them to the IBM SSR at the time of the installation.

The worksheets are holding information about your remote support preferences, and the network communication requirements that needs to be fulfilled by your local network.

15.2 IBM policies for remote support

The following guidelines are at the core of the IBM remote support strategies for the DS8880:

► When the DS8880 transmits service data to IBM, only logs and process memory dumps are gathered for troubleshooting.

► When a remote session with the DS8880 is needed, the HMC or Management Console always initiates an outbound connection to predefined IBM servers or ports.

► IBM maintains multiple-level internal authorizations for any privileged access to the DS8880 components. Only approved IBM service personnel can gain access to the tools that provide the security codes for HMC command-line access.

Although the Management Console is based on a Linux operating system, IBM disabled or removed all unnecessary services, processes, and IDs, including standard internet services, such as Telnet (Telnet server is disabled on the HMC), File Transfer Protocol (FTP), **r** commands (Berkeley r-commands and Remote Procedure Call (RPC) commands), and RPC programs.

15.3 Remote support advantages

The following benefits can be realized when you enable remote support on the DS8880:

► Serviceable events with related problem data are reported to IBM automatically and a support call is opened.

► IBM support personnel can start data analysis and problem isolation immediately, which can reduce the overall time that is required to fix a problem.

► If additional service data is needed, IBM Support can connect to the Management Console and offload the data for the next level of support.

► Remote support helps clients to maintain the highest availability of their data.

15.4 Remote support call home

This section details the call home characteristics.

15.4.1 Call home and heartbeat: Outbound

This section describes the call home and heartbeat capabilities.

Call home

Call home is the capability of the Management Console to report serviceable events to IBM. The Management Console also transmits machine-reported product data (MRPD) information to IBM through call home. The MRPD information includes installed hardware, configurations, and features. The call home is configured by the IBM SSR during the installation of the DS8880 by using the customer worksheets. A test call home is placed after the installation to register the machine and verify the call home function.

Heartbeat

The DS8880 also uses the call home facility to send proactive *heartbeat* information to IBM. The heartbeat configuration can be set by the IBM SSR to send heartbeat information to the customer (through Simple Network Management Protocol (SNMP) and email) in addition to IBM. A *heartbeat* is a small message with basic product information that is sent to IBM to ensure that the call home function works. The heartbeat can be scheduled every one to seven days based on the client's preference. When a scheduled heartbeat fails to transmit, a service call is placed for the SSR with an action plan to verify the call home function. The DS8880 uses an internet connection through Transport Layer Security (TLS), which is also known as Secure Sockets Layer (SSL), for call home functions.

15.4.2 Data offload: Outbound

For many DS8880 problem events, such as a hardware component failure, a large amount of diagnostic data is generated. This data can include text and binary log files, firmware information, inventory lists, and timelines. These logs are grouped into collections by the component that generated them or the software service that owns them.

The entire bundle is collected together in a *PEPackage*. A DS8880 PEPackage can be large, often exceeding 100 MB. In certain cases, more than one PEPackage might be needed to diagnose a problem correctly. In certain cases, the IBM Support center might need an extra memory dump that is internally created by the DS8880 or manually created through the intervention of an operator.

> **On Demand Dump:** The On-Demand Data Dump (ODD) provides a mechanism that allows the collection of debug data for error scenarios. With ODD, IBM can collect data with no impact to the host I/O after an initial error occurs. ODD can be generated by using the data storage command-line interface (DS CLI) command `diagsi -action odd` and then offloaded.

The Management Console is a focal point for gathering and storing all of the data packages. Therefore, the Management Console must be accessible if a service action requires the information. The data packages must be offloaded from the Management Console and sent in to IBM for analysis. The offload is performed through the internet through a TLS connection.

15.4.3 Outbound connection types

This section describes the outbound connection options that are available for call home and data offload.

Internet through a TLS connection

The preferred remote support connectivity method is internet TLS for management console to IBM communication. TLS is the encryption protocol that was originally developed as a secured web communication standard. Traffic through a TLS proxy is supported with or without authentication based on the client's proxy server configuration.

When the internet is selected as the outbound connectivity method, the Management Console (MC) uses a TLS connection over the internet to connect to the IBM.

For more information about IBM TLS remote support, see the *IBM DS8880 Introduction and Planning Guide*, GC27-8525, for planning and worksheets.

Standard FTP connection for data offload

The Management Console can be configured to support automatic data offload by using FTP over a network connection. This traffic can be examined at the client's firewall before it is moved across the Internet. For FTP, the Management Console must be connected to customer LAN with a path to the Internet from the repository server.

> **Important:** FTP offload of data is supported as an outbound service only. No active FTP server is running on the HMC that can receive connection requests.

When a direct FTP session across the Internet is not available or wanted, a client can configure the FTP offload to use a client-provided FTP proxy server. The client then becomes responsible for configuring the proxy to forward the data to IBM.

The client is required to manage its firewalls so that FTP traffic from the Management Console (or from an FTP proxy) can pass onto the Internet.

For more information, see the *IBM DS8880 Introduction and Planning Guide*, GC27-8525.

15.5 Remote Support Access (inbound)

IBM took many necessary steps to provide secure network access for the Management Console. The client can define how and when the IBM SSR can connect to the Management Console. When remote support access is configured, IBM Support can connect to the Management Console to start problem analysis and data gathering. This process allows data to be analyzed as fast as possible with an action plan that is created for an onsite IBM SSR, if needed.

Having inbound access enabled can greatly reduce the problem resolution time by not waiting for the SSR to arrive onsite to gather problem data and upload it to IBM. With the DS8880, inbound connectivity options are available to the client:

► External Assist On-Site Gateway
► Embedded remote access feature

The remote support access connection cannot be used to send support data to IBM.

The support data offload always uses the Call Home feature.

15.5.1 Assist On-site

Assist On-site (AOS) is an IBM remote access solution that relies on the IBM commercial product IBM BigFix® for Remote Control. The IBM DS8000 support uses the Port-Forwarding feature to maintain the DS8000 with an IP-based maintenance tool.

IBM Support encourages you to use Assist On-site as your remote access method.

The remote access connection is secured with TLS 1.2. In addition, a mechanism is implemented so that the HMC only communicates network wise as outbound connection while you must specifically allow IBM to connect to the HMC at any time. You can compare this function to that of a modem that picks up incoming calls. The DS8880 documentation refers to this as an unattended service.

The connection is under control of the DS8880 administrator at all time. Any DS8880 administrator can start and stop the AOS connection.

For more information, see 15.5.4, "Support access management through the DS CLI and the Storage Management GUI" on page 439.

When you prefer to have a centralized access point for IBM Support, then an Assist On-site Gateway might be the correct solution. With the AOS Gateway, install the AOS Software externally to a DS8880 HMC. You need to install the AOS software on a system you provide and maintain. IBM Support only provides the AOS software package. Through port-forwarding on an AOS Gateway, you can configure remote access to one or more DS8880s or other IBM Storage Systems.

A simple AOS connection to the DS8880 is shown in Figure 15-1. For more information about AOS, prerequisites, and installation, see *IBM Assist On-site for Storage Overview*, REDP-4889.
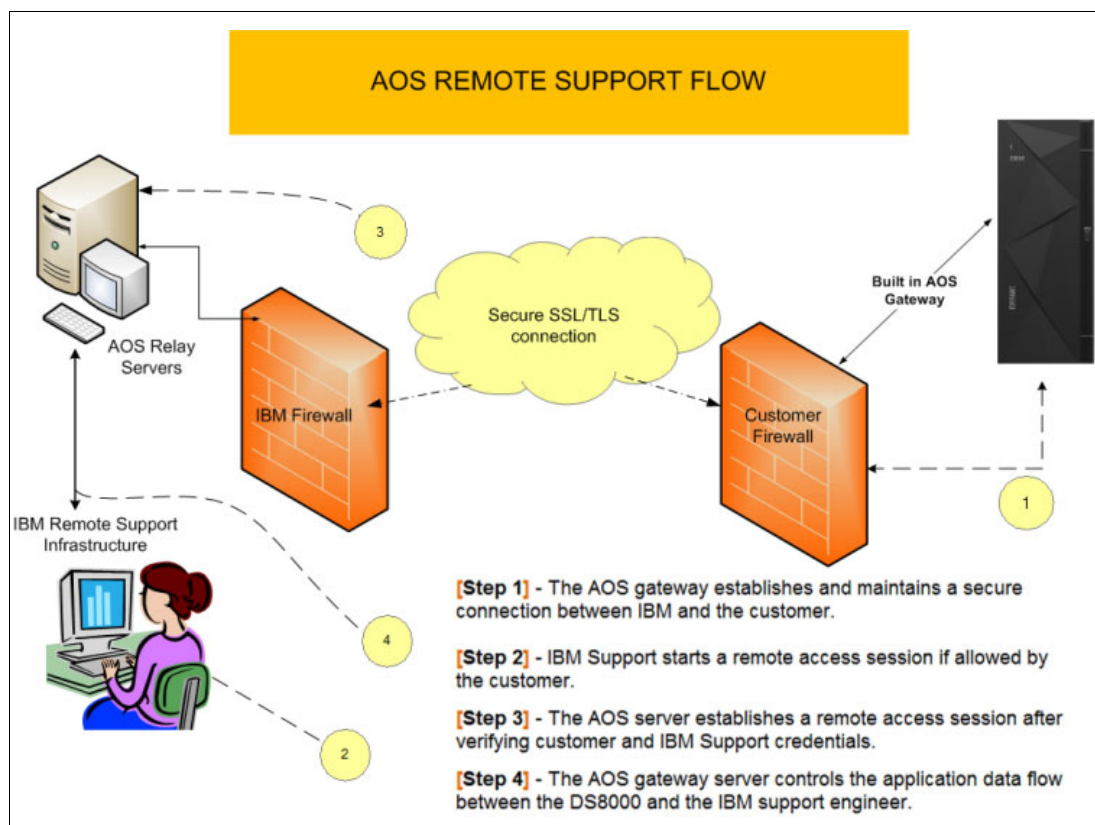

Figure 15-1   DS8880 AOS connection

15.5.2 DS8880 embedded AOS

AOS is an embedded feature, on DS8700, DS8800, DS8870, and DS8880. The AOS software package is preinstalled and customized on the Management Console. This technique eliminates the need to provide an additional system to operate an AOS Gateway. Embedded AOS is a secure, fast, broadband form of remote access. You can choose to allow unattended or attended remote access sessions. If you select attended remote access sessions, IBM Support contacts you or the storage operator to start the support session through DS CLI or the DS GUI.

The IBM SSR configures AOS during the installation or a later point in time by entering information that is provided in the inbound remote support worksheet. The worksheets can be found in the Installation and Planning Guide, or online in the Planning Section of the IBM DS8880 Knowledge Center at:

http://www.ibm.com/support/knowledgecenter/ST5GLJ_8.2.0/com.ibm.storage.ssic.help.doc/f2c_custmztnwkshts_1xznxr.html

In addition, your firewall needs to allow outbound traffic from the HMC to the AOS Infrastructure. The inbound remote support worksheet provides information about the required firewall changes.

For more information about AOS, see *IBM Assist On-site for Storage Overview*, REDP-4889.

15.5.3 DS8880 Remote Support Center

The HMC has been made Remote Support Center (RSC) ready starting with Microcode release R8.1. RSC uses SSH instead of TLS as used with Assist On-site.

15.5.4 Support access management through the DS CLI and the Storage Management GUI

All support connections can be enabled or disabled through the DS Storage Manager GUI or DSCLI. The following interfaces can be controlled:

► The web-based user interface for the IBM SSR on the HMC
► The SSH command-line interface access through the local or internal network
► The remote access through Assist On-site

Using the DS Storage Manager GUI to manage the service access

You can control the all service access through the DS Storage Manager GUI through the Access window, which can be opened by clicking **Settings** → **Support Menu**. Figure 15-2 shows the example of the Access window.
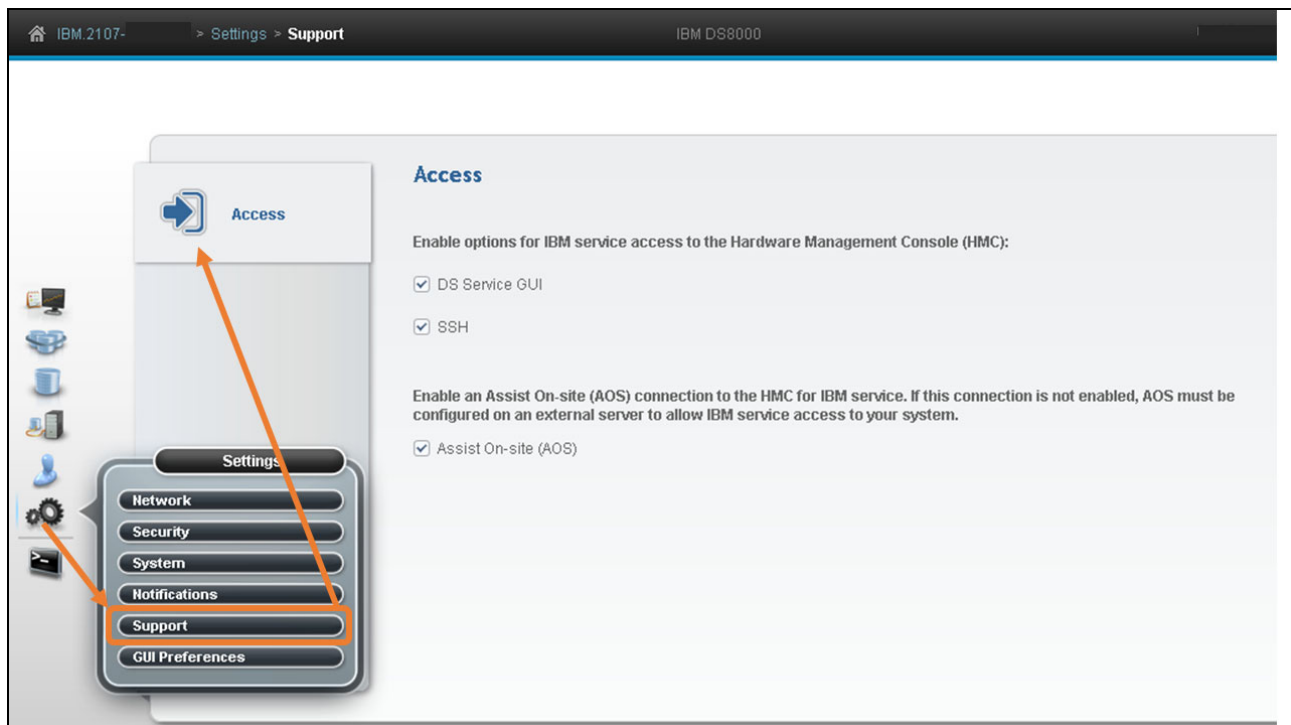

Figure 15-2   Control the Support access through the DS Storage Manager GUI

Using DSCLI to manage service access

You are able to manage the service access to the DS8880 by using DS CLI commands. The following user access security commands are available:

► `manageaccess`: This command manages the security protocol access settings of a Management Console for all communications to and from the DS8000 system. You can also use the `manageaccess` command to start or stop outbound virtual private network (VPN) connections instead of using the `setvpn` command.

► **chaccess**: The **chaccess** command changes one or more access settings of an HMC. Only users with administrator authority can access this command. See the command output in Example 15-1.

```
chaccess [-commandline enable | disable] [-wui enable | disable] [-modem enable
| disable] [-aos enable | disable] hmc1 | hmc2
```

Example 15-1   Output of chaccess command

```
Invoking the chaccess command
dscli> chaccess -cmdline enable -wui enable -hmc 1
The resulting output
hmc1 successfully modified.
```

> **Note:** With the release of the DS8880, VPN and modem support are no longer offered. The DS CLI retains the commands for compatibility with earlier versions.

► **lsaccess**: The **lsaccess** command displays the access settings and VPN status of the primary and backup Management Consoles:

```
lsaccess [hmc1 | hmc2]
```

See the output in Example 15-2.

Example 15-2   The lsaccess command output for a system with only one Management Console

```
dscli> lsaccess -hmc all -l
hmc   cmdline wui     modem cim     aos     vpn
======================================================
hmc1 enabled enabled -    disabled enabled disabled
dscli>
```

> **Important:** The hmc1 value specifies the primary HMC, and the hmc2 value specifies the secondary HMC, regardless of how **-hmc1** and **-hmc2** were specified during **dscli** start. A DS CLI connection might succeed even if a user inadvertently specifies a primary HMC by using **-hmc2** and the secondary backup HMC by using **-hmc1** at the DS CLI start.

Client notification of remote login

The Management Console code records all remote access in a log file. A client can use a DS CLI function to offload this file for audit purposes. The DS CLI function combines the log file that contains all service login information with an IBM enterprise storage server network interface (ESSNI) server audit log file that contains all client user login information to provide the client with a complete audit trail of remote access to a Management Console.

This on-demand audit log mechanism is sufficient for client security requirements about HMC remote access notification.

In addition to the audit log, email notifications and SNMP traps also can be configured at the Management Console to send notification in a remote support connection.

15.6 Audit logging

The DS8880 offers an audit log, It is an unalterable record of all actions and commands that were initiated by users on the storage system through the DS8000 Storage Management graphical user interface (GUI), DS CLI, DS Network Interface (DSNI), or Copy Service Manager. An audit log does not include commands that were received from host systems or actions that were completed automatically by the storage system. The audit logs can be exported and downloaded by the DS CLI or Storage Management GUI.

The DS CLI **offloadauditlog** command provides clients with the ability to offload the audit logs to the client's DS CLI workstation into a directory of their choice, as shown in Example 15-3.

Example 15-3   DS CLI command to download audit logs

```
dscli> offloadauditlog -logaddr smc1 c:\75ZA570_audit.txt
Date/Time: November 3, 2015 11:41:56 AM CET IBM DSCLI Version: 7.8.0.376 DS: -
CMUC00243I offloadauditlog: Audit log was successfully offloaded from smc1 to
c:\75ZA570_audit.txt.
```

The audit log can be exported by using the DS8000 Storage Management GUI on the Events window by clicking the **Diskette** icon and then selecting **Export Audit Log**, as shown in Figure 15-3.
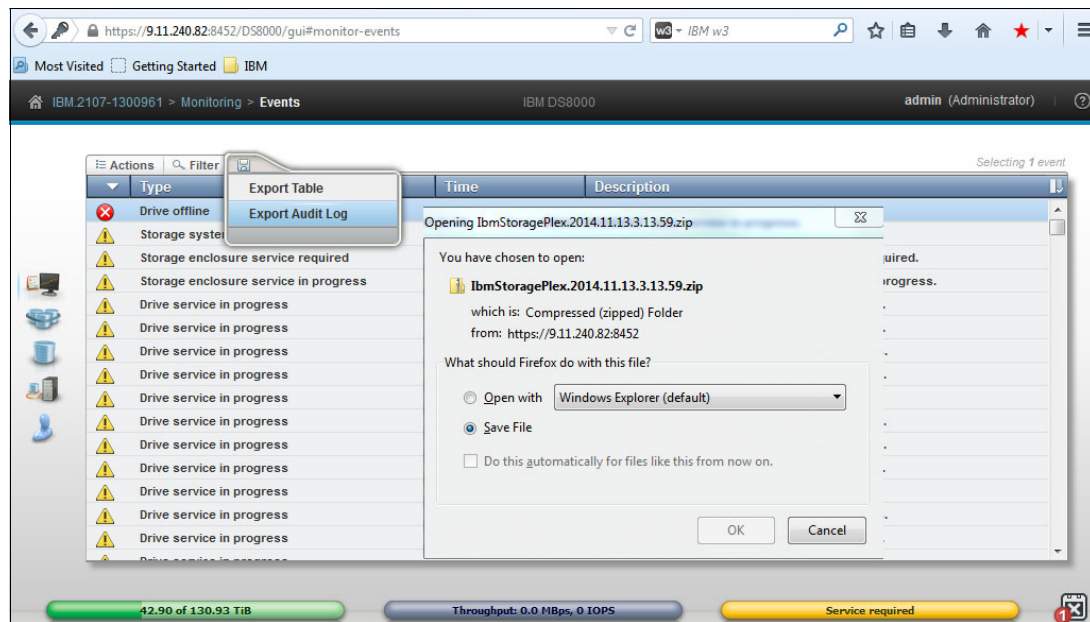

Figure 15-3   Export Audit Log

The downloaded audit log is a text file that provides information about when a remote access session started and ended, and the remote authority level that was applied. A portion of the downloaded file is shown in Example 15-4.

Example 15-4   Audit log entries that relate to a remote support event

```
MST,,1,IBM.2107-75ZA570,N,8036,Authority_to_root,Challenge Key = 'Fy31@C37';
Authority_upgrade_to_root,,,
U,2015/10/02 12:09:49:000
MST,customer,1,IBM.2107-75ZA570,N,8020,WUI_session_started,,,,
U,2015/10/02 13:35:30:000
MST,customer,1,IBM.2107-75ZA570,N,8022,WUI_session_logoff,WUI_session_ended_logged
off,,,
```

The *Challenge Key* that is presented to the IBM support representative is a part of a two-factor authentication method that is enforced on the Management Console. It is a token that is shown to the IBM SSR who connects to the DS8880. The representative must use the Challenge Key in an IBM internal system to generate a *Response Key* that is given to the HMC. The Response Key acts as a one-time authorization to the features of the HMC. The Challenge and Response Keys change when a remote connection is made.

The Challenge-Response process must be repeated if the representative needs higher privileges to access the Management Console command-line environment. No direct user login and no root login are on a DS8880.

Entries are added to the audit file only after the operation completes. All information about the request and its completion status is known. A single entry is used to log request and response information. It is possible, though unlikely, that an operation does not complete because of an operation timeout. In this case, no entry is made in the log.

The audit log entry includes the following information:

► Log users that connect or disconnect to the storage manager.

► Log user password and user access violations.

► Log commands that create, remove, or modify the logical configuration, including the command parameters and user ID.

► Log commands that modify storage facility image (SFI) and Storage Facility settings, including the command parameters and user ID.

► Log Copy Services commands, including command parameters and users.

> **Note:** IBM Copy Services Manager commands are not supported.

Audit logs are automatically trimmed (first-in first-out (FIFO)) by the subsystem so that they do not use more than 50 MB of disk storage.

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.

► *DB2 for z/OS and List Prefetch Optimizer*, REDP-4862:

  http://www.redbooks.ibm.com/abstracts/redp4862.html

► *DS8000 Copy Services*, SG24-8367:

  http://www.redbooks.ibm.com/abstracts/sg248367.html

► *DS8000 I/O Priority Manager*, REDP-4760:

  http://www.redbooks.ibm.com/abstracts/redp4760.html

► *DS8870 Easy Tier Application*, REDP-5014:

  http://www.redbooks.ibm.com/abstracts/redp5014.html

► *DS8880 High Performance Flash Enclosure Gen2*, REDP-5422:

  http://www.redbooks.ibm.com/abstracts/redp5422.html

► *Exploring the DS8870 RESTful API Implementation*, REDP-5187

  http://www.redbooks.ibm.com/abstracts/redp5187.html

► *Get More Out of Your IT Infrastructure with IBM z13 I/O Enhancements*, REDP-5134:

  http://www.redbooks.ibm.com/abstracts/redp5134.html

► *IBM Assist On-site for Storage Overview,* REDP-4889

  http://www.redbooks.ibm.com/abstracts/redp4889.html

- *IBM DS8000 Easy Tier*, REDP-4667:

  http://www.redbooks.ibm.com/abstracts/redp4667.html

- *DS8870 Data Migration Techniques,* SG24-8257

  http://www.redbooks.ibm.com/abstracts/sg248257.html

- *IBM DS8870 Easy Tier Heat Map Transfer,* REDP-5015:

  http://www.redbooks.ibm.com/abstracts/redp5015.html

- *IBM DS8870 and NIST SP 800-131a Compliance*, REDP-5069:

  http://www.redbooks.ibm.com/abstracts/redp5069.html

- *IBM DS8870 and VMware Synergy,* REDP-4915:

  http://www.redbooks.ibm.com/abstracts/redp4915.html

- *IBM DS8870 Multiple Target Peer-to-Peer Remote Copy*, REDP-5151:

  http://www.redbooks.ibm.com/abstracts/redp5151.html

- *IBM DS8870 and IBM z Systems Synergy*, REDP-5186:

  http://www.redbooks.ibm.com/abstracts/redp5186.html

- *IBM DS8880 Data-at-rest Encryption*, REDP-4500:

  http://www.redbooks.ibm.com/abstracts/redp4500.html

- *IBM DS8880 Integrated Copy Services Manager and LDAP Client on the HMC,* REDP-5356:

  http://www.redbooks.ibm.com/abstracts/redp5356.html

- *IBM DS8880 Product Guide (Release 8.2),* REDP-5344:

  http://www.redbooks.ibm.com/abstracts/redp5344.html

- *IBM DS8880 Thin Provisioning,* REDP-5343:

  http://www.redbooks.ibm.com/abstracts/redp5343.html

- *IBM GDPS Family: An Introduction to Concepts and Capabilities*, SG24-6374:

  http://www.redbooks.ibm.com/abstracts/sg246374.html

- *IBM Power System E850 Technical Overview and Introduction*, REDP-5222:

  http://www.redbooks.ibm.com/abstracts/redp5222.html

- *IBM Power System S822 Technical Overview and Introduction*, REDP-5102:

  http://www.redbooks.ibm.com/abstracts/redp5102.html

- *IBM Power Systems S814 and S824 Technical Overview and Introduction,* REDP-5097:

  http://www.redbooks.ibm.com/abstracts/redp5097.html

- *IBM System Storage DS8000 Copy Services Scope Management and Resource Groups*, REDP-4758:

  http://www.redbooks.ibm.com/abstracts/redp4758.html

- *IBM System Storage DS8000 Performance Monitoring and Tuning, SG24-8318*:

  http://www.redbooks.ibm.com/abstracts/sg248318.html

- *IBM System Storage DS8000: Remote Pair FlashCopy (Preserve Mirror)*, REDP-4504:

  http://www.redbooks.ibm.com/abstracts/redp4504.html

- *IBM System Storage DS8000: Host Attachment and Interoperability*, SG24-8887:

  http://www.redbooks.ibm.com/abstracts/sg248887.html

- *IBM System Storage DS8000: z/OS Distributed Data Backup*, REDP-4701:

  http://www.redbooks.ibm.com/abstracts/redp4701.html

- *Using IBM DS8870 in an OpenStack Environment*, REDP-5220

  http://www.redbooks.ibm.com/abstracts/redp5220.html

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

Other publications

These publications are also relevant as further information sources:

- *IBM DS8880 Introduction and Planning Guide*, GC27-8525

  https://www.ibm.com/support/docview.wss?uid=ssg1S7005228

- *IBM DS8000 Series Command-Line Interface User's Guide*, SC27-8526

- *IBM Storage Driver for OpenStack Version 1.8.0 User Guide*, SC27-4233

- *IBM DS Open Application Programming Interface Reference,* GC27-4211

  https://www.ibm.com/support/docview.wss?uid=ssg1S7001160

- *IBM DS8000 Host Systems Attachment Guide*, SC27-8527

  https://www.ibm.com/support/docview.wss?uid=ssg1S7005237

- *IBM System Storage Multipath Subsystem Device Driver User's Guide,* GC52-1309

- *IBM Security Key Lifecycle Manager Installation and Configuration Guide*, SC27-5335

- *IBM Security Key Lifecycle Manager Quick Start Guide*, GI13-2316

- "AMP: Adaptive Multi-stream Prefetching in a Shared Cache" by B. Gill

- "Optimal Multistream Sequential Prefetching in a Shared Cache" by B. Gill, et al.

- "Outperforming LRU with an adaptive replacement cache algorithm" by N. Megiddo, et al., in IEEE Computer, volume 37, number 4, pages 58 - 65, 2004

- "SARC: Sequential Prefetching in Adaptive Replacement Cache" by Binny Gill, et al., in "Proceedings of the USENIX 2005 Annual Technical Conference", pages 293 - 308

Online resources

These websites are also relevant as further information sources:

- DS8000 System Storage Interoperation Center (SSIC):

  https://www.ibm.com/systems/support/storage/ssic/

- IBM Support: Fix Central:

  https://www.ibm.com/support/fixcentral/swg/selectFixes?product=ibm/Storage_Disk/DS8880

► DS8000 IBM Knowledge Center:

https://www.ibm.com/support/knowledgecenter/ST5GLJ/ds8000_kcwelcome.html

► DS8000 Series Copy Services Fibre Channel Extension Support Matrix:

http://www.ibm.com/support/docview.wss?uid=ssg1S7003277

Help from IBM

IBM Support and downloads

ibm.com/support

IBM Global Services

ibm.com/services

IBM Training and Skills

ibm.com/training

IBM DS8880 Architecture and Implementation (Release 8.2.1)

Redbooks

SG24-8323-02

ISBN 0738442402

Printed in U.S.A.

ibm.com/redbooks