

IBM StoredIQ Data Sources and Volumes

Genifer Graff



 Analytics

ECM



IBM StoredIQ Data Sources And Volumes

Data sources and volumes are an important part of the IBM® StoredIQ® solution and an understanding of the data sources is required to effectively use StoredIQ.

This IBM Redpaper™ publication describes the data sources supported by the IBM StoredIQ Platform. Many data sources exist and the ability to provide insight into the content located within the supported data sources is an important benefit of StoredIQ.

Through this paper, you can learn about data sources and volumes and how to create StoredIQ volumes. You can also learn about configuring StoredIQ volumes and the options for configuring StoredIQ to index the content that is located in the data source.

As a reminder, the goal of StoredIQ is to create an information set (infoset) that can be used to provide insight into an organization's unstructured content at a given point in time. Infosets are derived from data sources that are harvested (indexed) by StoredIQ on the StoredIQ data server. After the data sources are known, an administrator defines the volumes that are associated with the data sources.

This paper covers the following topics:

- ▶ What a data source is
- ▶ What a volume is
- ▶ Configuring the server platforms
- ▶ Adding a volume
- ▶ Deleting a volume

Data sources

A data source is a location that contains unstructured content. Data sources include email servers, file shares, Box (enterprise file synchronization and sharing), social media, and enterprise content management systems. The unstructured content and images that exist at the data source are harvested by StoredIQ into an infoset. No content is moved or deleted in the creation of the harvest (index) of the data source.

After the infoset is created, the StoredIQ Data Workbench is used to provide insight to the information contained in the data sources. That insight into your data sources, without altering or moving content, is one of the benefits of the StoredIQ solution.

StoredIQ has a flexible and scalable architecture. This means that data sources can be located at different network addresses or even different physical locations.

Supported data sources

StoredIQ connects to many different data sources. Some of the data sources include Box, IBM FileNet®, Microsoft Exchange, and CMIS 1.0. IBM frequently adds new data sources to StoredIQ and any product documentation might become obsolete. For this reason, be sure to check the following website for the latest list of data sources:

http://www.ibm.com/support/knowledgecenter/SSSHEC_7.6.0/admin/ref/ref_supportedserverplatformsandprotocols.dita

StoredIQ can perform read, write, and delete operations on a data source. However, not all operations are available for all data sources. For example, StoredIQ might read from and write to a Box data source, but StoredIQ cannot delete content from a Box data source.

Table 1 lists the operations that can be performed on each data source and the supported version of the software data source, if applicable.

Table 1 Supported data sources

Platform or protocol	Versions	Operation type			Notes
		Read	Write	Delete	
Box	N/A	R	W	N/A	Box volumes are added through IBM StoredIQ Administrator.
CIFS	Windows XP, Windows Vista, Windows 7, Windows Server 2003, 2008, 2008 R2, Samba, Mac OS X, 10.7, 10.8, 10.9	R	W	D	
CMIS	1.0	R	W	N/A	The CMIS option references a solution that involves different products. As such, there is no specific supported version to cite.
EMC Centera	Centera API v3.1	R	W	D	Metadata and content written by IBM StoredIQ or EMC EmailXtender is supported. EmailXtender support is limited to email archived from Exchange.
EMC Documentum	Documentum 6.0, 6.5, and 6.7 Documentum with Retention Policy Services (RPS) 6.0, 6.5	R	W	D	Only Current documents in standard cabinets are indexed. Customer must supply DFC files to enable the connector.

Platform or protocol	Versions	Operation type			Notes
		Read	Write	Delete	
Hitachi HCAP	Hitachi Content Platform V1.8	R	W	D	
IBM Content Manager	IBM Content Manager 8.4.3 and later with the following platforms: <ul style="list-style-type: none"> ▶ IBM AIX® 5L 5.3, 6.1, 7.1 ▶ Red Hat Enterprise Linux 4.0, 5.0 ▶ SUSE Linux Enterprise Server 9, 10, 11 ▶ Solaris 9, 10 ▶ Windows Server 2003, 2008, 2008 R2 	N/A	N/A	N/A	
IBM Domino®	Domino/Notes 6.x, 7.x, 8.x	R	N/A	N/A	Email only. Email is converted to the .MSG format for processing.
IBM FileNet	FileNet 5.3	R	W	N/A	
Jive	Jive 5.0.2	R	N/A	N/A	
Microsoft Exchange	Exchange 2003	R	N/A	N/A	Messages (email), Contacts, Calendar items, Notes, Tasks, and Documents. WebDAV protocol.
Microsoft Exchange	Exchange 2007, 2010, 2013, Online	N/A	N/A	N/A	Messages (email), Contacts, Calendar items, Notes, Tasks, and Documents. Exchange web service interface.
Microsoft SharePoint	SharePoint 2003	R	W	D	WebDAV protocol.
Microsoft SharePoint	SharePoint 2007, 2010, 2013	N/A	N/A	N/A	All document versions are optional. SharePoint 2010 and 2013 supported list types: User profiles, User notes, Blog Post, Blog Comment, Discussion Post, Discussion Reply, Wiki Page, Calendar, Task/Project Task, Contact, Issue Tracker, Survey, Link, and Announcements. Content of custom lists is indexed generically as text. It is not modeled specifically, like standard list types. SharePoint web service interface.
NFS		R	W	D	
OpenText Livelink and Content Server	OpenText Livelink Enterprise Server 9.7, 9.7.1 Content Server 10.0.0	R	N/A	N/A	Connector is based on IBM Content Integrator (ICI) version 8.6.
Salesforce Chatter	Salesforce Partner API v26.0	R	N/A	N/A	

Data volume

A *volume* represents a data source or destination that is available in the network to the IBM StoredIQ Platform. Only a system administrator can define, configure, add, or remove volumes from an infoSet using the IBM StoredIQ Platform Administrator. Volumes behave the same way as hard disk drive partitions. An administrator can set up separate volumes that originate from the same server or across many servers.

Volume types

A volume is a data source or destination that is available in the network to be accessed by the StoredIQ Platform. StoredIQ supports four types of volumes:

- ▶ Primary volume
- ▶ Retention volume
- ▶ Export volume
- ▶ System volume

Primary volume

A primary volume is a storage location that knowledge workers can access to create, read, update, and delete unstructured content.

The following list shows the supported solutions that can be a StoredIQ primary volume:

- ▶ Box (Box volumes are added through IBM StoredIQ Administrator.)
- ▶ CIFS
- ▶ CMIS 1.0
- ▶ EMC Documentum (Customer must supply Documentum Foundation Classes (DFC) files to enable the connector.)
- ▶ IBM Content Manager
- ▶ IBM Domino (Supported for email only.)
- ▶ IBM FileNet
- ▶ Jive
- ▶ Microsoft Exchange
- ▶ Microsoft SharePoint
- ▶ NFS
- ▶ OpenText Livelink and Content Server
- ▶ Salesforce Chatter

Retention volume

A StoredIQ retention volume is a storage location that enforces retention and hold policies on content. The following locations can be retention volumes:

- ▶ CIFS
- ▶ EMC Centera
- ▶ Hitachi HCAP
- ▶ NFS

Export volume

An export volume is an unmanaged (not indexed) storage location where content is copied along with metadata with an audit detail report. The audit information is stored in a format that can be imported into other applications such as eDiscovery review applications. The two platforms that can be used as export volumes are as follows:

- ▶ CIFS
- ▶ NFS

System volume

A system volume is a storage location where files can be written to and read by IBM StoredIQ. It can be used to export volume metadata that is contained in the index on a data server. Exported volume data can be imported from a system volume to populate a volume index. The following two platforms can be system volumes:

- ▶ CIFS
- ▶ NFS

Configuring the server platforms

Before you configure volumes on StoredIQ Platform, you must configure the server platforms that you use for the volume types. Each server type has requisite permissions and settings. Because IBM StoredIQ Platform supports several types of volumes, server platforms must be configured to support the volume types you want.

Configure Box

Box volumes can be added only from IBM StoredIQ Administrator and have several configuration prerequisites that must be met.

In Box volumes, the concepts of ownership and access differ from a typical computer file system. When data objects are copied to Box, they do not retain their original owner. Instead, ownership changes to the user performing the copy of those objects. Additionally, the folder hierarchy of the source volume can be reproduced, but it is rooted in the home folder of the user. And finally, when configuring a copy action with Box as the target set, a preferable practice is to either designate a destination directory or select **Recreate Directory Structure**.

In the Box application, select these Scopes options:

- ▶ Read and write all files and folders
- ▶ Manage an enterprise
- ▶ Manage an enterprise's managed users
- ▶ Manage an enterprise's groups
- ▶ Manage an enterprise's properties
- ▶ Manage an enterprise's retention policies

Configure exchange servers

This section describes the necessary steps to connect to Microsoft Exchange servers. When you configure Exchange servers, you must consider various connections and permissions:

- ▶ **Secure connection:** If you want to connect to Exchange volumes over HTTPS, you can either select **Use SSL** or add the port number 443 after the server name. If you choose the port option, consider this example: qa03exch2000.qaw2k.local:443. In some cases, this secure connection can result in some performance degradation with high data volume or connections. If you enter the volume information without the 443 suffix, the default connection is over HTTP.
- ▶ **Permissions for Exchange 2003:** The following permissions must be set on the Exchange server to the mailbox store or the mailboxes from which you harvest:
 - Read
 - Execute
 - Read permissions
 - List contents
 - Read properties
 - List object
 - Receive as
- ▶ **Permissions for Exchange Server 2007, 2010, and 2013:** The Full Access permissions must be set on the Exchange server to the mailbox store or the mailboxes from which you will perform the harvest.
- ▶ **Recoverable items in Exchange 2010 and 2013:** To harvest the recoverable items folders in Exchange 2010 and 2013, you must be logged in as an administrator role.
- ▶ **Deleted items:** To harvest items that were deleted from the Exchange server, enable Exchange's transport dumpster settings. For more information, see *Microsoft Exchange Server 2010 Administrator's Pocket Consultant*, ISBN-13: 978-0735627123.
- ▶ **Windows Authentication:** For all supported versions, enable Integrated Windows Authentication on each Exchange server.
- ▶ **Public folders:** To harvest public folders in Exchange, the Read Items privilege is required.
- ▶ **Personal archives:** To harvest personal archives in Exchange, the Application Impersonation and Mailbox Search roles are required.

Configure Microsoft SharePoint

Microsoft SharePoint presents a unique set of configuration options to a StoredIQ administrator. When a SharePoint administrator configures SharePoint, certain privileges are required by a user account along with IBM StoredIQ Platform recommendations.

StoredIQ supports many versions of Microsoft SharePoint: 2003, 2007, 2010 and 2013.

Additionally, SharePoint 2007 and 2010 require the configuration of alternate-access mappings to map IBM StoredIQ Platform requests to the correct web sites.

To configure SharePoint, consider the connections and privileges described in this section.

Secure connection

If you want to connect to SharePoint volumes over HTTPS, you can either select the **Use SSL** check box or can add port number 443 after the server name when you set up the volume on IBM StoredIQ Platform.

If you choose the port option, an example of this is qa01.company.com:443. If you enter the volume information without the 443 suffix, the default connection is over HTTP.

Privileges

To run policies on SharePoint servers, you must use credentials with Full Control privileges. Use a site collection administrator to harvest subsites of a site collection.

Privileges required by user account

IBM StoredIQ Platform is typically used with SharePoint for one of these instances: to harvest and treat SharePoint as a source for policy actions or to use as a destination for policy actions, which means that you can write content into SharePoint.

Consider these points about attributes:

- ▶ Attributes are not set or reset on a SharePoint harvest or if you copy from SharePoint.
- ▶ Attributes are set only if you copy to SharePoint.

You must consider the following situations:

- ▶ If you plan to read only from the SharePoint (harvest and source copies from), then you must use user credentials with Read privileges on the site and all of the lists and data objects that you expect to process.
- ▶ If you plan to use SharePoint as a destination for policies, you must use user credentials with Contribute privileges on the site.
- ▶ More Privileges for Social Data: If you want to index all the social data for a user profile in SharePoint 2010, the user credentials must own privileges to Manage Social Data too.
- ▶ Recommended Privileges: Use a site collection administrator to ensure that all data is harvested from a site or site collection.

Alternate-access mappings for SharePoint 2007 and 2010

Alternate-access mappings map URLs presented by IBM StoredIQ Platform to internal URLs received by Windows SharePoint Services. An alternate-access mapping is required between the server name and optional port that is defined in the SharePoint volume definition and the internal URL of the web application. If SSL is used to access the site, ensure that the alternate-access mapping URL uses `https://` as the protocol.

See the Microsoft SharePoint 2007 or 2010 documentation to configure alternate-access mappings. These mappings are based on the public URL that is configured by the local SharePoint administrator and used by the IBM StoredIQ Platform SharePoint volume definitions.

For example, you access a SharePoint volume with the fully qualified domain name, from the intranet zone, as in this example:

```
http://itweb.storediqexample.com
```

An alternate-access mapping for the public URL (`http://itweb.storediqexample.com`) for the intranet zone must be configured for the SharePoint 2007 or 2010 web application that hosts the site to be accessed by the volume definition. If you access the same volume with SSL, the mapping added must be for the `https://itweb.storediqexample.com` URL instead.

Note that when you configure SharePoint volumes with non-qualified names, you enter the URL for a SharePoint site collection or site that is used by IBM StoredIQ Platform in the volume definition.

Consider the following conditions:

- ▶ The URL must be valid about the Alternate Access Mappings that are configured in SharePoint.
- ▶ If the host name in the URL does not convey the fully qualified domain to authenticate the configured user, an Active Directory server must be specified. The specified Active Directory must be a fully qualified domain name and is used for authentication.

Configure NFS Servers

To harvest and run policies on Network File System (NFS) servers, you must enable root access on the NFS server that is connected to IBM StoredIQ Platform.

Configure Windows Share (CIFS)

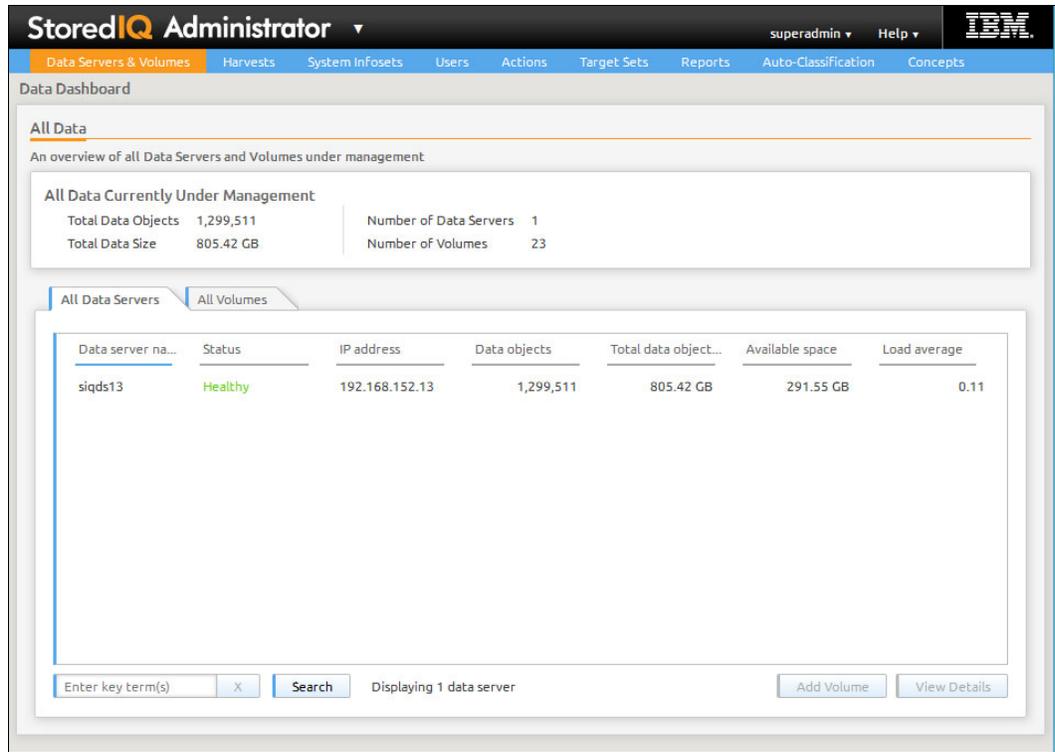
To harvest and run policies on Windows Share Common Internet File System (CIFS) servers, the StoredIQ user must be in the backup operator group on the Windows Share server that shows the shares on IBM StoredIQ Platform.

Adding a volume

Now that you have an understanding of supported data sources and the configuration options that must be made on some data sources, you can create volumes. First you create a primary volume. A *primary volume* serves as a primary data source in IBM StoredIQ Platform. You must have at least one primary volume within your configuration.

Use the following steps to create or add a volume:

1. Log in to the IBM StoredIQ Administrator (Figure 1).



The screenshot shows the IBM StoredIQ Administrator interface. At the top, there's a navigation bar with 'Data Servers & Volumes' selected. Below it, the 'Data Dashboard' section is visible. A summary box titled 'All Data Currently Under Management' shows: Total Data Objects: 1,299,511; Total Data Size: 805.42 GB; Number of Data Servers: 1; Number of Volumes: 23. Below this, there are two tabs: 'All Data Servers' (selected) and 'All Volumes'. A table lists the data servers:

Data server na...	Status	IP address	Data objects	Total data object...	Available space	Load average
siqds13	Healthy	192.168.152.13	1,299,511	805.42 GB	291.55 GB	0.11

At the bottom, there's a search bar with 'Enter key term(s)', a search button, and a status indicator 'Displaying 1 data server'. There are also 'Add Volume' and 'View Details' buttons.

Figure 1 StoredIQ Administrator Data Dashboard

2. The Data Dashboard opens (Figure 2 on page 10), which provides information about your StoredIQ environment. You see the data server and the status of the server. You also see all of the data objects under management of StoredIQ.

By selecting the **All Volumes** tab, you can see all volumes that have been defined. You might have many volumes defined or you might be creating your first primary volume. You can create the volume from this window.

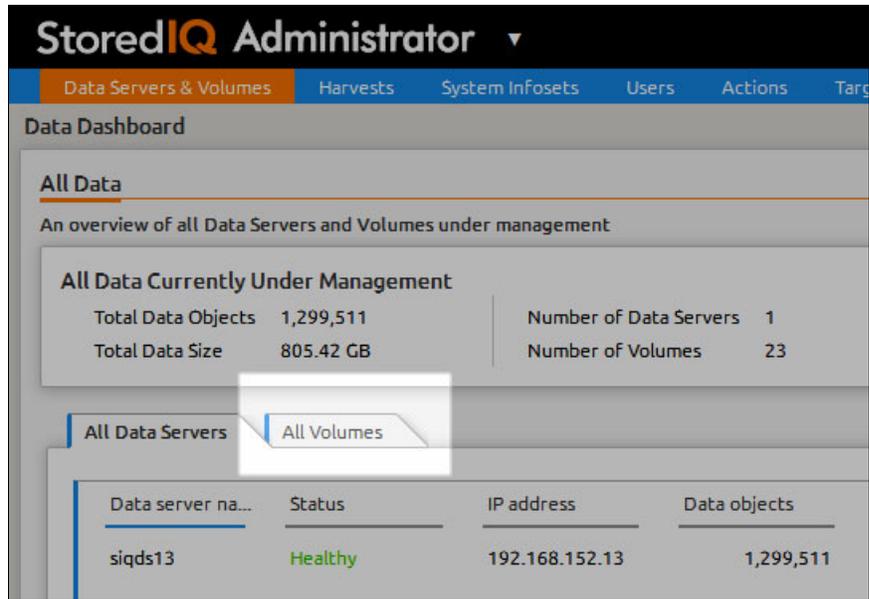


Figure 2 The All Volumes tab

3. To add a new volume, click **Add Volume** (Figure 3) at the lower right corner of the All Volumes tab.

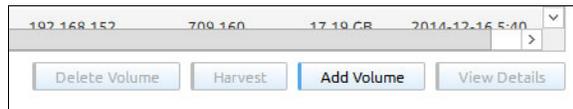


Figure 3 The Add a Volume button

4. The Add Volume dialog box is displayed. In this dialog box, you can choose the options for the volumes.

Remember, you must have a primary volume. The following sections contain the options for each of the primary volume types.

Add Box volume

With the Add Volume dialog window displayed (Figure 4), choose the source type of **Box**. This displays the options that must be set to create a Box primary volume.

Figure 4 The Add Volume window with the source type Box options

Table 2 lists the options and corresponding actions to add a Box volume.

Table 2 Box configuration options

Option	Action
Volume Type	Select Primary .
Source Type	Select Box .
Server	In the server text box, specify <code>api . box . com</code> as the Box server name.
Assign to Data Server	Select the data server that will be used for the index.
Volume Name	Specify a name for the volume.
Include Users	Select this option to scope the Box user accounts that will be included with this volume. Regular expressions are supported.
Include Metadata for contained objects	Select this option to include the metadata for contained objects.
Include content tagging for Full-Text index	Select this option to perform a full-text index.

Add CIFS volume

With the Add Volume dialog window displayed (Figure 5), choose the source type of **CIFS**. The options that must be set to create a CIFS primary volume are displayed.

The screenshot shows the 'Add Volume' dialog box with the following configuration:

- Volume Type:** Primary
- Source Type:** CIFS
- Assign To Data Server:** siqds13
- Server:** domain.name
- Username:** Enter the admin account
- Password:** Enter the proper password
- Volume Name:** Enter a name for the volume
- Share:** Enter share name
- Initial Directory:** Enter initial directory
- Indexing Options:**
 - Include metadata for contained objects
 - Include content tagging and Full-Text Index

Figure 5 The Add Volume window with the source type CIFS options

Table 3 lists the options and corresponding actions to add a Primary CIFS volume.

Table 3 CIFS configuration options

Option	Action
Volume Type	Select Primary .
Source Type	Select CIFS .
Server	Enter the fully qualified domain name of the server from which the volume is available for mounting.
Username	Enter the user name that is used to connect to the data source and mount the volume.
Password	Enter the password for the user.
Assign to Data Server	Select the data server that will be used for the index.
Volume Name	Specify a name for the volume.
Share	Enter the file share name of this volume.
Initial Directory	Enter the name of the initial directory from which the harvest will begin.
Include Metadata for contained objects	Select this option to include the metadata for contained objects.
Include content tagging for Full-Text index	Select this option to perform a full-text index.

Add CMIS volume

Content Management Interoperability Services (CMIS) is an open standard that allows for different content management systems. Use this option to connect to a content management system that supports CMIS but is not supported with another connector such as FileNet or Documentum.

With the Add Volume dialog window displayed (Figure 6), select the source type of **CMIS**. The options that must be set to create a CIFS primary volume are displayed.

Figure 6 The Add Volume window with the source type CMIS options

Table 4 lists the options and corresponding actions to add a CMIS volume.

Table 4 CMIS configuration options

Option	Action
Volume Type	Select Primary .
Source Type	Select CMIS .
Server	Enter the fully qualified domain name of the server from which the volume is available for mounting.
Port	Enter the Port number that will be used to connect to the system that supports CMIS.
Username	Enter the user name that is used to connect to the data source and mount the volume.
Password	Enter the password for the user.
Assign to Data Server	Select the Data Server that will be used for the index.
Volume Name	Enter a name for the volume.

Option	Action
Use SSL	Use the Secure Sockets Layer (SSL) standard security technology to establish an encrypted link between a web server and the browser.
Service	Enter the service name
Repository	Enter the name of the CMIS repository.
Include Metadata for contained objects	Select this option to include the metadata for contained objects.
Include content tagging for Full-Text index	Select this option to perform a full-text index.

Add NFS volume

A Network File System (NFS) is a distributed file system protocol that allows a server to share directories and files with clients over a network. NFS can also be a primary volume location for StoredIQ.

Figure 7 shows the Add Volume dialog window for adding NFS source type volumes.

Figure 7 The Add Volume window with the source type NFS options

Table 5 lists the options and corresponding actions to add an NFS volume.

Table 5 NFS configuration options

Option	Action
Volume Type	Select Primary .
Source Type	Select NFS .
Server	Enter the fully qualified domain name of the server from which the volume is available for mounting.
Assign to Data Server	Select the data server that will be used for the index.

Option	Action
Volume Name	Specify a name for the volume.
Export	Specify an export name for this volume.
Initial Directory	Enter the name of the initial directory from which the harvest must begin.
Include Metadata for contained objects	Select this option to include the metadata for contained objects.
Include content tagging for Full-Text index	Select this option to perform a full-text index.

Add Microsoft SharePoint volume

SharePoint is a web application platform for performing enterprise content management functions in the Microsoft Office server suite. The web application allows users to create team sites from a template and perform collaboration functions with content.

Figure 8 shows the Add Volume dialog window for adding a Microsoft SharePoint volume.

Figure 8 The Add Volume window with the source type SharePoint options

Table 6 lists the options and corresponding actions to add a Microsoft SharePoint primary volume.

Table 6 SharePoint configuration options

Option	Action
Volume Type	Select Primary .
Source Type	Select SharePoint .
Server	Enter the fully qualified domain name of the server from which the volume is available for mounting.
Username	Enter the user name that is used to connect to the data source and mount the volume.
Password	Enter the password for the user.

Option	Action
Assign to Data Server	Select the Data Server that will be used for the index.
Volume Name	Specify a name for the volume.
Site URL	Specify the site URL of the SharePoint Server.
Initial Directory	Specify the name of the initial directory from which the harvest must begin. To recurse subsites and include versioning, SharePoint volumes must be added in the IBM StoredIQ Data Server user interface. SharePoint volumes added on the application stack can be edited on the IBM StoredIQ Data Server user interface to include those options.
Include Metadata for contained objects	Select this option to include the metadata for contained objects
Include content tagging for Full-Text index	Select this option to perform a full-text index

Add Exchange volume

Microsoft Exchange Server is a calendaring and mail server developed by Microsoft and runs exclusively on the Microsoft Windows Server product line. This solution is used by many organizations as their email platform.

Figure 9 shows the Add Volume dialog windows for adding an Exchange volume.

The screenshot shows the 'Add Volume' dialog window with the following fields and options:

- Volume Type:** Primary (dropdown)
- Source Type:** Exchange (dropdown)
- Server:** domain.name (text input)
- Username:** Enter the admin account (text input)
- Password:** Enter the proper password (text input)
- Assign To Data Server:** siqds13 (dropdown)
- Volume Name:** Enter a name for the volume (text input)
- Server Version:** 2000/2003 (dropdown)
- Mailbox Server:** Enter a mailbox server name (text input)
- Active Directory Server:** Enter an active directory server name (text input)
- Use SSL
- Initial Directory:** Enter initial directory (text input)
- Virtual Root:** exchange (text input)
- Indexing Options:**
 - Include metadata for contained objects
 - Include content tagging and Full-Text Index

Figure 9 The Add Volume window with the source type Exchange options

Table 7 lists the options and corresponding actions to add a Microsoft Exchange server as a primary volume.

Table 7 Exchange configuration options

Option	Action
Volume Type	Select Primary .
Source Type	Select Exchange .
Server	Enter the fully qualified domain name of the server from which the volume is available for mounting.
Username	Enter the user name that is used to connect to the data source and mount the volume
Password	Enter the password for the user
Assign to Data Server	Select a Data Server that will be used for the index.
Volume Name	Specify a name for the volume.
Server Version	Select the version of Microsoft Exchange: 2000/2003, 2007, 2010/2013, or Exchange Online.
Mailbox Server	Enter the name (or names) of the mailbox server (or servers), separated by commas. This option is not available if Exchange Online is selected as the Server Version.
Active Directory Server	Enter the name of the Active Directory server. This option is not available if Exchange Online is selected as the Server Version.
Use SSL	Use the SSL (Secure Sockets Layer) standard security technology to establish an encrypted link between a web server and the browser. This option is not available if Exchange Online is selected as the Server Version.
Initial Directory	Enter the name of the initial directory from which the harvest must begin.
Virtual Root	The name defaults to the correct endpoint for the selected Exchange version.
Include Metadata for contained objects	Select this option to include the metadata for contained objects.
Include content tagging for Full-Text index	Select this option to perform a full-text index.

Add FileNet volume

IBM FileNet is a robust Enterprise Content Management solution to allow users to perform content management such as version control and document approval. This is an IBM solution.

Figure 10 on page 18 shows the Add Volume dialog window for adding a FileNet volume.

Figure 10 The Add Volume window with the source type IBM FileNet Content Manager options

Table 8 lists the options and corresponding actions to add an IBM FileNet Content Manager repository as a primary volume.

Table 8 FileNet Configuration options

Option	Action
Volume Type	Select Primary .
Source Type	Select FileNet .
Server	Enter the fully qualified domain name of the server from which the volume is available for mounting.
Port	Enter the port number that will be used to connect to the FileNet content repository.
Username	Enter the user name that is used to connect to the data source and mount the volume.
Password	Enter the password for the user.
Assign to Data Server	Select the Data Server that will be used for the index.
Volume Name	Specify a name for the volume.

Option	Action
Object Store	Enter the name of the FileNet object store that will be used as the primary volume.
Connection Type	Specify the connection type of HTTP or HTTPS depending on how your FileNet system was configured.
Path	Specify the directory path.
Stanza	Specify the appropriate stanza.
Scope	Enter the appropriate SQL “where” clause.
Include Metadata for contained objects	Select this option to include the metadata for contained objects.
Include content tagging for Full-Text index	Select this option to perform a full-text index.

Add IBM Content Manager volume

IBM Content Manager is an Enterprise Content Management solution to allow users to perform content management such as version control. This solution helps users manage all types of content, such as document images, electronic office documents, XML, audio, and video.

Figure 11 shows the Add Volume dialog window for adding an IBM Content Manager volume.

Figure 11 The Add Volume window with the source type IBM Content Manager options

Table 9 lists the options and corresponding actions to add IBM Content Manager as a primary volume.

Table 9 IBM Content Manager configuration options

Option	Action
Volume Type	Select Primary .
Source Type	Select IBM Content Manager .
Server	Enter the fully qualified domain name of the server from which the volume is available for mounting.
Port	Enter the port number that will be used to connect to the IBM Content Manager repository.
Username	Enter the user name that is used to connect to the data source and mount the volume
Password	Enter the password for the user.
Connection String	(Optional) Enter additional connection-string parameters.
Assign to Data Server	Select the Data Server that will be used for the index.
Volume Name	Specify a name for the volume.
Repository	Enter the name of the library server database.
Server Type	Select the type of server associated with the volume: IBM DB2® or Oracle. By default, DB2 is selected.
Schema	Enter the schema for this library server database.
Remote Database	Enter the name of the remote database.
Include Metadata for contained objects	Select this option to include the metadata for contained objects.
Include content tagging for Full-Text index	Select this option to perform a full text index.

Add OpenText Content Server (Livelink) volume

OpenText Content Server is the foundational repository and information governance backbone for a broad range of OpenText ECM products. In 2012, OpenText changed the name of the product from Livelink ECM to OpenText Content Server. The technology component, formerly known as Livelink ECM, became known as OpenText Content Server.

Figure 12 shows the Add Volume dialog window for adding an OpenText Content Server (Livelink) volume.

Figure 12 The Add Volume window with the source type Livelink options

Table 9 on page 20 lists the options and corresponding actions to add an OpenText Content Server volume.

Table 10 OpenText Livelink configuration options

Option	Action
Volume Type	Select Primary .
Source Type	Select Livelink .
Server	Enter the fully qualified domain name of the server from which the volume is available for mounting.
Port	Enter the port number that will be used to connect.
Username	Enter the user name that is used to connect to the data source and mount the volume.
Password	Enter the password for the user.
Assign to Data Server	Select the data server that will be used for the index.
Volume Name	Specify a name for the volume.

Option	Action
Database	Enter the database name.
Search Slice	Enter the search slice.
Initial Directory	Enter the name of the initial directory from which the harvest must begin.
Include Metadata for contained objects	Select this option to include the metadata for contained objects.
Include content tagging for Full-Text index	Select this option to perform a full-text index.

Add NewsGator volume

NewsGator is RSS software that allows users to receive notifications when a blog, news site or other page has been updated. NewsGator can be a primary volume as read-only.

Figure 13 shows the Add Volume dialog window for adding a NewsGator volume.

Figure 13 The Add Volume window with the source type NewsGator options

Table 11 lists the options and corresponding actions to add NewsGator as a primary volume.

Table 11 NewsGator configuration options

Option	Action
Volume Type	Select Primary .
Source Type	Select NewsGator .
Server	In the server text box, enter the fully qualified domain name of the server from which the volume is available for mounting.
Username	Enter the user name that is used to connect to the data source and mount the volume.

Option	Action
Password	Enter the password for the user.
Assign to Data Server	Select the Data Server that will be used for the index.
Volume Name	Specify a name for the volume.
Initial Directory	Enter the name of the initial directory from which the harvest must begin.
Include Metadata for contained objects	Select this option to include the metadata for contained objects.
Include content tagging for Full-Text index	Select this option to perform a full-text index.

Add Retention volume

A StoredIQ retention volume is immutable storage that enforces retention and hold policies. Although data is under management, it cannot be modified or deleted, and knowledge workers typically do not access retention storage directly. This storage typically has its own special API and protocol, although NAS vendors implemented retention and hold features that use standard CIFS and NFS protocols (NetApp SnapLock or EMC Celerra FLR are specific examples).

The storage platform typically does not implement a hierarchical namespace to store content, but instead relies on a globally unique identifier as a handle to metadata and content. Applications can write metadata and binary content in any internal format to satisfy their requirements.

Typically, IBM StoredIQ does not attempt to discover and manage data that is written by other applications to retention volumes. Application-specific knowledge is often required to interpret metadata and content. Retention storage is used by IBM StoredIQ to manage data on compliant immutable storage for retention and holds. It does so in a way that does not interfere with knowledge workers who create, update, and access content on primary volumes. IBM StoredIQ preserves source metadata when data is written to a retention volume.

The original source metadata is important for governance and legal discovery (custodian, time stamps, and so on) to replicate the content and metadata from the retention volume when needed.

To configure a retention volume, in the Add Volume dialog window (Figure 14), change the Volume Type to **Retention**. Remember, the two retention volume types are CIFS and NFS.

Figure 14 The Add Volume window with the source type Retention options

Table 12 lists the options and corresponding actions to add a Retention volume.

Table 12 Retention volume configuration options

Option	Action
Volume Type	Select Retention .
Source Type	Select CIFS or NFS .
Server	Enter the fully qualified domain name of the server from which the volume is available for mounting.
Username	CIFS Only: Enter the user name that is used to connect to the data source and mount the volume.
Password	CIFS Only: Enter the password for the user.
Assign to Data Server	Select the Data Server that will be used for the index.
Volume Name	Specify a name for the volume.
Share	CIFS Only: Enter the share name for this volume.
Export	NFS Only: Enter the export name for this volume.
Include Metadata for contained objects	Select this option to include the metadata for contained objects.
Include content tagging for Full-Text index	Select this option to perform a full-text index.

Add Discovery Export volume

Export volumes are an unmanaged (not indexed) storage location where content is copied along with metadata and audit detail in a format that can be imported by other applications. A common usage of export volumes is to stage native documents to be imported into a legal review tool in a format such as standard EDRM or a Concordance-compatible format.

To configure a Retention volume, on the Add Volume dialog window (Figure 15), change the Volume Type to **Discovery Export**. Remember, the two Discovery Export volume types are CIFS and NFS.

Figure 15 The Add Volume window with the source type Discovery Export options

Table 13 lists the options and corresponding actions to add a Discovery Export volume.

Table 13 Discovery Export Volume configuration options

Option	Action
Volume Type	Select Discovery Export .
Source Type	Select CIFS or NFS .
Server	Enter the fully qualified domain name of the server from which the volume is available for mounting.
Username	CIFS Only: Enter the user name that is used to connect to the data source and mount the volume.
Password	CIFS Only: Enter the password for the user.
Assign to Data Server	Select the data server that will be used for the index.
Volume Name	Specify a name for the volume.
Share	CIFS Only: Enter the share name for this volume.
Export	NFS Only: Enter the export name for this volume.

Deleting a volume

If you no longer want a StoredIQ volume, you can delete the volume through the StoredIQ Administrator.

When deleting a volume, be aware of the following information:

- ▶ Deleted volumes are removed from all volume lists.
- ▶ Applicable object counts and sizes within the IBM StoredIQ Administrator Dashboard adjust automatically when a volume is deleted.
- ▶ Object counts and sizes within user infosets will remain the same. User infosets are created at a specific point in time when the deleted data source was still available.
- ▶ Users exploring a specific data source and any generated reports will no longer reference the deleted volume.
- ▶ No exceptions are raised on previously executed actions. Instead, the content is no longer available. For example, if an infoset is copied that contained data objects from a volume that has been deleted, no exception is raised.
- ▶ If you mark a desktop volume for deletion, it is automatically removed from the Primary volume list; however, the status of that workstation is set to uninstall in the background. When the desktop client next checks in, it will see that change in status and uninstall itself.

Note: If a retention volume located on a Content Addressable Storage platform, such as Centera or Hitachi that contains content, the volume cannot be deleted because the IBM StoredIQ Platform is the source of record. Instead, you will see the Under Management link.

To delete a volume, display the list of volumes by clicking the **Add Volume** tab (Figure 16).

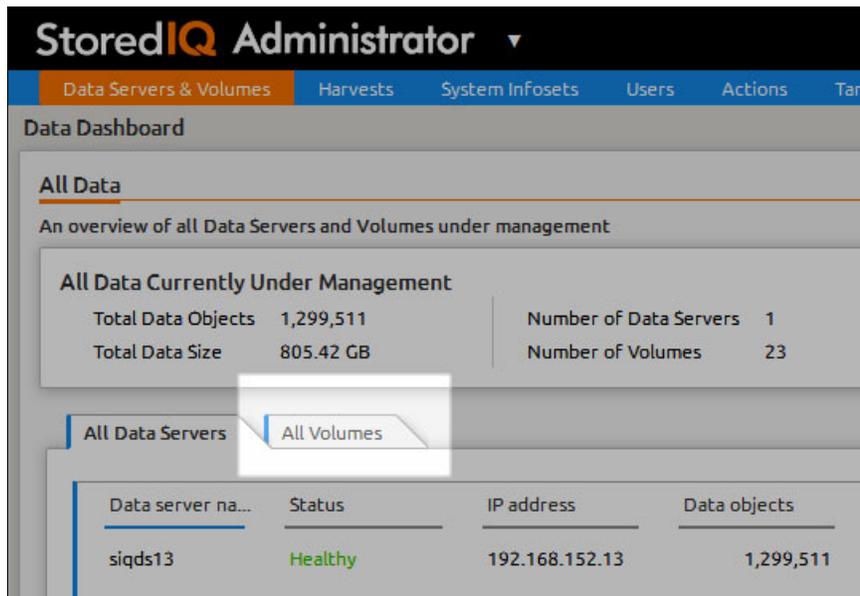


Figure 16 Data Dashboard for deleting a volume

Click **Delete Volume** at the lower right corner of the dashboard (Figure 17).

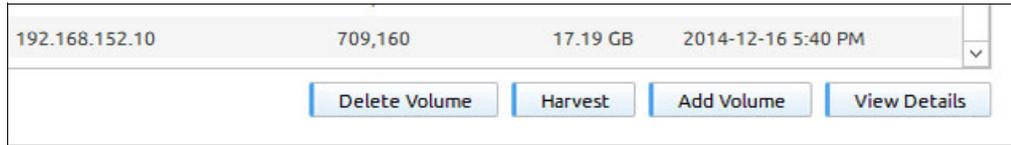


Figure 17 From the Add Volume tab, you can see the option to Delete a Volume

The Confirmation window opens (Figure 18). Click **Delete** to delete the volume.

Important: There is *no* undo function when a volume is deleted. If you choose the “Delete” option, the index for this volume is also deleted. To restore the index information, you must create a new volume.

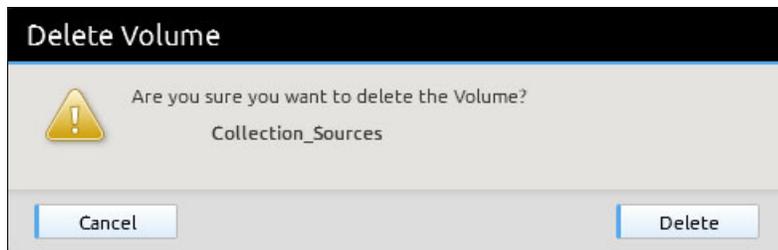


Figure 18 The Delete Volume Confirmation window

When a volume is deleted, the volume is no longer visible in the list of all volumes.

Authors

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



Genifer Graff has an extensive background in records management, eDiscovery, automated document capture, and Legacy Data Cleanup. She has worked with clients all over the world to install and implement these solutions as part of a comprehensive Information Lifecycle Governance Strategy. She has also assisted organizations in implementing global retention programs. Her focus is on helping organizations realize the cost and risk reduction benefits provided by an effective information governance program.

She is a member of an association for records and information management (ARMA) and has presented at the annual ARMA conference for the past five years. She also is a member of Association for Information and Image Management (AIIM) and presented at their annual conference in 2015. She also speaks at several AIIM and ARMA chapter meetings and seminars each year.

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

Jim Sikes, Worldwide Information Lifecycle Governance Tech Sales Leader, IBM System
Whei-Jen Chen, Project Leader, IBM International Technical Support Organization

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.

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

ibm.com/redbooks/residencies.html

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 IBM 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>

Notices

This information was developed for products and services offered in the U.S.A.

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, Armonk, NY 10504-1785 U.S.A.

The following paragraph does not apply to the United Kingdom or any other country where such provisions are inconsistent with local law: 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 states 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 supply in any way it believes appropriate without incurring any obligation to you.

Any performance data contained herein was determined in a controlled environment. Therefore, the results obtained in other operating environments may vary significantly. Some measurements may have been made on development-level systems and there is no guarantee that these measurements will be the same on generally available systems. Furthermore, some measurements may have been estimated through extrapolation. Actual results may vary. Users of this document should verify the applicable data for their specific environment.

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.

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 the names and addresses used by an actual business enterprise 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.

© Copyright International Business Machines Corporation 2015. All rights reserved.

Note to U.S. Government Users Restricted Rights -- Use, duplication or disclosure restricted by GSA ADP Schedule Contract with IBM Corp.

This document REDP-5316-00 was created or updated on December 22, 2015.

Send us your comments 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 U.S.A.

Trademarks

IBM, the IBM logo, and ibm.com are trademarks or registered trademarks of International Business Machines Corporation in the United States, other countries, or both. These and other IBM trademarked terms are marked on their first occurrence in this information with the appropriate symbol (® or ™), indicating US registered or common law trademarks owned by IBM at the time this information was published. Such trademarks may also be registered or common law trademarks in other countries. A current list of IBM trademarks is available on the Web at <http://www.ibm.com/legal/copytrade.shtml>

The following terms are trademarks of the International Business Machines Corporation in the United States, other countries, or both:

AIX®

DB2®

Domino®

FileNet®

IBM®

Redbooks®

Redpaper™

Redbooks (logo) ®

StoredIQ®

The following terms are trademarks of other companies:

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.

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



REDP-5316-00

ISBN 0738454869

Printed in U.S.A.

Get connected

