IBM StoredIQ

Data Server Administration Guide

Note

Before using this information and the product it supports, read the information in Notices.

This edition applies to Version 7.6.0.17 of product number 5724M86 and to all subsequent releases and modifications until otherwise indicated in new editions.

Contents

List of Tables ..... **vi**

About this publication ..... **viii**

IBM StoredIQ product library viii

Contacting IBM StoredIQ customer support viii

What IBM StoredIQ Data Server provides **1**

IBM StoredIQ Data Server user interface **3**

Navigation within IBM StoredIQ Data Server 3

Web interface icons 5

System administration **6**

Checking the system's status 6

Restarting the system 7

Configuration of IBM StoredIQ **9**

Configuring DA Gateway settings 9

Configuring network settings 10

Configuring mail settings 11

Configuring SNMP settings 11

Configuring notification from IBM StoredIQ 12

Configuration of multi-language settings 12

Setting the system time and date 14

Setting system backup configurations 14

Restoring system backups 14

Backing up the IBM StoredIQ image 15

Managing LDAP connections 15

Managing user accounts 16

Importing encrypted NSF files from Lotus Notes 17

Configuring harvester settings 18

Configuring full-text index settings 20

Specifying data object types 22

Configuring audit settings 22

Configuring hash settings 22

Configuring desktop settings 24

Downloading the IBM StoredIQ Desktop Data Collector installer from the application 24

Upgrading the IBM StoredIQ Desktop Data Collector agent 24

Managing the Encrypted file system recovery agent 25

Volumes and data sources **26**

Volume indexing 26

Server platform configuration 26

Configuration of IBM Connections 33

Configuration of retention servers 34

Creating primary volumes 36

Creating retention volumes 76

Creating system volumes 82

Export and import of volume data 83

Creating discovery export volumes 86

Deleting volumes 87

Action limitations for volume types....87
Volume limitations for migrations.... 91

Data harvesting.... 92
Harvest of properties and libraries....92
Lightweight harvest parameter settings.... 93

Job configuration....95
Creating a job.... 95
Creating a job to discover retention volumes....96
Editing a job....96
Starting a job.... 97
Saving a job.... 97
Running a predefined job....97
Deleting a job.... 98
Monitoring processing.... 98

Desktop collection.... 101
IBM StoredIQ Desktop Data Collector client installation....101
IBM StoredIQ Desktop Data Collector installation methods.... 101
Configuring the IBM StoredIQ Desktop Data Collector collection.... 102
Delete policy: special notes.... 103

Folders....104
Creating a folder.... 104
Deleting a folder.... 104
Moving a folder.... 104
Renaming a folder....104
Copying items to different folders....105
Saving items into different folders.... 105
Filtering items within the folder view....105

Audits and logs....106
Harvest audits....106
Import audits.... 109
Event logs....110
Policy audits....111

Deploying customized web services.... 117
Uninstalling an existing instance of the web service....117
Verifying the location of the hosted web service....117
Configuration of administration knobs....118

Supported file types.... 120
Supported file types by name.... 120
Supported file types by category.... 128
SharePoint supported file types....138

Supported server platforms and protocols....140

Event log messages.... 147
ERROR event log messages....147
INFO event log messages....157
WARN event log messages....165

Notices....172

Privacy policy considerations 173
Terms and conditions for product documentation..... 174
Trademarks..... 174
Index..... 176

List of Tables

1. IBM StoredIQ primary tabs 3
2. Dashboard subtab settings and descriptions 3
3. Configuration settings and descriptions 4
4. IBM StoredIQ dashboard icons 5
5. IBM StoredIQ Folders icons 5
6. System and Application configuration options 9
7. Supported languages 12
8. CIFS/SMB or SMB2 (Windows platforms) primary volumes: fields, required actions, and notes 37
9. NFS v2 and v3 primary volumes: fields, required actions, and notes 39
10. Exchange primary volumes: fields, required actions, and notes 41
11. SharePoint primary volumes: fields, required actions, and notes 44
12. Documentum primary volumes: fields, required actions, and notes 46
13. Discovery Accelerator primary volumes: fields, required actions, and notes 48
14. Domino primary volumes: fields, required actions, and notes 50
15. FileNet primary volumes: fields, required actions, and notes 52
16. NewsGator primary volumes: fields, required actions, and notes 54
17. Livelink primary volumes: fields, required actions, and notes 55
18. Jive primary volumes: fields, required actions, and notes 57
19. Chatter primary volumes: fields, required actions, and notes 59
20. IBM Content Manager primary volumes: fields, required actions, and notes 61
21. CMIS primary volumes: fields, required actions, and notes 63
22. HDFS primary volumes: fields, required actions, and notes 65
23. Connections volumes: Add Volume dialog box options for primary volumes 67
24. SharePoint volumes as primary volumes: Fields and examples 73
25. Enterprise Vault retention volume: fields, required actions, and notes 76
26. CIFS (Windows platforms) retention volume: fields, required actions, and notes 77
27. NFS v3 retention volume: fields, required actions, and notes 78
28. Centera retention volume: fields, required actions, and notes 79
29. Hitachi retention volume: fields, required actions, and notes 80
30. System volume fields, descriptions, and applicable volume types 82
31. Discovery export volumes: fields, required actions, and applicable volume types 86
32. Action limitations for volume types 88
33. Out-of-the-box job types 95
34. Harvest/Volume cache details: Fields, descriptions, and values 98
35. Discovery export cache details: Fields, descriptions, and values 98
36. Harvest audit by volume: Fields and descriptions 106
37. Harvest audit by time: Fields and descriptions 106
38. Harvest overview summary options: Fields and descriptions 107
39. Harvest overview results options: Fields and descriptions 107

40. Harvest overview detailed results: Fields and descriptions....107
41. Imports by volumes details: Fields and descriptions....109
42. Policy audit by name: Fields and descriptions....111
43. Policy audit by volume: Fields and descriptions....111
44. Policy audit by time: Fields and descriptions....111
45. Policy audit by discovery exports: Fields and descriptions....112
46. Discovery export runs by discovery export: Fields and descriptions....112
47. Policy audit details: Fields and descriptions....114
48. Policy audit execution details: Fields and descriptions....114
49. Policy audit data object details: Fields and descriptions....115
50. Types of and reasons for policy audit messages....115
51. IBM StoredIQ administration knobs....118
52. Supported file types by name....120
53. Supported archive file types by category....128
54. Supported CAD file types by category....129
55. Supported database file types by category....129
56. Supported email file types by category....130
57. Supported graphic file types by category....130
58. Supported multimedia file types by category....133
59. Supported presentation file types by category....133
60. Supported spreadsheet file types by category....134
61. Supported system file types by category....135
62. Supported text and markup file types by category....135
63. Supported word-processing file types by category....136
64. Supported SharePoint object types....138
65. Attribute summary....139
66. ERROR event log messages....147
67. INFO event log messages....157
68. WARN event log messages....165

About this publication

IBM® StoredIQ® Data Server Administration Guide describes how to manage the administrative tasks such as administering appliances, configuring IBM StoredIQ, or creating volumes and data sources.

IBM StoredIQ product library

The following documents are available in the IBM StoredIQ product library.

- *IBM StoredIQ Overview Guide*
- *IBM StoredIQ Deployment and Configuration Guide*
- *IBM StoredIQ Data Server Administration Guide*
- *IBM StoredIQ Administrator Administration Guide*
- *IBM StoredIQ Data Workbench User Guide*
- *IBM StoredIQ Policy Manager User Guide*
- *IBM StoredIQ Insights User Guide*
- *IBM StoredIQ Integration Guide*

Contacting IBM StoredIQ customer support

For IBM StoredIQ technical support or to learn about available service options, contact IBM StoredIQ customer support at this phone number:

- 1-866-227-2068

Or, see the Contact IBM web site at http://www.ibm.com/contact/us/.

IBM Knowledge Center

The IBM StoredIQ documentation is available in IBM Knowledge Center.

Contacting IBM

For general inquiries, call 800-IBM-4YOU (800-426-4968). To contact IBM customer service in the United States or Canada, call 1-800-IBM-SERV (1-800-426-7378).

For more information about how to contact IBM, including TTY service, see the Contact IBM website at http://www.ibm.com/contact/us/.

What IBM StoredIQ Data Server provides

IBM StoredIQ provides access to data-server functions. It allows administrators to configure system and application settings, manage volumes, administer harvests, configure jobs and desktop collection, manage folders, access audits and error logs, and deploy customized settings.

The administrator is the person responsible for managing the IBM StoredIQ installation at a customer site. This individual has strong understanding of data sources, harvests, indexes, data servers, jobs, infosets, reports, concepts, and actions. This list provides an overview as to how IBM StoredIQ works.

Configuring system and application settings

Using IBM StoredIQ Administrator, the administrator can identify what data servers are deployed, their location, what data is being managed, and the status of each data server in the system. Volume management is a central component of IBM StoredIQ. With IBM StoredIQ Administrator, the administrator can see:

- Configure the DA gateway.
- View and modify network settings, including host name, IP address, NIS domain membership, and use.
- View and modify settings to enable the generation of email notification messages.
- Configure SNMP servers and communities.
- Manage notifications for system and application events.
- View and modify date and time settings for IBM StoredIQ.
- Set backup configurations.
- Manage LDAP connections.
- Manage users.
- Upload Lotus Notes user IDs so that encrypted NSF files can be imported into IBM StoredIQ.
- Specify directory patterns to exclude during harvests.
- Specify options for full-text indexing.
- View, add, and edit known data object types.
- View and edit settings for policy audit expiration and removal.
- Specify options for computing hash settings when harvesting.
- Specify options to configure the desktop collection service.

Managing volumes and data sources

A volume represents a data source or destination that is available on the network to the IBM StoredIQ appliance, and they are an integral to IBM StoredIQ indexing your data. Only administrators can define, configure, and add or remove volumes to IBM StoredIQ Data Server.

Administering harvests

Harvesting (or indexing) is the process or task by which IBM StoredIQ examines and classifies data in your network. Within IBM StoredIQ Data Server, you can specify harvest configurations.

Configuring jobs

Harvests and discover retention volume jobs can be configured and run from the IBM StoredIQ Data Server.

Configuring Desktop collection

When you configure desktop settings, you are enabling or disabling encryption within IBM StoredIQ. The IBM StoredIQ Desktop Data Collector enables desktops as a volume type or data source, allowing them to be used just as other types of added data sources. It can collect PSTs and compressed files.

Managing folders

The **Library** and **Workspace** can be managed and used.

Accessing audits and logs

IBM StoredIQ Data Server contains audit and log categories that can be viewed and downloaded as needed.

Deploying customized web services

IBM StoredIQ Data Server can be used deploy SharePoint custom web services.

IBM StoredIQ Data Server user interface

Describes the IBM StoredIQ Data Server web interface and outlines the features within each tab. References to sections where you can find additional information on each topic are also provided.

Navigation within IBM StoredIQ Data Server

The primary tabs and subtabs found within the user interface provide you with the access to data server functionality.

Primary IBM StoredIQ Data Server tabs

IBM StoredIQ users do most tasks with the web interface. The menu bar at the top of the interface contains three primary tabs that are described in this table.

Table 1: IBM StoredIQ primary tabs

Tab name	Description
Administration	Allows Administrators to do various configurations on these subtabs: **Dashboard**, **Data Sources**, and **Configuration**.
Folders	Create folders and jobs; run jobs.
Audit	Examines a comprehensive history of all harvests, run policies, imports, and event logs.

Administration tab

The **Administration** tab includes these subtabs: **Dashboard**, **Data Sources**, and **Configuration**.

- **Dashboard**: The **Dashboard** subtab provides an overview of the system's current, ongoing, and previous processes and its status. This table describes administrator-level features and descriptions.
- **Data sources**: The **Data sources** subtab is where administrators define servers and volumes. They can be places that are indexed or copied to. Various server types and volumes can be configured for use in managing data. Administrators can add Enterprise Vault sites, Centera pools, and FileNet servers through the **Specify servers** area. Volumes are configured and imported in the **Specify volumes** section.
- **Configuration**: The administrator configures system and application settings for IBM StoredIQ through the **Configuration** subtab.

Table 2: Dashboard subtab settings and descriptions

Dashboard setting	Description
Page refresh	Choose from 30-second, 60-second, or 90-second intervals to refresh the page.
Today's job schedule	View a list of jobs that are scheduled for that day with links to the job's summary.
System summary	View a summary of system details, including system data objects, contained data objects, volumes, and the dates of the last completed harvest.
Jobs in progress	View details of each job step as it is running, including estimated time to completion, average speed, total system and contained objects that are encountered, harvest exceptions, and binary processing information.

Table 2: Dashboard subtab settings and descriptions (continued)	
Dashboard setting	**Description**
Harvest statistics	Review the performance over the last hour for all harvests.
Event log	Review the last 500 events or download the entire event log for the current date or previous dates.
Appliance status	Provides a status view of the appliance. Restart the appliance through the about appliance link. View cache details for volumes and discovery exports.

Table 3: Configuration settings and descriptions	
Configuration setting	**Description**
System	• **DA Gateway settings**: Configure the DA Gateway host or IP address. • **Network settings**: Configure the private and public network interfaces. • **Mail server settings**: Configure what mail server to use and how often to send email. • **SNMP settings**: Configure Simple Network Management Protocol (SNMP) servers and communities. • **System time and date**: Set the system time and date on the appliance. • **Backup configuration**: Back up the system configuration of the server to the IBM StoredIQ gateway server. • **Manage LDAP connections**: Add, edit, and remove LDAP connections. • **Manage users**: Add, remove, and edit users. • **Lotus Notes user administration**: Add a Lotus Notes User.
Application	• **Harvester settings**: Set basic parameters and limits, data object extensions and directories to skip, and reasons to run binary processing. • **Full-text settings**: Set full-text search limits for length of word and numbers and edit stop words. • **Data object types**: Set the object types that appear in the disk use by data object type report. • **Audit settings**: Configure how long and how many audits are kept. • **Hash settings**: Configure whether to compute a hash when harvesting and which hash. • **Desktop settings**: Configure the desktop collection service.

Folders tab

Within the **Folders** tab, any type of user can create and manage application objects.

Audit tab

IBM StoredIQ audit feature allows Administrators to review all actions that are taken with the data server, including reviewing harvests and examining the results of actions.

Web interface icons

The following tables describe the icons that are used throughout IBM StoredIQ web interface.

IBM StoredIQ icons

Table 4: IBM StoredIQ dashboard icons	
Dashboard icon	**Description**
User account	The **User account** icon accesses your user account, provides information about version and system times, and logs you out of the system. For more information, see Logging In and Out of the System.
Inbox	The inbox link provides you the access to the PDF audit reports.
Help	Clicking the **Help** icon loads IBM StoredIQ technical documentation in a separate browser window. By default, the technical documentation is loaded as HTML help.

Folders icons

*Table 5: IBM StoredIQ **Folders** icons*	
Folder icon	**Description**
New	Use the **New** icon to add jobs and folders.
Action	Use the **Action** icon to act on workspace objects, including the ability to move, and delete jobs and folders.
Job	Jobs tasks such as harvesting are either a step or a series of steps. For more information, see “Job configuration” on page 95.
Folder	Folders are a container object that can be accessed and used by administrators. For more information, see Folders.
Folder Up	Folders are a container object that can be accessed and used by administrators. By default, you view the contents of the Workspace folder; however, by clicking this icon, you move to the parent folder in the structure. For more information, see Using Folders.

Audit icons

No specialized icons are used on the **Audit** tab.

System administration

System administration entails checking the system's status and restarting the appliance.

Checking the system's status

You can check the system's status for information about various appliance details.

1. Go to **Administration > Dashboard > Appliance status**.
2. Click **About appliance** to open the Appliance details page. The Appliance details page shows the following information:
   - Node
   - Harvester processes
   - Status
   - Software version
   - View details link
3. Click the **View details** link for the controller. The following table defines appliance details and describes the data that is provided for the node.

Option	Description
View appliance details	Shows software version and details of harvester processes running on the controller for the appliance.
Application services	Shows a list of all services and status, which includes this information: • Service: the name of each service on the appliance component • PID: the process ID associated with each service • Current memory (MB): the memory that is being used by each service • Total memory (MB): total memory that is being used by each service and all child services • Processor percentage: the percentage of processor usage for each service. This value is zero when a service is idle. • Status: the status of each service. Status messages include Running, Stopped, Error, Initializing, Unknown.
System services	Shows a list of basic system information details and memory usage statistics. System information includes this information: • System time: current time on the appliance component • GMT Offset: the amount of variance between the system time and GMT. • Time up: the time the appliance component was running since the last restart, in days, hours, minutes, and seconds • System processes: the total number of processes that are running on the node • Number of processors: the number of processors in use on the component • Load Average (1 Minute): the average load for system processes during a 1-minute interval • Load Average (5 Minutes): the average load for system processes during a 5-minute interval

Option	Description
	• Load Average (10 Minutes): the average load for system processes during a 10-minute interval Memory details include this information: • Total: total physical memory on the appliance component • In use: how much physical memory is in use • Free: how much physical memory is free • Cached: amount of memory that is allocated to disk cache • Buffered: the amount of physical memory that is used for file buffers • Swap total: the total amount of swap space available (in use plus free) • System services Swap in use: the total amount of swap space that is used • Swap free: the total amount of swap space free • Database Connections • Configured • Active • Idle • Network interfaces • Up or down status for each interface
Storage	Storage information for a controller includes this information: • Volume • Total space • Used space • Percentage
Controller and compute node status	Indicator lights show component status • Green light: Running • Yellow light: The node is functional but is in the process of rebuilding; performance can be degraded during this time. Note: The rebuild progresses faster if the system is not being used. • Red light: not running Expand the node to obtain details of the appliance component by clicking the image.

Restarting the system

Restarting the system must be done periodically.

Note: The web application is temporarily unavailable if you restart it. Additionally, whenever volume definitions are edited or modified, you must restart the system.

1. Select **Administration > Dashboard > Appliance status**. On the Appliance status page, you have two options:
   - Click the **Controller** link.
   - Click **About Appliance**.
2. The **Restart services** and **Reboot** icons appear at the bottom of the window. These icons are available on the **View details** page and on each of the tabs. Click either of these options:

- **Restart services**: Restart all system services that are running on the node.
- **Reboot**: to restart the components.

Configuration of IBM StoredIQ

An administrator can modify the Application and Network areas to configure IBM StoredIQ.

The **Configuration** subtab (**Administration > Configuration**) is divided into **System** and **Application** sections.

Table 6: System and Application configuration options

Section	Configuration options
System	• Configure the DA gateway. • View and modify network settings, including host name, IP address, NIS domain membership, and use. • View and modify settings to enable the generation of email notification messages. • Configure SNMP servers and communities. • Manage notifications for system and application events. • View and modify date and time settings for IBM StoredIQ. • Set backup configurations. • Configure LDAP connections. • Manage users. • Upload Lotus Notes user IDs so that encrypted NSF files can be imported into IBM StoredIQ.
Application	• Specify directory patterns to exclude during harvests. • Specify options for full-text indexing. • View, add, and edit known data object types. • View and edit settings for policy audit expiration and removal. • Specify options for computing hash settings when harvesting. • Specify options to configure the desktop collection service.

Configuring DA Gateway settings

The DA Gateway settings are configured as part of general system-configuration options.

1. Go to **Administration > Configuration > System > DA Gateway Settings**.
2. If secure gateway communication (via stunnel) was enabled during deployment, the **Host** field displays 127.0.0.1. If secured gateway communication was not enabled during deployment, the **Host** field displays the IP address configured during deployment. You can update the IP address or enter the host name fully qualified domain name of the StoredIQ gateway server instead.
   For example, enter `mgmt.example.com` or `192.168.10.10`.
3. Enter the name of the data server in the **Node name** text box.
4. Click **OK**.
   Services must be restarted.

Configuring network settings

Describes how to configure the network settings that are required to operate IBM StoredIQ Data Server.

1. Go to **Administration > Configuration > System > Network settings**.
2. Click **Controller Settings**. Set or modify the following **Primary Network Interface** options.
   - **IP type**: Set to static or dynamic. If it is set to dynamic, the IP address, Netmask, and Default Gateway fields are disabled.
   - **IP address**: Enter the IP address.
   - **Netmask**: Enter the network mask of the IP address.
   - **Default gateway**: Enter the IP address of the default gateway.
   - **Hostname**: Enter the fully qualified domain name that is assigned to the appliance.
   - **Ethernet speed**: Select the **Ethernet** speed.
   - **Separate network for file/email servers**: Specify the additional subnet for accessing file/email servers. If you are using the web application from one subnet and harvesting from another subnet, select this check box.

   For information about enabling or disabling the ports, see Default open ports.

   Restart the system for any primary network interface changes to take effect. See Restarting and Rebooting the Appliance.
3. In **Controller Settings**, set or modify the following **DNS Settings** options.
   - **Nameserver 1**: Set the IP address of the primary DNS server for name resolution.
   - **Nameserver 2**: Set the IP address of the secondary DNS server for name resolution.
   - **Nameserver 3**: Set the IP address of the tertiary DNS server for name resolution.

   DNS settings take effect after they are saved. Changes to the server's IP address take effect immediately. Because the server has a new IP address, you must reflect this new address in the browser address line before next step.
4. Click **OK**.
5. Click **Server name resolution**. Set these options for the data server:
   a) **CIFS file server name resolution**: These settings take effect upon saving.
      - **LMHOSTS**: Enter the IP host name format.
      - **WINS Server**: Enter the name of the WINS server.

   b) **NIS (for NFS)**: These settings take effect upon saving.
      - **Use NIS**: Select this box to enable NIS to set UID/GID to friendly name resolution in an NFS environment.
      - **NIS Domain**: Specify the NIS domain.
      - **Broadcast for server on local network**: If the NIS domain server is on the local network and can be discovered by broadcasting, select this box. This option does not work if the NIS domain server is on another subnet.
      - **Specify NIS server**: If not using broadcast, specify the IP address of the NIS domain server here.

   c) **Doc broker settings** (for Documentum)
      - Enter **Host** for doc broker.
6. Click **OK**.

Configuring mail settings

Mail settings can be configured as part of system configuration options.

1. Go to **Administration > Configuration > System > Mail Server settings**.
2. In **Mail server**, enter the name of the SMTP mail server.
3. In **From address**, enter a valid sender address. If the sender is invalid, some mail servers reject email. A sender address also simplifies the process of filtering email notifications that are based on the sender's email.
4. Click **OK** to save changes.

Configuring SNMP settings

You can configure the system to make Object Identifier (OID) values available to Simple Network Management Protocol (SNMP) client applications. At the same time, you can receive status information or messages about system events in a designated trap. For information about environmental circumstances that are monitored by IBM StoredIQ, see the following table.

1. Go to **Administration > Configuration > System > SNMP settings**.
2. To make OID values available to SNMP client applications, in the **Appliance Public MIB** area:
   a) Select the **Enabled** check box to make the MIB available, that is, to open port 161 on the controller.
   b) In the **Community** field, enter the community string that the SNMP clients use to connect to the SNMP server.
   c) To view the MIB, click **Download Appliance MIB**. This document provides the MIB definition, which can be provided to an SNMP client application.
3. To capture messages that contain status information in the **Trap destination** area:
   a) In the **Host** field, enter the common name or IP address for the host.
   b) In the **Port** field, enter the port number. Port number 162 is the default.
   c) In the **Community** field, enter the SNMP community name.
4. To modify the frequency of notifications, complete these fields in the **Environmental trap delivery** area:
   a) **Send environmental traps only every __ minutes**.
   b) **Send environmental traps again after __ minutes**.
5. Click **OK**. Environmental traps that are monitored by IBM StoredIQ are described in this table.

Option	Description
siqConsoleLogLineTrap	A straight conversion of a console log line into a trap. It uses these parameters: messageSource, messageID, severity, messageText.
siqRaidControllerTrap	Sent when the RAID controller status is anything but normal. Refer to the MIB for status code information. It uses this parameter: nodeNum.
siqRaidDiskTrap	Sent when any attached raid disk's status is anything but OK. It uses this parameter: nodeNum.
siqBbuTrap	Battery Backup Unit (BBU) error on the RAID controller detected. It uses this parameter: nodeNum.
siqCacheBitTrap	Caching indicator for RAID array is off. It uses this parameter: nodeNum.
siqNetworkTrap	Network interface is not UP when it must be. It uses this parameter: nodeNum.
siqDbConnTrap	Delivered when the active Postgres connection percentage exceeds an acceptable threshold. It uses this parameter: nodeNum.

Option	Description
siqFreeMemTrap	Delivered when available memory falls too low. It uses this parameter: nodeNum.
siqSwapUseTrap	Sent when swap use exceeds an acceptable threshold. Often indicates memory leakage. It uses this parameter: nodeNum.
siqCpuTrap	Sent when processor load averages are too high. It uses this parameter: nodeNum.
siqTzMismatchTrap	Sent when the time zone offset of a node does not match the time zone offset of the controller. It uses this parameter: nodeNum.

Configuring notification from IBM StoredIQ

You can configure the system to notify you using email or SNMP when certain events occur.

For a list of events that can be configured, see "Event log messages" on page 147.

1. Go to **Administration > Configuration > System > Manage notifications**.
2. Click **Create a notification**.
3. In the **Event number:** field, search for events by clicking **Browse** or by typing the event number or sample message into the field.
4. Select the event level by clicking the **ERROR**, **WARN**, or **INFO** link.
5. Scroll through the list, and select each event by clicking it. The selected events appear in the **create notification** window. To delete an event, click the delete icon to the right of the event.
6. In the **Destination:** field, select the method of notification: **SNMP**, or **Email address**, or both. If you choose email address, enter one or more addresses in the **Email address field**. If you choose SNMP, the messages are sent to the trap host identified in the **SNMP settings** window, with a trap type of **siqConsoleLogLineTrap**.
7. Click **OK**.
8. To delete an item from the Manage notifications window, select the check box next to the event, and then click **Delete**.

   You can also request a notification for a specific event from the dashboard's event log. Click the **Subscribe** link next to any error message and a prepopulated edit notification screen that contains the event is provided.

Configuration of multi-language settings

The following table lists the languages that are supported by the IBM StoredIQ.

Table 7: Supported languages

Language	Code	Lemmas	Stop words
Arabic	ar	X	
Catalan	ca		
Chinese	zh	X	
Czech	cs	X	
Danish	da	X	
Dutch	nl	X	
English	en	X	X

Table 7: Supported languages (continued)			
Language	**Code**	**Lemmas**	**Stop words**
Finnish	fi	X	
French	fr	X	X
German	de	X	X
Greek	el	X	
Hebrew	he	X	
Hungarian	hu		
Icelandic	is		
Italian	it	X	
Japanese	ja	X	
Korean	ko	X	
Malay	ms		
Norwegian (Bokmal)	nb	X	
Norwegian (Nynorsk)	nn	X	
Polish	pl	X	
Portuguese	pt	X	X
Romanian	ro		
Russian	ru	X	
Spanish	es	X	X
Swedish	sv	X	
Thai	th	X	
Turkish	tr	X	
Vietnamese	vi		

By default, English is the only language that Multi-language Support identifies during a harvest and it is also the default search language. Both the identified language (or languages) and the search default language can be changed in the `siq-findex.properties` file on the data server. This properties file exists in the following directories on each data server:

```
/usr/local/tomcat/webapps/storediq/WEB-INF/classes/siq-findex.properties
/usr/local/tomcat/webapps/query-converter/WEB-INF/classes
```

All versions of the `siq-findex.properties` file must be kept in sync, within one data server as well as across all data servers, for searches to be consistent and correct.

To change the language that the harvester can identify, use the **index.presetLanguageIDs** field, which is the second-to-last line of the file: `index.presetLanguageIDs = en,fr,de,pt`. The first language in the list is the default language, which is assigned to a document whose language cannot be identified.

To change the default search language, use the **search.defaultLanguage** field, which is the last line of the file. The search language is used to determine which language's rules, that is, **stop words**, **lemmas**, **character normalization**, apply in a search. Only one language can be set as the default for search. However, the default language can be manually overwritten in a full-text search: `lang:de[umkämpft großteils]`

After you change this property file, you must restart the data server and reharvest the volumes that are to be searched. If the data server is the DataServer - Distributed type, run the following command on the data server after restarting it:

```
/etc/deepfile/dataserver/es-update-findex-props.py
```

Setting the system time and date

The system's time and date can be modified as needed.

A system restart is required for any changes that are made to the system time and date. See Restarting and Rebooting the Appliance.

1. Go to **Administration > Configuration > System > System time and date**.
2. Enter the current date and time.
3. Select the appropriate time zone for your location.
4. Enable **Use NTP to set system time** to use an NTP server to automatically set the system date and time for the data server.

   If NTP is used to set the system time, then the time and date fields set automatically. However, you must specify the time zone.
5. Enter the name or IP address of the NTP server.
6. Click **OK** to save changes.

Setting system backup configurations

To prepare for disaster recovery, you can back up the system configuration of an IBM StoredIQ data server to an IBM StoredIQ gateway server. This process backs up volume definitions, discovery export records, and data-server settings. It does not back up infosets, data maps, or indexes. The preferred method is to take a snapshot of the virtual machine to use as backup.

The gateway must be configured manually to support this backup.

1. Configure a working IBM StoredIQ data server, gateway, and application stack.

   The installation and configuration procedure are provided in Configuration of IBM StoredIQ.
2. Using an SSH tool, log on to the gateway as `root`.
3. Select **Appliance Tools** > **OK** > **Enable NFS Share for Gateway**.
   A dialog box appears, stating that the system is checking exports.
4. Within a browser, enter the IP address of the data server, `https://<data server name or IP address>`. Log in as an administrator.
5. Start the system backup. Go to **Administration** > **Configuration** > **System** > **Backup configuration**. Click **Start backup**.
6. Check the event log. Go to **Navigate to the Administration** > **Dashboard** page and check the Event log for the status information.

   You can also examine the files created by the backup procedure. To see these files, go to the `/deepfs/backup` directory on the gateway server.

Restoring system backups

To prepare for disaster recovery, you can back up the system configuration of a IBM StoredIQ data server to a IBM StoredIQ gateway server. You can later restore a system's configuration.

1. Build a new data server, which serves as the restore point for the backup. When building this new data server, use the same IP address and server name as the original data server. Verify that the original server is shut down and not on the network.

See Installing the data server and Configuring the IBM StoredIQ data server for the installation and configuration procedures.

2. Ensure the gateway still has its configured NFS mount.
   a) Using an SSH tool, log onto the gateway as `root`.
   b) Select **Appliance Tools > OK > Enable NFS Share for Gateway**.
   A dialog box appears, stating that the system is checking exports.
3. Using an SSH tool, log onto the data server as `root`. At the data prompt, issue the `su util` command.
4. Go to `Appliance Tools > OK > Restore Configuration From Backups`.
5. Enter the gateway's server name or IP address, and then press **Enter**.
6. Provide the full system restore date, or leave the space empty in order to restore the most recent system backup.
7. Enter Y and then press **Enter** to confirm the system's restoration.

Backing up the IBM StoredIQ image

Backing up the IBM StoredIQ image is a good method for disaster recovery. It is also a best practice before you start any upgrades on your image. If you need to back up the IBM StoredIQ image, you must complete the following steps.

An active IBM StoredIQ image must not be backed up by using VMWare VCenter or other product backup utilities. If you do so, the data servers might hang and become unresponsive. Running a backup snapshot on an active IBM StoredIQ image might result in transaction integrity issues.

If a backup snapshot of IBM StoredIQ image is needed, follow these steps:

1. Enter this command to stop all IBM StoredIQ services on all data servers and gateway: `service deepfiler stop`

   Note: Wait 10 minutes after a harvest before you use this command to stop services.
2. Enter this command to stop the `postgresql` database service on all data servers and gateway: `service postgresql stop`
3. Enter this command to stop the IBM StoredIQ services on the application stack: `service appstack stop`
4. Contact the VMWare VCenter administrator to have the IBM StoredIQ image manually backed up. Confirm the work completion before you proceed to the next step.
5. Enter this command to restart the IBM StoredIQ services on all data servers and gateway: `service deepfiler restart`
6. Enter this command to start the IBM StoredIQ services on the application stack: `service appstack start`

   To prepare for disaster recovery, another method is to back up the system configuration of the IBM StoredIQ data server to an IBM StoredIQ gateway server.

Managing LDAP connections

Configure and manage connections to the LDAP server so that you can create IBM StoredIQ users who use LDAP authentication when logging in to IBM StoredIQ.

You must be logged in to IBM StoredIQ Data Server.

To be able to create *LDAP users* in IBM StoredIQ Data Server, at least one LDAP connection must be configured and defined as default connection for authentication.

1. On the **Administration > Configuration** page, click **Manage LDAP connections**.
2. Select one of the following options:
   - To add connections:

a. Click **Add connection**. Then, provide the following connections details:

Parameter	Value
LDAP server	The URL of the LDAP server in the form of an IP address or the FQDN (fully qualified domain name).
Principal	The security principal for this connection in this format: `cn=common_name,ou=organizational_unit,dc=domain_component` If you configure this connection as default connection, the principal must have admin privileges on the LDAP server.
Password	The principal's password.

b. Make this connection the default connection for LDAP authentication. One of the configured connections must be set as default connection before you can add LDAP users. The default connection is used when validating other LDAP users with the LDAP server. Therefore, the security principal that you specify for this connection must have admin privileges on the LDAP server.

c. Click **OK**.

d. Add further connections by repeating steps “2.a” on page 16 to “2.c” on page 16.

- To edit a connection, click the respective entry and update the settings as required. Remember that you must enter the password again to apply the changes.
- To delete a connection, click the respective entry and then click **Delete** on the connection details window. You can delete only connections that are not in use.

Managing user accounts

User account administration includes creating, modifying, and deleting administrative accounts.

The system comes with a default administrative account that you can use for the initial setup. The default system administrator is `admin`. The default password for this account is `admin`. For security purposes, change the password as soon as possible. Also, create additional system administrators for routine administration so that their actions can be audited.

Note: If someone tries to log in while the Database Compactor appliance is doing database maintenance, the administrator can override the maintenance procedure and use the system. For more information, see “Job configuration” on page 95.

1. In a browser window, enter the IP address or host name of IBM StoredIQ Data Server.
2. Log in with your IBM StoredIQ credentials.

   To log in for the first time after the deployment, use the default administrative account. For regular accounts, use the email address that is defined in your account settings to log in. Local users must provide the password they configured in IBM StoredIQ. LDAP users are authenticated by using the LDAP server and must therefore provide their LDAP password.
3. On the **Administration > Configuration** page, click **Manage users**.

   You have the following options:

   - Change the password of the default administrative account.

     From the list, click **The Administrator account**, and then select **Change the “admin” password**.

     This default administrative account is always a local account.
   - Create an account.

     Click **Create new user**. Provide the name and an email address, and select the authentication type and the appropriate notification setting.

For a user with the authentication type **LDAP**, you must also provide the LDAP principal in this format: cn=*common_name*,ou=*organizational_unit*,dc=*domain_component*

For details, see your LDAP documentation.

When you click **OK**, LDAP user information is verified with the LDAP server to make sure that the user exists and that the attributes are valid. Therefore, at least a default LDAP connection must be configured before you can create LDAP users.

Users with the authentication type **Local** must create and maintain a password for authenticating to IBM StoredIQ. The welcome message that they receive provides instructions for creating this password. LDAP users authenticate with their LDAP passwords.

- Edit an account.

  In the list, click the user name of the account you want to edit. Then, click **Edit user** and change settings as required. For an LDAP user, the information is verified with the LDAP server when you click **OK**.

- Lock or unlock a local or LDAP user account.

  In the list, click the user name of the account you want to lock or unlock. Then, click either **Lock account** or **Unlock account**. A locked account is marked accordingly. An account also becomes locked after three failed login attempts. The user cannot log in while the account is locked.

- Change your password.

  This option is available only for local user accounts. Passwords of LDAP users must be changed by using LDAP administration tools.

  In the list, click your user name to open your account. Then, click **Change password**. Alternatively, open your account by selecting **Your account** from the user menu in the navigation bar.

  When you need to change your password while you are not logged in, you can click the **Forgot your password?** link in the login window. You will then receive an email with instructions for creating a new password. Again, this applies to local user accounts only.

- Reset other users' passwords.

  This option is available only for local user accounts. Passwords of LDAP users must be reset by using LDAP administration tools.

  In the list, click the user name of the account. Then, click **Reset password**. The user receives an email with the information that the password was reset and instructions for creating a new password.

- Delete an account.

  In the list, click the user name of the account. Then, click **Delete**.

  When you delete an LDAP user, only the IBM StoredIQ account is deleted. The user account on the LDAP server remains unchanged.

Importing encrypted NSF files from Lotus Notes

IBM StoredIQ can decrypt, import, and process encrypted NSF files from IBM Lotus Domino v7. The feature works by comparing a list of user.id and key pairs that were imported into the system with the key values that lock each encrypted container or email. When the correct match is found, the file is unlocked with the key. After the emails or containers are unlocked, IBM StoredIQ analyzes and processes them in the usual fashion. This topic provides procedural information about how to import encrypted NSF files from IBM Lotus Domino.

These use cases are supported:

- Multiple unencrypted emails within a journaling database that was encrypted with a single user.id key
- Multiple unencrypted emails in an encrypted NSF file

- Multiple encrypted emails within an unencrypted NSF file
- Multiple encrypted emails with the same or different user.id keys, contained in an encrypted NSF file
- Encrypted emails from within a journaling database

1. On the primary data server, go to **Administration > Configuration > Lotus Notes user administration**.
2. Click **Upload a Lotus user ID file**.
   a) In the dialog that appears, click **Browse**, and go to a user file.
   b) In the **Lotus Notes password:** field, type the password that unlocks the selected file.
   c) In the **Description:** field, enter a description for the file.
   d) Click **OK**. Repeat until the keys for all encrypted items are uploaded. When the list is compiled, you can add new entries to it then.
   e) To delete an item from the list, from the **Registered Lotus users** screen, select the check box next to a user, and then click **Delete**. In the confirmation dialog that appears, click **OK**.

   Note: After you upload user IDs, restart services.

Configuring harvester settings

You can use several different harvester settings to fine-tune your index process.

1. Go to **Administration > Configuration > Application > Harvester settings**.
2. To configure **Basic** settings, follow these steps:
   a) **Harvester Processes**: Select either **Content processing** or **System metadata only**.
   b) **Harvest miscellaneous email items**: Select to harvest contacts, calendar appointments, notes, and tasks from the Exchange server.
   c) **Harvest non-standard Exchange message classes**: Select to harvest message classes that do not represent standard Exchange email and miscellaneous items.
   d) **Include extended characters in object names**: Select to allow extended characters to be included in data object names during a harvest.
   e) **Determine whether data objects have NSRL digital signature**: Select to check data objects for NSRL digital signatures.
   f) **Enable parallel grazing**: Select to harvest volumes that were already harvested and are going to be reharvested.

      If the harvest completes normally, parallelized grazing enables harvests to begin where they left off when interrupted and to start at the beginning.
   g) **Index generated text**: Select for the generated text that is extracted by OutsideIn and OCR libraries to be indexed and available for full-text search.
3. Specify **Skip Content processing**.

   In **Data object extensions to be skipped**, specify those file types that you want the harvest to ignore by adding data object extensions to be skipped.
4. To configure **Locations to ignore**, enter each directory that must be skipped. IBM StoredIQ accepts only one entry per line and that regular expressions can be used.
5. To configure **Limits**, follow these steps:
   a) **Maximum data object size**: Specify the maximum data object size to be processed during a harvest.

      During a harvest, files that exceed the maximum data object size are not read. As a result, if full-text/content processing is enabled for the volume, they are audited as skipped: **Configured max. object size**. These objects still appear in the volume cluster along with all file system metadata. Since they were not read, the hash is a hash of the file-path and size of the object, regardless of what the hash settings are for the volume (full/partial/off).

b) **Max entity values per entity**: For any entity type (date, city, address and the like), the system records, per data object, the number of values set in this field.
The values do not need to be unique. For example, if the maximum value is 1,000, and the harvester collects 1,000 instances of the same date (8/15/2009) in a Word document, the system stops counting dates. This setting applies to all user-defined expressions (keyword, regular expression, scoped, and proximity) and all standard attributes.

c) **Max entity values per data object**: Across all entity types, the total (cumulative) number of values that is collected from a data object during a harvest. A 0 in this field means "unlimited".

This setting applies to all user-defined expressions (key-word, regular expression, scoped, and proximity) and all standard attributes.

6. Configure **Binary Processing**.

a) **Run binary processing when text processing fails**: Select this option to run binary processing.

The system runs further processes against content that failed in the harvesting. You can select options for when to start this extended processing and how to scan content. Binary processing does not search image file types such as .GIF and .JPG for text extraction.

b) **Failure reasons to begin binary processing**: Select the check boxes of the options that define when to start extended processing.

Binary processing can enact in extracting text from a file failure in these situations:

- when the format of the file is unknown to the system parameters;
- when the data object type is not supported by the harvester scan;
- when the data object format does not contain actual text.

c) **Data object extensions**: Set binary processing to process all data files or only files of entered extensions. To add extensions, enter one per line without a period.

d) **Text encoding**: Set options for what data to scan and extract at the start of binary processing.

This extended processing can accept extended characters and UTF-16 and UTF-32 encoded characters as text. The system searches UTF-16 and UTF-32 by default.

e) **Minimums**: Set the minimum required number of located, consecutive characters to begin processing for text extraction.
For example, if you enter 4, the system begins text processing when four consecutive characters of a particular select text encoding are found. This setting helps find and extract helpful data from the binary processing, reducing the number of false positives.

7. Click **OK**.

Changes to harvester settings do not take effect until the appliance is rebooted or the application services are restarted.

Optical character recognition processing

Optical character recognition (OCR) processing enables text extraction from graphic image files that are stored inside archives where the **Include content tagging and full-text index option** is selected.

After content typing inside the IBM StoredIQ processing pipeline, enabling OCR processing routes the following file types through an optical character recognition engine OCR to extract recognizable text.

- Windows or OS/2 bitmap (BMP)
- Tag image bitmap file (TIFF)
- Bitmap (CompuServe) (GIF)
- Portable Network Graphics (PNG)
- Joint Picture Experts Group (JPG)

The text that is extracted from image files is processed through the IBM StoredIQ pipeline in the same manner as text extracted from other supported file types. Policies with a specific feature to write out extracted text to a separate file for supported file types do so for image files while OCR processing is enabled.

The OCR processing rate of image files is approximately 7-10 KB/sec per IBM StoredIQ harvester process.

Configuring full-text index settings

Use the full-text index settings feature to customize your full-text index.

Before you configure or search the full-text index, consider the following situations:

- Full-text filters that contain words might not return all instances of those words: You can limit full-text indexing for words that are based on their length. For example, if you choose to full-text index words limited to 50 characters, then no words greater than 50 characters are indexed.
- Full-text filters that contain numbers might not return all instances of those numbers: This situation can occur when number searches are configured as follows:
  - The length of numbers to full-text index was defined. If you configure the full-text filter to index numbers with 3 digits or more and try to index the numbers 9, 99, 999, and the word `stock`, only the number 999 and the word `stock` are indexed. The numbers 9 and 99 are not indexed.
  - Number indexing in data objects that are limited by file extensions. For example, if you choose to full-text index the number 999 when it appears in data objects with the file extensions .XLS and .DOC, then a full-text filter returns only those instances of the number 999 that exist in data objects with the file extensions .XLS and .DOC. Although the number 999 can exist in other data objects that are harvested, these data objects do not have the file extensions .XLS or .DOC.

1. Go to **Administration > Configuration > Application > Full-text settings**.
2. To configure **Limits**:
   a) **Do not limit the length of the words that are indexed**: Select this option to have no limits on the length of words that are indexed.
   b) **Limit the length of words indexed to___characters**: Select this option to limit the length of words that are indexed. Enter the maximum number of characters at which to index words. Words with more characters than the specified amount are not indexed.
3. To configure **Numbers**:
   - **Do not include numbers in the full-text index**: Select this option to have no indexed numbers. This option is selected by default.
   - **Include numbers in the full-text index**: Select this option to have numbers to be indexed.
   - **Include numbers in full-text index but limit them by**: Select this option to have only certain numbers indexed. Define these limits as follows:
     - **Number length**: Include only numbers that are longer than ____ characters. Enter the number of characters a number must contain to be indexed. The Number length feature indexes longer numbers and ignores shorter numbers. By not indexing shorter numbers, such as one- and two-character numbers, you can focus your filter on meaningful numbers. These numbers can be account numbers, Social Security numbers, credit card numbers, license plate numbers, or telephone numbers.
     - **Extensions**: Index numbers that are based on the file extensions of the data objects in which they appear. Select **Limit numbers for all extensions** to limit numbers in all file extensions to the character limits set in **Number length**. Alternatively, select **Limit numbers for these extensions** to limit the numbers that are selected in **Numbers length** only to data objects with certain file extensions. Enter the file extensions one per line that must have limited number indexing. Any data object with a file extension that is not listed has all indexed numbers.
4. To configure **Include word lemmas in index**, select whether to identify and index the lexical forms of words as well.
   For example, `employ` is the lemma for words such as `employed`, `employment`, `employs`. If you use lemmas and search for the word `employed`, IBM StoredIQ denotes any found instances of `employment`, `employ`, `employee`, and so on, when it views the data object.

- **Do not include word lemmas in index (faster indexing)**: By not indexing lemmas, data sources are indexed slightly faster and the index size on disk is smaller.
- **Include word lemmas in index (improved searching)**: By indexing lemmas, filter results can be more accurate, although somewhat slower. Without lemmas, a filter for `trade` would need to be written as `trade`, `trades`, `trading`, or `traded` to get the same effect, and even then a user might miss an interesting variant.

5. Configure **Stop words**.

   Stop words are common words that are found in data objects and are indexed like other words. This allows users to find instances of these words where it matters most. A typical example would be a search expression of 'to be or not to be' (the single quotation marks are a specific usage here). Typically, IBM StoredIQ ignores stop words in search expressions, but because single quotation marks as syntax elements, a user can find Shakespeare's "Hamlet." Indexing stop words slightly increases the amount of required storage space, but relevant documents might be missed without these words present in the index. By default, the following words are considered stop words for the English language: `a`, `an`, `and`, `are`, `as`, `at`, `be`, `but`, `by`, `for`, `if`, `in`, `into`, `is`, `it`, `no`, `not`, `of`, `on`, `or`, `such`, `that`, `the`, `their`, `then`, `there`, `these`, `they`, `this`, `to`, `was`, `will`, `with`.

   To add a stop word, enter one word per line, without punctuation, which includes hyphens and apostrophes.

   Note: As of the IBM StoredIQ 7.6.0.3 release, stop words on the configuration page are for the English language only.

6. Select **Enable OCR image processing** to control at a global level whether Optical Character Recognition (OCR) processing is attempted on image files, such as `PNG, BMP, GIF, TIFF,` and images that are produced from scanned image PDF files. A scanned image PDF file is a PDF file with a document that is scanned into it. Through OCR processing, the images are extracted from scanned image PDF files and texts are extracted from the image files.

   The quality of the text extraction relies on the resolution setting on the image files and images from scanned image PDF files. Thus, the resolution setting must be at 300 dots per inch (DPI) or higher. For text in images that is rotated, small font, or unclear text cannot be extracted.

   If you select this option, you must restart services. See Restarting and Rebooting the Appliance.

7. Select **Always process PDFs for images** to control at a global level to extract text from scanned image PDF files.

   A scanned image PDF is a special type of PDF that is created by scanning a document into PDF and is different from a normal PDF. A scanned image PDF contains one image per entire page and no other elements such as plain text. In contrast, a normal PDF can contain a mix of plain-text elements, embedded objects, and images per page. Text extraction from a scanned image PDF is processing intensive and involves two steps:

   a. Retrieving images from scanned image PDF
   b. Extracting text from the retrieved images

   To identify a PDF as a scanned image PDF and then extract text from it, you must select both the **Enable OCR image processing** option and the **Always process PDFs for images** option. However, to extract text from image files such as PNG, BMP, GIF, and TIFF, you need to select only the **Enable OCR image processing** option (as described in step “6” on page 21) because only step b needs to be performed on these files.

   You can set a maximum number of images for processing by entering the respective count for the **Limit number of images in scanned image PDF to** option. However, this setting does affect text extraction only. The default value is zero, which means that text is extracted from all images that are retrieved from scanned image PDFs.

   If you select this option, you must restart services. See Restarting and Rebooting the Appliance.

8. Click **OK**.

Specifying data object types

On the Data object types page, you can add new data object types and view and edit known data object types. These data objects appear in the Disk usage (by data object type) report. Currently, there are over 400 data object types available.

1. Go to **Administration > Configuration > Application > Data object types**.
2. In the **add data object type** section, enter one or more extensions to associate with the data object type. These entries must be separates by spaces.
   For example, enter `doc txt xls`.
3. Enter the name of the data object type to be used with the extension or extensions.
   For example, enter `Microsoft Word`.
4. Click **Add** to add the extension to the list.

Configuring audit settings

Audit settings can be configured to determine the number of days and number of policy audits to be kept before they are deleted.

1. Go to **Administration > Configuration > Application > Audit settings**.
2. Specify the number of days to keep the policy audits before automatically deleting them.
3. Specify the maximum number of policy audits to keep before automatically deleting them.
4. Specify the file limit for drill-down in policy audits.
5. Click **OK** to save changes.

Configuring hash settings

Hashes are used to identify unique content. Configure the type of hash to compute when harvesting.

By default, IBM StoredIQ computes a SHA-1 hash for each object encountered during harvesting. If the SHA-1 hash is based on the content of the files, it can be used to identify unique files (and duplicates).

For computing such a hash, document content must be fetched over the network even for harvests where only file system metadata is collected. To avoid this, you can disable content based hashing for file system metadata only indexing. This provides the fastest indexing rate at the expense of the ability to identify unique content. In this case, the information used to compute the hash is based on volume and object metadata.

If you change the hash settings between harvests, the next harvest uses the updated settings for any new or modified documents. For example, you might not have content based hashes created initially, but some time after the harvest completed you decide to enable content based hashing. In this case, a full-text harvest (if the volume allows for that) generates regular content based hashes for all documents that are indexed during the harvest.

The hash setting does not impact data object preview.

1. Go to **Administration > Configuration > Application > Hash settings**.
2. Determine whether you want to generate a content based hash.
   - For content based hashes, leave the **Compute data object hash** option selected.

     With this setting, content based hashes are generated as selected for full-text and metadata harvests (see step “4” on page 23).
   - For metadata based hashes, clear the **Compute data object hash** check box.
3. For creating a hash for email, select what email attributes are considered to compute the hash.

Email has characteristics that present a challenge when attempting to identify unique messages based on a hash. Using a pure content based hash, it is likely that emails with identical user-visible content do not share the same SHA-1 hash. Therefore, you can select from a set of attributes the ones to contribute to the hash. By using specific fields to compute the email hash, an email located in a local PST archive in a file system, for example, can be identified as a duplicate of a message in an Exchange mailbox even though they are stored in completely different binary formats.

By default, the following information contributes to the hash:

- The information in the To, From, CC, and BCC attributes
- The email subject
- The content of the email body
- The content of any email attachments

The email hash selections operate independently from the data object hash settings; that is, a data object can have a binary hash or an email hash, but not both.

4. For content based hashes, select whether you want to generate a full or a partial hash. This option is not available if you cleared the **Compute data object hash** check box.

   IBM StoredIQ offers two strategies for computing a content based hash. The default option is to read the entire contents of each file as input to computing a SHA-1 hash for the file (*full hash*). If the content of a file must be read to satisfy other content based index options (container processing or full-text indexing), a full content based hash is always computed.

   If you want only a file system metadata index with the ability to identify unique files, you have the option to create a hash from parts of the file content (*partial hash*). With a partial hash , only a maximum of 128 KB of a file's content is read to compute the hash. This minimizes the amount of data read reducing the workload on the data source and network and effectively increasing the indexing rate.

   For a partial hash, up to four 32 KB blocks from each file are read to compute the hash. If a file is less than 128 KB in size, the entire file content is evaluated. Content to compute the hash for files with a size greater than 128 KB is read as follows:

   - 1 x 32 KB block taken from the beginning of the file
   - 2 x 32 KB blocks equally spaced between the beginning and end of the file
   - 1 x 32 KB block taken from the end of the file

   The resulting four 32 KB blocks are used as input to compute the hash. The partial hash might not be appropriate for all use cases but might be sufficient for use cases such as storage management.

   - For a full hash, leave **Entire data object content (required for data object typing)** selected.

     IBM StoredIQ uses Oracle Outside In Technology filters to determine the object type based on content and to extract additional metadata and text.

     IBM StoredIQ implements its own support for text files, web archives (MHT), IBM Notes® email, and EMC EmailXtender and SourceOne archives.

     If a particular data object cannot be handled with the available text extraction methods, IBM StoredIQ can selectively use binary processing to extract strings from a file. File processed in this way have a *binary processing* attribute associated with them to allow the content to be filtered based on this processing attribute. It can be useful to segregate these files because binary processing can yield a high rate of false positives relative to other content extraction techniques.

     You can configure binary processing in the harvester settings.
   - For a partial hash, select **Partial data object content**.

5. Click **OK**.

Configuring desktop settings

When you configure desktop settings, you are enabling or disabling encryption within IBM StoredIQ. IBM StoredIQ Desktop Data Collector(desktop client or client) enables desktops as a volume type or data source, allowing them to be used just as other types of added data sources. The client is provided as a standard MSI file, and is installed according to the typical method (such as Microsoft Systems Management Service (SMS)) used within your organization. The client can collect PSTs and compressed files and other data objects and it is capable of removing itself when its work is completed.

1. Go to **Administration > Configuration > Application > Desktop settings**.
2. In the **Desktop Services** area, select the **Encrypt all traffic to/from desktops** check box.
3. Select either **Enabled** or **Disabled** to enable or disable desktop services.
4. Click **Apply**.

Downloading the IBM StoredIQ Desktop Data Collector installer from the application

The IBM StoredIQ Desktop Data Collector installer can be downloaded from the application.

The port needs to be opened before downloading the Desktop Data Collector agent. You must log in by using ssh and run the following command on the data server:

```
python /usr/local/storediq/bin/util/port_handler.pyc -e 'desktop'
```

You do not need to restart the services.

1. Go to **Administration > Configuration > Application > Desktop settings**.
2. In the **Download the Desktop Agent installer** area, click **Download the desktop client installer**.
3. When the download is complete, click **Save File**.

After the desktop client is installed on a desktop and connected to and registered with the data server, that desktop is available as a data source within the list of primary volumes. Connectivity and the correct IP address are required.

After the desktop volume appears on the volumes page, edit the include directory by using regular expressions to target the harvest to specific directories on the desktop.

Upgrading the IBM StoredIQ Desktop Data Collector agent

The IBM StoredIQ Desktop Data Collector agent must be upgraded to open and harvest encrypted files.

During collection, if IBM StoredIQ Desktop Data Collector finds an Encrypted File System-encrypted file, the IBM StoredIQ Desktop Data Collector installs a recovery agent certificate, allowing the client to open the encrypted file and harvest from it.

1. Go to **Administration > Configuration > Application > Desktop settings**.
2. In the **Upgrades** area, select either **Automatic upgrade** or **Available versions**.
3. For **Automatic upgrade** options:
   - **Upgrades disabled**: All upgrades are disabled, that is, none is applied.
   - **Upgrade all workstations**: All workstations are upgraded.
4. For **Available versions** options:
   - Select **Manually publish new version**, and then select that version.
   - Select **Automatically publish the latest version**.
5. Click **Apply**.

Managing the Encrypted file system recovery agent

During IBM StoredIQ Desktop Data Collector collection, if IBM StoredIQ Desktop Data Collector finds an Encrypted File System-encrypted file, the IBM StoredIQ Desktop Data Collector installs a recovery agent certificate so that you can open the encrypted file.

1. Go to **Administration > Configuration > Application > Desktop settings**.
2. In the **Encrypted file system recovery agent users** area, click **Add encrypted file system user**. The Upload Recovery Agent Certificate dialog box appears.
3. In the **Select a .PFX file to upload:** text box, click **Browse** to go to the wanted .PFX file. By default, the system takes a .PFX file.
4. Enter the .PFX password in the **.PFX password:** text box. This password protects the file itself.
5. In the **Username:** text box, enter the user name for the user, a SAM compatible/NT4 Domain name-style user name.
   For example, enter `MYCOMPANY\esideways`. This user name is the credential of the user to whom this recovery agent belongs.
6. In the **Password:** text box, enter the password for the user.
7. Enter a description in the **Description:** text box.
8. Click **OK**. The file is uploaded, and the added user is visible within the **User name** column.

   Note:

   After users are added, they can also be edited or deleted with the **Edit** or **Delete** options.

Volumes and data sources

Volumes or data sources are integral to IBM StoredIQ to index your data.

A volume represents a data source or destination that is available on the network to the IBM StoredIQ appliance. A volume can be a disk partition or group of partitions that is available to network users as a single designated drive or mount point. IBM StoredIQ volumes have the same function as partitions on a hard disk drive. When you format the hard disk drive on your PC into drive partitions A, B, and C, you are creating three partitions that function like three separate physical drives. Volumes behave the same way that disk partitions on hard disk drive behave. You can set up three separate volumes that originate from the same server or across many servers. Only administrators can define, configure, and add or remove volumes to IBM StoredIQ.

Volume indexing

When you define volumes, you can determine the type and depth of index that is conducted.

Three levels of analysis are as follows.

- System metadata index. This level of analysis runs with each data collection cycle and provides only system metadata for system data objects in its results. It is useful as a simple inventory of what data objects are present in the volumes you defined and for monitoring resource constraints, such as file size, or prohibited file types, such as the .MP3 files.
- System metadata plus containers. In a simple system metadata index, container data objects (compressed files, PSTs, emails with attachments, and the like) are not included. This level of analysis provides container-level metadata in addition to the system metadata for system data objects.
- Full-text and content tagging. This option provides the full local language analysis that yields the more sophisticated entity tags. Naturally, completing a full-text index requires more system resources than a metadata index. Users must carefully design their volume structure and harvests so that the maximum benefit of sophisticated analytics is used, but not on resources that do not require them. Parameters and limitations on “full-text” indexing are set when the system is configured.

Server platform configuration

Before you configure volumes on IBM StoredIQ, you must configure the server platforms that you use for the different volume types. Each server type has prerequisite permissions and settings.

Because IBM StoredIQ supports several different types of volumes, server platforms must be configured to support those volume types.

Defining server aliases

You can define server aliases for your volumes.

If your server naming conventions aren't very descriptive, you can assign an alias to the server for easier reference. Another use case for an alias might be to allow for mapping several volumes to the same data source such that volumes overlap one another. Exercise care when using server aliases for the second use case because the same physical data object might appear on multiple volumes.

1. Navigate to **Administration** > **Data sources** > **Specify servers** and click **Server aliases**.
2. In the **Add an alias** section, enter a server and an appropriate alias, and click **Add**.

   As server, specify the fully qualified domain name (FQDN) or IP address as appropriate for the server type.

   The information is displayed in the list of server aliases.

In IBM StoredIQ Data Server, you can now use the alias instead of the FQDN or IP address to map a volume to the data source.

At any time, you can edit a server alias. However, in this case, any mappings in which the alias is used will break. You cannot delete an alias that is in use.

Configuring Windows Share (CIFS)

Windows Share (CIFS) must be configured to harvest and run policies.

- To harvest and run policies on volumes on Windows Share (CIFS) servers, the user must be in the backup operator group on the Windows Share server that shows the shares on IBM StoredIQ and also needs to have full control share-level permissions.

Configuring NFS

NFS must be configured to harvest and run policies.

- To harvest and run policies on NFS servers, you must enable root access on the NFS server that is connected to IBM StoredIQ.

Configuration of 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 the **Use SSL** check box or add port number 443 after the server name. If you choose the latter option, an example is `qa03exch2000.qaw2k.local:443`. In some cases, this secure connection can result in some performance degradation due to SSL running large. If you enter the volume information without the 443 suffix, the default connection is 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 2007, 2010, 2013, and Online. The Full Access permissions must be granted on the Exchange server for each mailbox from which you harvest.
- 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*. Configuration information is also available online at www.microsoft.com. It applies only to on-premises versions of Exchange.
- Windows Authentication. For all on-premises versions, enable Integrated Windows Authentication on each Exchange server.
- Public folders. To harvest public folders in Exchange, the **Read Items** privilege is required. It applies to Exchange 2003 and 2007.
- An Exchange 2013 service account must belong to an administrative group or groups granted the following administrator roles:
  - Mailbox Search
  - ApplicationImpersonation
  - Mail Recipients
  - Mail Enabled Public Folders

– Public Folders

An Exchange Online service account must belong to an administrative group or groups granted the following administrator roles, which are required as part of the Service account:

– Mailbox Search
– ApplicationImpersonation
– Mail Recipients
– Mail Enabled Public Folders
– MailboxSearchApplication
– Public Folders

Note: It is possible to create a new Exchange Admin Role specific to IBM StoredIQ that includes only these roles.

The current Exchange Online authentication uses basic authentication over SSL. Volume credentials that are supplied for Exchange Online are only as secure as the SSL session.

Note: Exchange Online connection uses claims-based authentication only. OAuth is not supported currently.

Deleted items might persist because of Exchange Online's retention policies. Exchange Online is a cloud-based service; items are deleted by an automated maintenance task. Items that are deleted manually might persist until the automated job completes.

Enabling integrated Windows authentication on Exchange servers
Windows authentication can be integrated on Exchange servers.

1. From Microsoft Windows, log in to the Exchange Server.
2. Go to **Administrative Tools > Internet Information Services (IIS) Manager**.
3. Go to **Internet Information Services > Name of Exchange Server > Web Sites > Default Web Site**.
4. Right-click **Default Web Site**, and then click the **Directory Security** tab.
5. In the **Authentication and access control** pane, click **Edit**.
6. Select **Properties**. The **Authentication Methods** window appears.
7. In the **Authentication access** pane, select the **Integrated Windows authentication** check box.
8. Click **OK**.
9. Restart IIS services.

Improving performance for IIS 6.0 and Exchange 2003
Within IBM StoredIQ, performance can be improved for IIS 6.0 and Exchange 2003.

1. From Microsoft Windows, log on to the Exchange Server.
2. Go to **Administrative Tools > Internet Information Services (IIS) Manager**.
3. Select **Internet Information Services > <Name of Exchange Server> > Web Sites > Application Pools**.
4. Right-click **Application Pools** and select **Properties**.
5. On the **Performance** tab, locate the **Web Garden** section.
6. If the number of worker processes is different from the default value of **1**, then change the number of worker processes to **1**.
7. Click **OK**.
8. Restart IIS Services.

Configuration of SharePoint

When you configure SharePoint, certain privileges are required by user account along with IBM StoredIQ recommendations. Additionally, SharePoint 2007, 2010, 2013, and 2016 require the configuration of alternate-access mappings to map IBM StoredIQ requests to the correct websites.

To configure SharePoint, consider these connections and privileges:

Secure Connection

If you want to connect to SharePoint volumes over HTTPS, you can either select the **Use SSL** check box or add port number 443 after the server name when you set up the volume on IBM StoredIQ. If you choose the latter option, an example is `qa01.company.com:443`. In some cases, this secure connection can result in some performance degradation due to Secure Socket Layer (SSL) running large. If you enter the volume information without the 443 suffix, the default connection is over HTTP.

Note: SharePoint Online connection uses claims-based authentication only. OAuth is not supported currently.

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 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 with . Consider these points:

- 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 denote 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, then the user credentials must own privileges to **Manage Social Data** as well.
- Privileges: Use a site collection administrator to ensure that all data is harvested from a site or site collection.

Alternate-access mappings

Alternate-access mappings map URLs presented by IBM StoredIQ 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.

Refer to Microsoft SharePoint 2007, 2010, 2013, or 2016 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 SharePoint volume definitions.

For example, you are accessing a SharePoint volume with the fully qualified domain name, `http://itweb.storediqexample.com`, from the intranet zone. An alternate-access mapping for the public URL `http://itweb.storediqexample.com` for the intranet zone must be configured for the SharePoint 2007, 2010, 2013, or 2016 web application that hosts the site to be accessed by the volume definition. If you are accessing the same volume with SSL, the mapping added must be for the URL `https://itweb.storediqexample.com` instead.

Note: When you configure SharePoint volumes with non-qualified names, you are entering the URL for a SharePoint site collection or site that is used by IBM StoredIQ 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.

Configuring Documentum

Documentum has configuration requirements when it is used as a server platform.

- To run harvests and copy from Documentum servers, you must use the **Contributor** role.

Installing Documentum client jars to the data server

Documentum can be used as a data source within IBM StoredIQ, but it must be downloaded and then the RPM installed and run on the target data server.

1. Using this command, create a directory on your target data server to hold the Documentum .JAR files: `mkdir /deepfs/documentum/dfc`
2. Copy the Documentum .JAR files to this location on your target data server: `/deepfs/documentum/dfc`
3. Download and install rpm-build on your target data server.
   a) To locate the rpm-build, go to `http://mirror.centos.org/centos/6/os/x86_64/Packages/rpm-build-4.8.0-59.el6.x86_64.rpm`
   b) Copy `rpm-build-4.8.0-59.el6.x86_64.rpm` to `/deepfs/documentum` on your target data server.
   c) Using this command, change directories and install rpm-build: `cd /deepfs/documentum`
   d) Using this command, install rpm-build: `rpm --nodeps -i rpm-build-4.8.0-59.el6.x86_64.rpm`
4. Using this command, run the script to create the Documentum package: `/usr/local/bin/build-dfc-client-rpm /deepfs/documentum/dfc`
   The output is placed in `/deepfs/documentum` and is named something similar to `siq-war-dfc-client-202.0.0.0p46-1.x86_64.rpm`
5. Using this command, run the newly created rpm on your target data server: `rpm -i siq-war-dfc-client-202.0.0.0p46-1.x86_64.rpm`

You must build the Documentum RPM only once. Copy the created file `siq-war-dfc-client-202.0.0.0p46-1.x86_64.rpm` to each data server you plan to connect to Documentum and then run the rpm as described in step “5” on page 30.

Configuring Enterprise Vault sites

You must take the following actions before you can add an Enterprise Vault volume.

Enterprise Vault sites must be configured.

1. Go to **Administration > Data sources > Specify servers**.
2. Click **Enterprise Vault sites**, and then click **Add new Enterprise Vault site**.
3. In the **Site name** field, enter a unique logical site name.
   The name appears in the screens that are used to configure Enterprise Vault volumes.
4. In the **Enterprise Vault site alias** field, enter the FQDN of the Enterprise Vault Server. Add each server only one time.
5. In the **User name** field, enter the login name. If the user is a domain user, then enter the login name as **domain\user**. Use the Enterprise Vault Service Account or a user with equivalent privileges.
6. In the **Password** field, enter the user’s password to authenticate with Active Directory.

7. Click **OK** to save the site.

Checking on the Enterprise Vault servers

You must also check the Enterprise Vault servers before you add an Enterprise Vault volume.

1. Ensure that the following ports on the Enterprise Vault servers are accessible from the data server.
   - If the Enterprise Vault server is installed in Windows Server 2003, then the RPC dynamic port range 1024-5000 must be accessible for both TCP and UDP.
   - If the Enterprise Vault server is installed in Windows Server 2008, Window Vista or later versions, then the RPC dynamic port range 49152-65535 must be accessible for both TCP and UDP.
2. Confirm that the following Enterprise Vault DCOM API classes are registered on the Enterprise Vault Windows servers.
   - In Windows Server 2003, classes are automatically registered so no further action is needed.
   - In Windows Server 2008 or later versions, confirm that the following classes are registered in the Windows registry. If not, they must be registered.
     – Change the Windows registry to register DCOM API classes to run in a DII Surrogate. The following API classes must be registered for StoredIQ to work with the Enterprise Vault server.
       - Enterprise Vault ContentManagementAPI with Class GUID - *{E4BE20A4-9EF1-4B05-9117-AF43EAB4B295}*
       - Enterprise Vault Retention API with Class GUID - *{744FC7D7-6933-4696-AC3F-9EFC1E00C96B}*
       - Enterprise Vault Directory Connection API with Class GUID - *{4EC6FF78-C97A-11D1-90E0-0000F879BE6A}*
     – Check whether the classes are registered on the Enterprise Vault server, follow these steps:
       a. Log in to the Windows Server where the Enterprise Vault server is installed by using the Vault Service Account credentials.
       b. Open **regedit**.
       c. Find the AppID key for *EVContentManagementAPI* Class.
          1) Go to `HKEY_CLASSES_ROOT\Wow6432Node\CLSID\{E4BE20A4-9EF1-4B05-9117-AF43EAB4B295}`.
          2) Copy the AppID value for this key. For example, `{07E456D0-DFEB-4677-A38F-FEA238945A2A}`.
          3) Go to `HKEY_CLASSES_ROOT\Wow6432Node\AppID` and find the AppID key with the value you copied in the previous step. In the example, look for a key named `{07E456D0-DFEB-4677-A38F-FEA238945A2A}`.
          4) Check whether this key has a String Value (REG_SZ) called `DllSurrogate` with empty Data.
          5) If that string value is not present, then add String Value manually.
             a) Right-click inside the window that displays values for this key.
             b) Select **New** > **String Value**.
             c) Name it as `DllSurrogate`.
       d. Repeat Step c for Enterprise Vault Retention API by using the CLSID *{744FC7D7-6933-4696-AC3F-9EFC1E00C96B}*.
       e. Repeat Step c for Enterprise Vault Directory Connection API by using the CLSID *4EC6FF78-C97A-11D1-90E0-0000F879BE6A}*.

     Note: After you change the registry name, you do not need to restart IIS or the server. Modifying registry entries is needed only for Windows 2008 and later versions.

Configuration of Discovery Accelerator

Before you configure Discovery Accelerator primary volumes, you must configure Discovery Accelerator customer information and Enterprise Vault sites sequentially so that certain configuration items can appear in the volume configuration lists.

You must log in to the Discovery Accelerator server and run the `ImportExport.exe` tool in the installation folder to obtain the appropriate Customer IDs and customer database names.

- **Discovery Accelerator Web-Services Interface**: In order for IBM StoredIQ to interface with Enterprise Vault with the Discovery Accelerator web services, the following configurations must be made on the Discovery Accelerator service.
  1. Log in to the Discovery Accelerator client interface as a Vault User or Discovery Administrator.
  2. Click the **Configuration** tab and expand the **API options** on the **Settings** page.
  3. In the **API settings** group, ensure that the **API Enabled** setting is enabled.
  4. Configure a **Temporary Storage Area** as needed. Ensure that it has sufficient free space and that any authenticated users that define volumes against Discovery Accelerator have Full Control permissions on this storage area. Additionally, consider configuring the **Temporary Storage Area Cleanup Interval**. Depending on the size of the cluster that is deployed, typically, the default value of 30 minutes can be sufficient. If greater than four nodes in the cluster, this interval must be reduced for more frequent cleanups to free up storage space.
- **Improve performance for IIS 6.0**: If the Discovery Accelerator server runs over IIS 6.0, an existing bug in IIS causes severe performance degradation when used along with Kerberos authentication. The hotfix described in Microsoft Knowledge Base article 917557 (http://support.microsoft.com/kb/917557) can be applied to the server in this case.

Discovery Accelerator permissions

The credentials that are used for referencing the Enterprise Vault Site are credentials of the Vault User or any other administrator. IBM StoredIQ validates that the credentials are strong enough for it to:

- Login remotely to the specified server
- Perform DCOM operations over RPC remotely

To harvest a Discovery Accelerator volume successfully, a user must have the following privileges:

- A role that is defined in the Discovery Accelerator Web Application
- Review messages permission for the case that is used in the volume definition
- Folder review permissions on a case, if a folder (sometimes also called a Research Folder) in the case is going to be harvested
- Permission to set all of the review marks that are selected for the volume definition

Configuring security settings for Enterprise Vault servers

The following procedure addresses only the standard security settings that must be configured on the Windows Servers hosting Enterprise Vault to allow it to interact with IBM StoredIQ.

1. Go to **Administration > Data sources > Specify servers**.
2. Click **Discovery Accelerator customers**, and then click **Add new Discovery Accelerator customer**.
3. In the **Customer name** field, enter a unique display name of the DA customer. This name appears in the screens that are used to configure Enterprise Vault volumes.
4. In the **Discovery Accelerator server** field, enter the DNS name of the physical server that runs Discovery Accelerator.
5. In the **Customer virtual directory** field, enter the IIS Virtual Directory where the Discovery Accelerator web service is located.
6. Click **OK** to save the site.

Configuring NewsGator

When NewsGator is used as a server platform, several privileges must be configured.

Privileges Required by User Account: The user account to harvest or copy from a NewsGator volume must have the Legal Audit permission on the NewsGator Social Platform Services running on the SharePoint farm.

1. Log in as an administrator to your SharePoint Central Administration Site.
2. Under **Application Management**, select **Manage Service Applications**.
3. In the Manage Service Applications screen, select the **NewsGator Social Platform Services** row.
4. From the toolbar, select **Administrators**.
5. Add the user account that is used for the NewsGator harvest to the list of administrators. Ensure that the account has the **Legal Audit** permission.

Configuration of IBM Connections

IBM Connections can be harvested and the *Copy from* action to a CIFS target is supported. Discovery Exports are also supported.

Note: Not all Profile fields are harvested, such as mobile number, pager number, and fax number. Custom attributes are supported. Libraries in Connections are links to FileNet objects; these files can be harvested.

The *Copy from* action is supported only with a CIFS target. Any harvested Connections instance has the following directory structure. It is a logical structure of hierarchy, not the actual way that data is stored.

```
Home
  Communities
  Files
  Forums
  Wikis
  Activities
  Blogs
  Status
  Bookmarks
  Events
  Comments
  Profiles
```

Note: When you create a Connections volume, the use of an initial directory, Start directory or End directory beyond two levels of recursion, is not supported. For example, `Home/Files` is supported, but `Home/Files/User1` is not. Additionally, harvest scoping, which is the advanced option in IBM StoredIQ Data Server, is not supported.

A Connections volume that is created in IBM StoredIQ version 7.6.0.10 must be fully reharvested after an upgrade for the objects to be viewed.

Each of the subdirectories has elements under the user name directory. So, if User A created a forum, the directory to find it is `home/forums/userA/<Forum Name>`. If a user created a forum inside a community that is owned by User B, the directory to find it is `home/communities/userB/<Community Name>/forums/<Forum Name>`.

For more information about Connections attributes and their use examples, see the topic about Connections attributes in the IBM StoredIQ Data Workbench documentation.

Setting up the administrator access on Connections

IBM Connections needs an actual user account, not `wasadmin`, to be set up with admin and search-admin privileges. The following procedure describes how to set up the administrator access on Connections.

This procedure needs to be done in the WebSphere® Application Server Administrative Console by the administrator.

1. In the Administrative Console, follow these steps.

a) Go to **Users and Groups** > **Administrative user roles**.

b) Select **Add...** > **Administrator role**.

c) Search for the Connections user account that is used to add the Connections volume in IBM StoredIQ and add it to the role.

d) Click **OK** and select **Save directly to the master configuration**.

2. Follow these steps for each of these applications: Activities, Blogs, Communities, Dogear, Files, Forums, News, Profiles, RichTextEditors, Search, URLPreview, and Wikis.

a) In the Administrative Console, go to **Applications** > **Application Types** > **WebSphere enterprise applications**.

b) Select an application from the list.

c) Select **Security role to user/group mapping** > **Search-admin** > **Map Users...**.

d) Search for the Connections user account that is used to add the Connections volume in IBM StoredIQ and add it to the role.

e) Click **OK** > **OK**.

f) Select **Save directly to the master configuration**.

3. Follow these steps for each of these applications: Activities, Blogs, Common, Communities, Files, Forums, Homepage, Metrics, News, Profiles, PushNotification, RichTextEditors, Search, URLPreview, WidgetContainer, and Wikis.

a) In the Administrative Console, go to **Applications** > **Application Types** > **WebSphere enterprise applications**.

b) Select an application from the list.

c) Select **Security role to user/group mapping** > **admin** > **Map Users...**.

d) Search for the Connections user account that is used to add the Connections volume in IBM StoredIQ and add it to the role.

e) Click **OK** > **OK**.

f) Select **Save directly to the master configuration**.

Configuration of retention servers

IBM StoredIQ supports various types of retention servers.

Retention servers must be configured before you add retention volumes to IBM StoredIQ. As IBM StoredIQ supports different types of retention servers, various configuration options are provided.

Creating Centera pools

When a Centera pool is created, an empty volume set is automatically created and associated with the Centera pool. This feature enables the harvesting of unknown Centera volumes so business policies can be applied to data objects on already retained storage servers. The volume set cannot be edited or deleted from the manage volume sets page.

Advanced retention feature must be enabled.

- Permissions: To support all IBM StoredIQ features, the following effective access profile rights to a Centera pool must be enabled:
  - Read
  - (D)elete
  - (Q)uery
  - (E)xist
  - (W)rite
  - Retention

– (H)old

- ***Centera Pools***: If you have an integrated Centera server, you need to create a Centera pool before you can add a Centera volume. Unlike other volumes, Centera servers are not placed into volume sets but into Centera Pools. After you create a pool, the Centera pool is in the list of available choices when a Centera volume is added.

1. Go to **Administration > Data sources > Specify volumes**.
2. In **Centera pools**, click **Create new pool** to open the Centera pool editor.
3. Enter a unique name for the Centera pool in the **IBM StoredIQ pool profile name** text box.
4. Choose how to define the pool. Select either **Structured Output** or **Single text field** and enter a connection string.
   - **Specify Access**: Enter a profile name and Secret for the pool.
   - **Use .pea file**: This file is a pool-entry authorization.
5. If you chose **Structured input**, click **Add new connection** in the **Connections** section.
6. Enter the IP address for an access point on the server, and then click **Add**. Multiple access points can be specified for failover in the event of a problematic access node.
7. Click **OK** to save the pool. The pool is now available in the **Add volumes** dialog when a Centera volume is added.

Configuration of Enterprise Vault

When you configure Enterprise Vault as a retention volume, you must configure both security settings and DCOM.

Before you use Enterprise Vault to create retention volumes, you must configure Discovery Accelerator customers and Enterprise Vault sites sequentially so that certain configuration items can appear in the retention volume configuration lists. You must log in to Discovery Accelerator and run the `ImportExport.exe` tool to obtain the appropriate Customer IDs and customer database names.

Note: If you use a system other than IBM StoredIQ to import data into Enterprise Vault but still want to use IBM StoredIQ for exporting out of Discovery Accelerator, you must define an Enterprise Vault site within IBM StoredIQ and then use that site to define a Discovery Accelerator volume. When you define the Enterprise Vault site, configuring DCOM configuration is a prerequisite.

Enabling remote DCOM

Remote DCOM is required on all Enterprise Vault servers and Discovery Accelerator.

You can use the DCOM Config utility (`DCOMCnfg.exe`) found in Administrative Tools in Control Panel to configure DCOM settings as Component Services. This utility displays the settings that enable certain users to connect to the computer remotely through DCOM. Members of the Administrators group are allowed to connect remotely to the computer by default.

If the Enterprise Vault Service Account or the user whose credentials are used to define the Enterprise Vault Site IBM StoredIQ does not have permissions to connect remotely for DCOM, then follow this procedure on the target server.

1. Run **dcomcnfg** as a user with administrator privileges.
2. In the Component Services dialog box, expand **Component Services**, expand **Computers**, and then right-click **My Computer** > **Default Properties**. If not already enabled, select the **Enable Distributed COM on this Computer** check box.
3. Click the **COM Security** tab.
4. Under **Launch and Activation Permissions**, click **Edit Limits**.
5. In the Launch Permission dialog box, follow these steps if the user name does not appear in the Groups or user names list:
   a) In the Launch Permission dialog box, click **Add**.
   b) In the Select Users, Computers, or Groups dialog box, add the user name and then click **OK**.

c) In the Launch Permission dialog box, select your user and in the **Allow** column under **Permissions for User**, select **Remote Launch** and select **Remote Activation**, and then click **OK**.

Allowing DCOM traffic through the Windows firewall
DCOM traffic must be allowed through the Windows firewall for Enterprise Vault servers and Discovery Accelerator.

To allow DCOM traffic over the network on the target server, the DCOM TCP port (135) must be open on the Firewall. This command opens this port if it is closed:

```
netsh firewall add portopening protocol=tcp port=135
name=DCOM_TCP135
```

You can also use the Firewall User Interface to open the port.

1. In the Control Panel, double-click **Windows Firewall**.
2. Click **Change Settings**, and then click the **Exceptions** tab.
3. In the **Exceptions** window, select the check box for **DCOM** to enable DCOM traffic through the firewall. If there is no such check box, you can
   a) Click **Add Port**.
   b) In the dialog box, enter **Name** as DCOM and **Port number** as 135. Ensure that **TCP** is selected.
   c) Click **OK**.

Hitachi HCAP configuration requirements

In order for IBM StoredIQ to access the Hitachi HCAP server, certain configuration requirements must be met.

The IBM StoredIQ application accesses the Hitachi HCAP server with HTTP. The HTTP gateway must be enabled on the server. Depending on the current allow/deny lists for the HTTP gateway, you might need to add the IBM StoredIQ data server's IP addresses to the **Allow IP addresses** list.

Creating primary volumes

Primary volumes can be created as data sources by using these volume types: CIFS/SMB or SMB2, NFS v2 and v3, Exchange, SharePoint, Documentum, Discovery Accelerator, Domino, FileNet, NewsGator, Livelink, Jive, Chatter, IBM Content Manager, CMIS, and HDFS (Hadoop).

1. Go to **Administration > Data sources > Specify volumes > Volumes**.
2. On the **Primary volume list** page, click **Add primary volumes**.
3. Enter the information that is described in the following tables, which are based on your server type.
   Individual tables describe the CIFS/SMB or SMB2, NFS v2 and v3, Exchange, SharePoint, Documentum, Discovery Accelerator, Domino, FileNet, NewsGator, Livelink, Jive, Chatter, IBM Content Manager, CMIS volume, and HDFS options.
4. Click **OK** to save the volume.
5. Select one of the following options:
   - **Add another volume on the same server.**
   - **Add another volume on a different server.**
   - **Finished adding volumes.**

   This table describes the fields that are available in the **Add volume** dialog box when you configure primary volumes.

   Note: Case-sensitivity rules for each server type apply. Red asterisks within the user interface denote the fields.

Table 8: CIFS/SMB or SMB2 (Windows platforms) primary volumes: fields, required actions, and notes		
Field name	**Required action**	**Special notes**
Server type	In the **Server type** list, select **CIFS (Windows platform)**.	Both SMB and SMB2 are supported. Depending on the setup of your SMB server, some additional SMB configuration might be required on the IBM StoredIQ data server. For details, see "Configuring SMB properties" on page 70.
Server	In the **Server** text box, enter the fully qualified name of the server where the volume is available for mounting.	
Connect as	In the **Connect as** text box, enter the logon ID that is used to connect and mount the defined volume.	
Password	In the **Password** text box, enter the password that is used to connect and mount the defined volume.	
Volume	In the **Volume** text box, enter the name of the share to be mounted.	
Initial directory	In the **Initial directory** text box, enter the name of the initial directory from which the harvest must begin.	With this feature, you can select a volume further down the directory tree rather than selecting an entire volume.
Index options	Select either or both of the **Index options** check boxes. • **Include system metadata for data objects within containers.** • **Include content tagging and full-text index.**	These options are selected by default.
Validation	To validate volume accessibility, select **Validation**.	When selected (the default state), IBM StoredIQ tests to see whether the volume can be accessed.
Include directories	In **Include directories**, specify a regular expression for included directories for each harvest (if it was specified).	These directories are defined as sets of "first node" directories, relative to the specified (or implied) starting directory, that is considered part of the logical volume.

Table 8: CIFS/SMB or SMB2 (Windows platforms) primary volumes: fields, required actions, and notes (continued)

Field name	Required action	Special notes
Start directory	In **Start directory**, designate a start directory for the harvest. The start directory involves volume partitioning to break up a large volume. If an initial directory is defined, the start directory must be underneath the initial directory. In the case of directories E-H, E is the start directory and H is the end directory.	
End directory	In **End directory**, determine the end directory for the harvest. The end directory is also part of volume partitioning and is the last directory harvested.	
Access Times	In **Access times**, select one of these options: • **Reset access times but do not synchronize them.** (This setting is the default setting.) • **Do not reset or synchronize access times.** • **Reset and synchronize access times on incremental harvests.**	

Table 8: CIFS/SMB or SMB2 (Windows platforms) primary volumes: fields, required actions, and notes (continued)

Field name	Required action	Special notes
Constraints	In **Constraints**, select one of these options: • **Only use __ connection process (es)**: Specify a limit for the number of harvest connections to this volume. If the server is also being accessed for attribute and full-text searches, you might want to regulate the load on the server by limiting the harvester processes. The maximum number of harvest processes is automatically shown. This maximum number is set on the system configuration tab. • **Control the number of parallel data object reads**: Designate the number of parallel data object reads. • **Scope harvests on these volumes by extension**: Include or exclude data objects that are based on extension.	

Table 9: NFS v2 and v3 primary volumes: fields, required actions, and notes

Field name	Required action	Special notes
Server type	In the **Server type** list, select **NFS v2** or **NFS v3**.	
Server	In the **Server** text box, enter the fully qualified name of the server where the volume is available for mounting.	
Volume	In the **Volume** text box, enter the name or names of the volume to be mounted.	
Initial directory	In the **Initial directory** text box, enter the name of the initial directory from which the harvest must begin.	With this feature, you can select a volume further down the directory tree rather than selecting an entire volume.

Table 9: NFS v2 and v3 primary volumes: fields, required actions, and notes (continued)		
Field name	**Required action**	**Special notes**
Index options	Select either or both of the **Index options** check boxes. • **Include system metadata for data objects within containers.** • **Include content tagging and full-text index.**	These options are selected by default.
Validation	To validate volume accessibility, select **Validation**.	When selected (the default state), IBM StoredIQ tests to see whether the volume can be accessed.
Include directories	In **Include directories**, specify a regular expression for included directories for each harvest (if it was specified).	These directories are defined as sets of "first node" directories, relative to the specified (or implied) starting directory, that is considered part of the logical volume.
Start directory	In **Start directory**, designate a start directory for the harvest. The start directory involves volume partitioning to break up a large volume. If an initial directory is defined, the start directory must be underneath the initial directory. In the case of directories E-H, E would be the start directory and H would be the end directory.	
End directory	In **End directory**, determine the end directory for the harvest. The end directory is also part of volume partitioning and is the last directory harvested.	
Access times	In **Access times**, select one of these options: • **Reset access times but do not synchronize them.** (This setting is the default setting.) • **Do not reset or synchronize access times.** • **Reset and synchronize access times on incremental harvests.**	

Table 9: NFS v2 and v3 primary volumes: fields, required actions, and notes (continued)		
Field name	**Required action**	**Special notes**
Constraints	In **Constraints**, select one of these options: • **Only use __ connection process (es)**: Specify a limit for the number of harvest connections to this volume. If the server is also being accessed for attribute and full-text searches, you might want to regulate the load on the server by limiting the harvester processes. The maximum number of harvest processes is automatically shown. This maximum number is set on the system configuration tab. • **Control the number of parallel data object reads**: Designate the number of parallel data object reads. • **Scope harvests on these volumes by extension**: Include or exclude data objects that are based on extension.	

Table 10: Exchange primary volumes: fields, required actions, and notes		
Field name	**Required action**	**Special notes**
Server type	In the **Server type** list, select **Exchange**.	
Version	In the **Version** list, select the appropriate version. Options include **2000/2003**, **2007**, **2010/2013/2016**, and **Online**.	
Server	In the **Server** text box, enter the fully qualified name of the server where the volume is available for mounting.	**Note:** If you selected **Online** as the **Server** option, this field fills automatically with the online server name. For Exchange primary volumes, it is the fully qualified domain name where the OWA is. Multiple Client Access servers on Exchange 2007 are supported. The server load must be balanced at the IP or DNS level.

Table 10: Exchange primary volumes: fields, required actions, and notes (continued)		
Field name	**Required action**	**Special notes**
Mailbox server	When you configure multiple client access servers, enter the name of one or more mailbox servers, which are separated by a comma.	For Exchange primary volumes, it is the fully qualified domain name where the mailbox to be harvested is. **Note:** If you selected **Online** as the **Server** option, this field is not available.
Active Directory server	In the **Active Directory server** text box, enter the name of the Active Directory server.	It must be a fully qualified Active Directory server. **Note:** If you selected **Online** as the **Server** option, this field is not available.
Protocol	To use SSL, select the **Protocol** check box.	**Note:** If you selected **Online** as the **Server** option, the **Use SSL** check box is automatically selected, and this field cannot be edited.
Connect as	In the **Connect as** text box, enter the logon ID that is used to connect and mount the defined volume.	
Password	In the **Password** text box, enter the password that is used to connect and mount the defined volume.	
Volume	In the **Volume** text box, enter the name or names of the volume to be mounted.	For Exchange, enter a friendly name for the volume.
Folder	In **Folder**, select either of the **Mailboxes** or **Public folders** options.	
Initial directory	In the **Initial directory** text box, enter the name of the initial directory from which the harvest must begin.	For Exchange, this field must be left blank if you are harvesting all mailboxes. If you are harvesting a single mailbox, enter the email address for that mailbox.
Virtual root	In **Virtual root**, the name defaults to the correct endpoint for the selected Exchange version.	

Table 10: Exchange primary volumes: fields, required actions, and notes (continued)		
Field name	**Required action**	**Special notes**
Personal archives	Select **Harvest personal archive** to harvest personal archives.	**Note:** This check box is seen only when **Exchange 2010/2013/2016** or **Online** is selected.
Index options	Select either or both of the **Index options** check boxes. • **Include system metadata for data objects within containers.** • **Include content tagging and full-text index.**	These options are selected by default.
Remove journal envelope	When selected, the journal envelope is removed.	
Validation	To validate volume accessibility, select **Validation**.	When selected (the default state), IBM StoredIQ tests to see whether the volume can be accessed.
Include directories	In **Include directories**, specify a regular expression for included directories for each harvest (if it was specified).	These directories are defined as sets of "first node" directories, relative to the specified (or implied) starting directory that is considered part of the logical volume.
Start directory	In **Start directory**, designate a start directory for the harvest. The start directory involves volume partitioning to break up a large volume. If an initial directory is defined, the start directory must be underneath the initial directory. In the case of directories E-H, E would be the start directory and H would be the end directory.	The parameters are date ranges that are used to scope the harvest, the format of which is `YYYY-MM-DD`.
End directory	In **End directory**, determine the end directory for the harvest. The end directory is also part of volume partitioning and is the last directory harvested.	The parameters are date ranges that are used to scope the harvest, the format of which is `YYYY-MM-DD`.

Table 10: Exchange primary volumes: fields, required actions, and notes (continued)		
Field name	**Required action**	**Special notes**
Constraints	In **Constraints**, select one of these options: • **Only use __ connection process (es)**: Specify a limit for the number of harvest connections to this volume. If the server is also being accessed for attribute and full-text searches, you might want to regulate the load on the server by limiting the harvester processes. The maximum number of harvest processes is automatically shown. This maximum number is set on the system configuration tab. • **Control the number of parallel data object reads**: Designate the number of parallel data object reads.	

Table 11: SharePoint primary volumes: fields, required actions, and notes		
Field name	**Required action**	**Special notes**
Server type	In the **Server type** list, select **SharePoint**.	
Version	Select either **2003, 2007, 2010, 2013, 2016** or **Online**.	
Server	In the **Server** text box, enter the fully qualified name of the server where the volume is available for mounting.	When you add SharePoint volumes that contain spaces in the URL, see Special Note: Adding SharePoint Volumes.
Active Directory server	In the **Active Directory server** text box, enter the name of the Active Directory server.	It must be a fully qualified Active Directory server. This option is not available for SharePoint Online.
Protocol	To use SSL, select the **Protocol** check box.	For SharePoint Online, Use SSL is automatically populated and cannot be edited.
Connect as	In the **Connect as** text box, enter the logon ID that is used to connect and mount the defined volume.	

Table 11: SharePoint primary volumes: fields, required actions, and notes (continued)		
Field name	**Required action**	**Special notes**
Password	In the **Password** text box, enter the password that is used to connect and mount the defined volume.	
Volume	In the **Volume** text box, enter the name or names of the volume to be mounted.	When you add SharePoint volumes that contain spaces in the URL, see Special Note: Adding SharePoint Volumes.
Initial directory	In the **Initial directory** text box, enter the name of the initial directory from which the harvest must begin.	• With this feature, you can select a volume further down the directory tree rather than selecting an entire volume. • When you add SharePoint volumes that contain spaces in the URL, see Special Note: Adding SharePoint Volumes.
Index options	Select either or both of the **Index options** check boxes. • **Include system metadata for data objects within containers.** • **Include content tagging and full-text index.**	These options are selected by default. **Note:** For SharePoint Online, full-text indexing of OneNote notebook objects, that is, Notes, is not supported currently. FSMD-based searches for these files are supported.
Subsites	Select **Recurse into subsites**.	
Versions	Select **Include all versions**.	IBM StoredIQ supports indexing versions from SharePoint. For more information, see Special Note: Adding SharePoint Volumes.
Validation	To validate volume accessibility, select **Validation**.	When selected (the default state), IBM StoredIQ tests to see whether the volume can be accessed.
Include directories	In **Include directories**, specify a regular expression for included directories for each harvest (if it was specified).	These directories are defined as sets of "first node" directories, relative to the specified (or implied) starting directory, that is considered part of the logical volume.

Table 11: SharePoint primary volumes: fields, required actions, and notes (continued)		
Field name	**Required action**	**Special notes**
Start directory	In **Start directory**, designate a start directory for the harvest. The start directory involves volume partitioning to break up a large volume. If an initial directory is defined, the start directory must be underneath the initial directory. In the case of directories E-H, E would be the start directory and H would be the end directory.	
End directory	In **End directory**, determine the end directory for the harvest. The end directory is also part of volume partitioning and is the last directory harvested.	
Constraints	In **Constraints**, select one of these options: • **Only use __ connection process (es)**: Specify a limit for the number of harvest connections to this volume. If the server is also being accessed for attribute and full-text searches, you might want to regulate the load on the server by limiting the harvester processes. The maximum number of harvest processes is automatically shown. This maximum number is set on the system configuration tab. • **Control the number of parallel data object reads**: Designate the number of parallel data object reads.	

Table 12: Documentum primary volumes: fields, required actions, and notes		
Field name	**Required action**	**Special notes**
Server type	In the **Server type** list, select **Documentum**.	For Documentum, you must specify the doc broker. See “Configuring Documentum” on page 30.
Doc base	In the **Doc base** text box, enter the name of the Documentum repository.	A Documentum repository contains cabinets, and cabinets contain folders and documents.

Table 12: Documentum primary volumes: fields, required actions, and notes (continued)		
Field name	**Required action**	**Special notes**
Connect as	In the **Connect as** text box, enter the logon ID that is used to connect and mount the defined volume.	
Password	In the **Password** text box, enter the password that is used to connect and mount the defined volume.	
Volume	In the **Volume** text box, enter the name or names of the volume to be mounted.	For Documentum, enter a friendly name for the volume.
Harvest	To enable harvesting, select **Harvest all document versions**.	
Initial directory	In the **Initial directory** text box, enter the name of the initial directory from which the harvest must begin.	• With this feature, you can select a volume further down the directory tree rather than selecting an entire volume.
Index options	Select either or both of the **Index options** check boxes. • **Include system metadata for data objects within containers.** • **Include content tagging and full-text index.**	These options are selected by default.
Validation	To validate volume accessibility, select **Validation**.	When selected (the default state), IBM StoredIQ tests to see whether the volume can be accessed.
Include directories	In **Include directories**, specify a regular expression for included directories for each harvest (if it was specified).	These directories are defined as sets of "first node" directories, relative to the specified (or implied) starting directory that is considered part of the logical volume.

Table 12: Documentum primary volumes: fields, required actions, and notes (continued)		
Field name	**Required action**	**Special notes**
Constraints	In **Constraints**, select one of these options: • **Only use __ connection process (es)**: Specify a limit for the number of harvest connections to this volume. If the server is also being accessed for attribute and full-text searches, you might want to regulate the load on the server by limiting the harvester processes. The maximum number of harvest processes is automatically shown. This maximum number is set on the system configuration tab. • **Control the number of parallel data object reads**: Designate the number of parallel data object reads.	

Table 13: Discovery Accelerator primary volumes: fields, required actions, and notes		
Field name	**Required action**	**Special notes**
Server type	In the **Server type** list, select **Discovery Accelerator**.	
Customer	In the **Customer** list, select the customer for this volume.	
Active Directory server	In the **Active Directory server** text box, enter the name of the Active Directory server.	It must be a fully qualified Active Directory server.
Protocol	To use SSL, select the **Protocol** check box.	API client uses HTTP over SSL to communicate with Discovery Accelerator Server configuration.
Connect as	In the **Connect as** text box, enter the logon ID that is used to connect and mount the defined volume.	
Password	In the **Password** text box, enter the password that is used to connect and mount the defined volume.	
Site	In the **Site** list, select the site of the volume.	

Table 13: Discovery Accelerator primary volumes: fields, required actions, and notes (continued)		
Field name	**Required action**	**Special notes**
Volume	In the **Volume** text box, enter the name or names of the volume to be mounted.	
Discovery Accelerator case	In the **Discovery Accelerator case** text box, enter the name of the Discovery Accelerator.	This text box is populated from Discovery Accelerator when connection is established.
Scope of data	In the **Scope of data** options, select either **All data in this case** or **Only selected review marks**.	
Index options	Select either or both of the **Index options** check boxes. • **Include system metadata for data objects within containers.** • **Include content tagging and full-text index.**	**Include system metadata for data objects within containers** is selected by default.
Validation	To validate volume accessibility, select **Validation**.	When selected (the default state), IBM StoredIQ tests to see whether the volume can be accessed.
Include research folders	In **Include research folders**, specify a regular expression for included research folders for each harvest (if it was specified).	For Discovery Accelerator, this regular expression helps restrict the volume to one or more Research Folders in the case.

Table 13: Discovery Accelerator primary volumes: fields, required actions, and notes (continued)		
Field name	**Required action**	**Special notes**
Constraints	In **Constraints**, select one of these options: • **Only use __ connection process (es)**: Specify a limit for the number of harvest connections to this volume. If the server is also being accessed for attribute and full-text searches, you might want to regulate the load on the server by limiting the harvester processes. The maximum number of harvest processes is automatically shown. This maximum number is set on the system configuration tab. • **Control the number of parallel data object reads**: Designate the number of parallel data object reads. • **Scope harvests on these volumes by extension**: Include or exclude data objects that are based on extension.	

Table 14: Domino primary volumes: fields, required actions, and notes		
Field name	**Required action**	**Special notes**
Server type	In the **Server type** list, select **Domino**.	For Domino, you must first upload at least one user.id. See Adding Domino as a Primary Volume.
Server	In the **Server** text box, enter the fully qualified name of the server where the volume is available for mounting.	For Domino, select the appropriate user name, which was entered with the **Configuration** subtab in the Lotus Notes user administration area.
Connect as	In the **Connect as** text box, enter the logon ID that is used to connect and mount the defined volume.	For Domino, select the user name for the primary user ID. The user ID must be configured on the System Configuration screen under the Lotus Notes user administration link.

Table 14: Domino primary volumes: fields, required actions, and notes (continued)

Field name	Required action	Special notes
Password	In the **Password** text box, enter the password that is used to connect and mount the defined volume.	For Domino, enter the password for the primary user ID.
Volume	In the **Volume** text box, enter the name or names of the volume to be mounted.	For Domino, enter a friendly name for the volume.
Harvest	• To harvest mailboxes, select the **Harvest mailboxes** option. • To harvest mail journals, select the **Harvest mail journals** option. • To harvest all applications, select the **Harvest all applications** option.	This option obtains the list of all known Domino users and their NSFs. It then harvests those mailboxes unless it was pointed to a single mailbox with the initial directory.
Initial directory	Enter the name of the initial directory.	
Index options	Select either or both of the **Index options** check boxes. • **Include system metadata for data objects within containers.** • **Include content tagging and full-text index.**	These options are selected by default.
Validation	To validate volume accessibility, select **Validation**.	When selected (the default state), IBM StoredIQ tests to see whether the volume can be accessed.
Include directories	In **Include directories**, specify a regular expression for included directories for each harvest (if it was specified).	These directories are defined as sets of "first node" directories, relative to the specified (or implied) starting directory, that is considered part of the logical volume.

Table 14: Domino primary volumes: fields, required actions, and notes (continued)		
Field name	**Required action**	**Special notes**
Start directory	In **Start directory**, designate a start directory for the harvest. The start directory involves volume partitioning to break up a large volume. If an initial directory is defined, the start directory must be underneath the initial directory. In the case of directories E-H, E would be the start directory and H would be the end directory.	
End directory	In **End directory**, determine the end directory for the harvest. The end directory is also part of volume partitioning and is the last directory harvested.	
Constraints	In **Constraints**, select one of these options: • **Only use __ connection process (es)**: Specify a limit for the number of harvest connections to this volume. If the server is also being accessed for attribute and full-text searches, you might want to regulate the load on the server by limiting the harvester processes. The maximum number of harvest processes is automatically shown. This maximum number is set on the system configuration tab. • **Control the number of parallel data object reads**: Designate the number of parallel data object reads.	

Table 15: FileNet primary volumes: fields, required actions, and notes		
Field name	**Required action**	**Special notes**
Server type	In the **Server type** list, select **FileNet**.	Within IBM StoredIQ Data Server, the FileNet domain must be configured before any FileNet volumes are created. See “Configuring FileNet” on page 74.

Table 15: FileNet primary volumes: fields, required actions, and notes (continued)		
Field name	**Required action**	**Special notes**
FileNet config	Use the **FileNet config** list to select the FileNet server you would like to use for this configuration.	For more information, see “Configuring FileNet” on page 74.
Connect as	In the **Connect as** text box, enter the logon ID that is used to connect and mount the defined volume.	
Password	In the **Password** text box, enter the password that is used to connect and mount the defined volume.	
Domain	Domain name automatically populates.	
Object store	In the **Object store** list, select the wanted object store.	The object store must exist before the creation of a FileNet primary volume.
Volume	In the **Volume** text box, enter the name or names of the volume to be mounted.	
Index options	Select either or both of the **Index options** check boxes. • **Include system metadata for data objects within containers.** • **Include content tagging and full-text index.**	These options are selected by default.

Table 15: FileNet primary volumes: fields, required actions, and notes (continued)		
Field name	**Required action**	**Special notes**
Constraints	In **Constraints**, select one of these options: • **Only use __ connection process (es)**: Specify a limit for the number of harvest connections to this volume. If the server is also being accessed for attribute and full-text searches, you might want to regulate the load on the server by limiting the harvester processes. The maximum number of harvest processes is automatically shown. This maximum number is set on the system configuration tab. • **Control the number of parallel data object reads**: Designate the number of parallel data object reads.	

Table 16: NewsGator primary volumes: fields, required actions, and notes		
Field name	**Required action**	**Special notes**
Server type	In the **Server type** list, select **NewsGator**.	
Server	In the **Server** text box, enter the fully qualified name of the server where the volume is available for mounting.	
Protocol	To use SSL, select the **Protocol** check box.	
Connect as	In the **Connect as** text box, enter the logon ID that is used to connect and mount the defined volume.	
Password	In the **Password** text box, enter the password that is used to connect and mount the defined volume.	
Volume	In the **Volume** text box, enter the name or names of the volume to be mounted.	Enter a friendly name for the volume.

Table 16: NewsGator primary volumes: fields, required actions, and notes (continued)		
Field name	**Required action**	**Special notes**
Index options	Select either or both of the **Index options** check boxes. • **Include system metadata for data objects within containers.** • **Include content tagging and full-text index.**	These options are selected by default.
Validation	To validate volume accessibility, select **Validation**.	When selected (the default state), IBM StoredIQ tests to see whether the volume can be accessed.

Table 17: Livelink primary volumes: fields, required actions, and notes		
Field name	**Required action**	**Special notes**
Server type	In the **Server type** list, select **Livelink**.	
Server	In the **Server** text box, enter the fully qualified name of the server where the volume is available for mounting.	
Port	In the **Port** text box, enter the name of the port.	
Database	In the **Database** text box, enter the name of the database.	
Search slice	In the **Search slice** text box, enter the name of the search slice.	
Connect as	In the **Connect as** text box, enter the logon ID that is used to connect and mount the defined volume.	
Password	In the **Password** text box, enter the password that is used to connect and mount the defined volume.	
Volume	In the **Volume** text box, enter the name or names of the volume to be mounted.	
Initial directory	In the **Initial directory** text box, enter the search slice and the name of the initial directory from which the harvest must begin.	With this feature, you can select a volume further down the directory tree rather than selecting an entire volume.

Table 17: Livelink primary volumes: fields, required actions, and notes (continued)		
Field name	**Required action**	**Special notes**
Index options	Select either or both of the **Index options** check boxes. • **Include system metadata for data objects within containers.** • **Include content tagging and full-text index.**	These options are selected by default.
Validation	To validate volume accessibility, select **Validation**.	When selected (the default state), IBM StoredIQ tests to see whether the volume can be accessed.
Include directories	In **Include directories**, specify a regular expression for included directories for each harvest (if it was specified).	These directories are defined as sets of "first node" directories, relative to the specified (or implied) starting directory that is considered part of the logical volume.
Start directory	In **Start directory**, designate a start directory for the harvest. The start directory involves volume partitioning to break up a large volume. If an initial directory is defined, the start directory must be underneath the initial directory. In the case of directories E-H, E would be the start directory and H would be the end directory.	
End directory	In **End directory**, determine the end directory for the harvest. The end directory is also part of volume partitioning and is the last directory harvested.	

Table 17: Livelink primary volumes: fields, required actions, and notes (continued)		
Field name	**Required action**	**Special notes**
Constraints	In **Constraints**, select one of these options: • **Only use __ connection process (es)**: Specify a limit for the number of harvest connections to this volume. If the server is also being accessed for attribute and full-text searches, you might want to regulate the load on the server by limiting the harvester processes. The maximum number of harvest processes is automatically shown. This maximum number is set on the system configuration tab. • **Control the number of parallel data object reads**: Designate the number of parallel data object reads. • **Scope harvests on these volumes by extension**: Designate which data objects must be harvested by entering those object extensions.	

Table 18: Jive primary volumes: fields, required actions, and notes		
Field name	**Required action**	**Special notes**
Server type	In the **Server type** list, select **Jive**.	
Server	In the **Server** text box, enter the fully qualified name of the server where the volume is available for mounting.	
Protocol	To use SSL, select the **Protocol** check box.	
Connect as	In the **Connect as** text box, enter the logon ID that is used to connect and mount the defined volume.	
Password	In the **Password** text box, enter the password that is used to connect and mount the defined volume.	

Table 18: Jive primary volumes: fields, required actions, and notes (continued)		
Field name	**Required action**	**Special notes**
Volume	In the **Volume** text box, enter the name or names of the volume to be mounted.	
Initial directory	In the **Initial directory** text box, enter the name of the initial directory from which the harvest must begin.	With this feature, you can select a volume further down the directory tree rather than selecting an entire volume.
Index options	Select either or both of the **Index options** check boxes. • **Include system metadata for data objects within containers.** • **Include content tagging and full-text index.**	These options are selected by default.
Versions	Select **Include all versions**.	
Validation	To validate volume accessibility, select **Validation**.	When selected (the default state), IBM StoredIQ tests to see whether the volume can be accessed.
Include directories	In **Include directories**, specify a regular expression for included directories for each harvest (if it was specified).	These directories are defined as sets of "first node" directories, relative to the specified (or implied) starting directory that is considered part of the logical volume.
Start directory	In **Start directory**, designate a start directory for the harvest. The start directory involves volume partitioning to break up a large volume. If an initial directory is defined, the start directory must be underneath the initial directory. In the case of directories E-H, E would be the start directory and H would be the end directory.	
End directory	In **End directory**, determine the end directory for the harvest. The end directory is also part of volume partitioning and is the last directory harvested.	

Table 18: Jive primary volumes: fields, required actions, and notes (continued)		
Field name	**Required action**	**Special notes**
Constraints	In **Constraints**, select one of these options: • **Only use __ connection process (es)**: Specify a limit for the number of harvest connections to this volume. If the server is also being accessed for attribute and full-text searches, you might want to regulate the load on the server by limiting the harvester processes. The maximum number of harvest processes is automatically shown. This maximum number is set on the system configuration tab. • **Control the number of parallel data object reads**: Designate the number of parallel data object reads.	

Table 19: Chatter primary volumes: fields, required actions, and notes		
Field name	**Required action**	**Special notes**
Server type	In the **Server type** list, select **Chatter**.	For Chatter, see “Configuring Chatter messages” on page 74.
Connect as	In the **Connect as** text box, enter the logon ID that is used to connect and mount the defined volume.	
Password	In the **Password** text box, enter the password that is used to connect and mount the defined volume.	
Auth token	In the **Auth token** text box, enter the token that is used to authenticate the Chatter volume.	The auth token must match the user name that is used in the **Connect as** field. Auth tokens can be generated online on Salesforce. See Configuring chatter messages.
Volume	In the **Volume** text box, enter the name or names of the volume to be mounted.	

Table 19: Chatter primary volumes: fields, required actions, and notes (continued)		
Field name	**Required action**	**Special notes**
Initial directory	In the **Initial directory** text box, enter the name of the initial directory from which the harvest must begin.	With this feature, you can select a volume further down the directory tree rather than selecting an entire volume.
Index options	Select either or both of the **Index options** check boxes. • **Include system metadata for data objects within containers.** • **Include content tagging and full-text index.**	These options are selected by default.
Validation	To validate volume accessibility, select **Validation**.	When selected (the default state), IBM StoredIQ tests to see whether the volume can be accessed.
Include directories	In **Include directories**, specify a regular expression for included directories for each harvest (if it was specified).	These directories are defined as sets of "first node" directories, relative to the specified (or implied) starting directory, that is considered part of the logical volume.
Start directory	In **Start directory**, designate a start directory for the harvest. The start directory involves volume partitioning to break up a large volume. If an initial directory is defined, the start directory must be underneath the initial directory. In the case of directories E-H, E would be the start directory and H would be the end directory.	
End directory	In **End directory**, determine the end directory for the harvest. The end directory is also part of volume partitioning and is the last directory harvested.	

Table 19: Chatter primary volumes: fields, required actions, and notes (continued)		
Field name	**Required action**	**Special notes**
Constraints	In **Constraints**, select one of these options: • **Only use __ connection process (es)**: Specify a limit for the number of harvest connections to this volume. If the server is also being accessed for attribute and full-text searches, you might want to regulate the load on the server by limiting the harvester processes. The maximum number of harvest processes is automatically shown. This maximum number is set on the system configuration tab. • **Control the number of parallel data object reads**: Designate the number of parallel data object reads.	

Table 20: IBM Content Manager primary volumes: fields, required actions, and notes		
Field name	**Required actio**	**Special notes**
Server type	In the **Server type** list, select **IBM Content Manager**.	
Server	In the **Server** text box, enter the fully qualified host name of the library server database.	
Port	In the **Port** text box, enter the port that is used to access the library server database.	
Repository	In the **Repository** text box, enter the name of the library server database.	
Database type	In the **Database Type** list, select the type of database that is associated with the volume. Options include **DB2** and **Oracle**. By default, **DB2** is selected.	
Schema	In the **Schema** text box, enter the schema of the library server database.	
Remote database	Optional: In the **Remote database** text box, enter the name of the remote database.	

Table 20: IBM Content Manager primary volumes: fields, required actions, and notes (continued)		
Field name	**Required actio**	**Special notes**
Connection String	Optional: In the **Connection String** text box, enter any additional parameters.	
Harvest itemtype	In the **Harvest itemtype** text box, enter the name of the item types to be harvested, separated by commas.	Harvest type is required to harvest the CM8 volume.
Copy to itemtype	The **Copy to itemtype** text box is a read-only field. Only SiqDocument is supported.	In this release, the attribute lengths are increased for some. For more information, see "IBM Content Manager attributes" on page 75
Connect as	In the Connect as text box, enter the logon ID that is used to connect and mount the defined volume. **Note:** This Content Manager user ID must have access to all documents to be able to create documents in the SiqDocument item type for copy to. If the SiqDocument item type does not exist, a Content Manager administration ID must be used, as this ID creates the item type when the volume is created.	When selected (the default state), IBM StoredIQ tests to see whether the volume can be accessed.
Password	In the **Password** text box, enter the password that is used to connect and mount the defined volume.	
Volume	In the **Volume** text box, enter the name or names of the volume to be mounted.	
Index options	Select either or both of the Index options check boxes. • **Include system metadata for data objects within containers.** • **Include content tagging and full-text index.**	These two options are selected by default.
Validation	To validate volume accessibility, select **Validation**.	When selected (the default state), IBM StoredIQ tests to see whether the volume can be accessed.

Table 20: IBM Content Manager primary volumes: fields, required actions, and notes (continued)		
Field name	**Required actio**	**Special notes**
Constraints	In Constraints, select one of these options: • **Only use __ connection process (es)**: Specify a limit for the number of harvest connections to this volume. If the server is also being accessed for attribute and full-text searches, you might want to regulate the load on the server by limiting the harvester processes. The maximum number of harvest processes is automatically shown. This maximum number is set on the system configuration tab. • **Control the number of parallel data object reads**: Designate the number of parallel data object reads. • **Scope harvests on these volumes by extension**: Designate which data objects must be harvested by entering those object extensions.	

Table 21: CMIS primary volumes: fields, required actions, and notes		
Field name	**Required action**	**Special notes**
Server type	In the **Server type** list, select **CMIS**.	
Server	In the **Server** text box, enter the fully qualified name of the server where the volume is available for mounting.	
Port	In the **Port** text box, enter the name of the port.	
Repository	In the **Repository** text box, enter the name of the repository.	
Service	In the **Service** text box, enter the name of the service.	
Protocol	To use SSL, select the **Protocol** check box.	

Table 21: CMIS primary volumes: fields, required actions, and notes (continued)		
Field name	**Required action**	**Special notes**
Connect as	In the **Connect as** text box, enter the logon ID that is used to connect and mount the defined volume.	
Password	In the **Password** text box, enter the password that is used to connect and mount the defined volume.	
Volume	In the **Volume** text box, enter the name or names of the volume to be mounted.	
Index options	Select either or both of the **Index options** check boxes. • **Include system metadata for data objects within containers.** • **Include content tagging and full-text index.**	These options are selected by default.
Validation	To validate volume accessibility, select **Validation**.	When selected (the default state), IBM StoredIQ tests to see whether the volume can be accessed.
Constraints	In **Constraints**, select one of these options: • **Only use __ connection process (es)**: Specify a limit for the number of harvest connections to this volume. If the server is also being accessed for attribute and full-text searches, you might want to regulate the load on the server by limiting the harvester processes. The maximum number of harvest processes is automatically shown. This maximum number is set on the system configuration tab. • **Control the number of parallel data object reads**: Designate the number of parallel data object reads.	

Table 22: HDFS primary volumes: fields, required actions, and notes

Field name	Required action	Special notes
Server type	In the **Server type** list, select **HDFS**.	
Server	In the **Server** text box, enter the host name or IP address.	
Port	In the **Port** text box, enter the name of the port.	
Repository	In the **Repository** text box, enter the name of the repository.	
Option string	This option is supported: VerifyCertificate=True.	This option is used to indicate that the validity of the HDFS server's SSL certificate is verified when SSL is used. Values are True, False, or default value. If no value is specified, value is False. To validate the certificate on the HDFS server, the user needs to specify this option and set the value to True.
Protocol	To use SSL, select the **Protocol** check box.	
Connect as	In the **Connect as** text box, enter the logon ID that is used to connect and mount the defined volume.	
Password	In the **Password** text box, enter the password that is used to connect and mount the defined volume.	Authentication to HDFS is not supported. If your HDFS server requires a password, StoredIQ is not able to connect to it.
Volume	In the **Volume** text box, enter the name or names of the volume to be mounted.	
Initial directory	In the **Initial directory** text box, enter the name of the initial directory from which the harvest must begin.	To avoid the interrogator timeout issue, see **Note** at the end of this table.
Index options	Select either or both of the **Index options** check boxes. • **Include system metadata for data objects within containers.** • **Include content tagging and full-text index.**	These options are selected by default.

Table 22: HDFS primary volumes: fields, required actions, and notes (continued)		
Field name	**Required action**	**Special notes**
Validation	To validate volume accessibility, select **Validation**.	
Include directories	In **Start directory**, designate a start directory for the harvest. The start directory involves volume partitioning to break up a large volume. If an initial directory is defined, the start directory must be underneath the initial directory. For example, in the case of directories E-H, E would be the start directory and H would be the end directory.	
Start directory	In **Start directory**, designate a start directory for the harvest. The start directory involves volume partitioning to break up a large volume. If an initial directory is defined, the start directory must be underneath the initial directory. In the case of directories E-H, E would be the start directory and H would be the end directory.	
End directory	In **End directory**, determine the end directory for the harvest. The end directory is also part of volume partitioning and is the last directory harvested.	

Table 22: HDFS primary volumes: fields, required actions, and notes (continued)		
Field name	**Required action**	**Special notes**
Contraints	In **Constraints**, select one of these options: • **Only use __ connection process (es)**: Specify a limit for the number of harvest connections to this volume. If the server is also being accessed for attribute and full-text searches, you might want to regulate the load on the server by limiting the harvester processes. The maximum number of harvest processes is automatically shown. This maximum number is set on the system configuration tab. • **Control the number of parallel data object reads**: Designate the number of parallel data object reads. • **Scope harvests on these volumes by extension**: Designate which data objects must be harvested by entering those object extensions.	

Note: If you harvest HDFS volumes with many files in a directory, then an interrogator timeout might occur resulting in a *Skipped directory* exception in the harvest audit. HDFS responds to StoredIQ slowly when it handles large directories and processes more responses from HDFS. The slow response from HDFS is caused by high CPU usage on HDFS NameNode. Therefore, if interrogator timeout occurs and high CPU usage on the HDFS server is observed, you can allocate more CPU resources to the HDFS server.

To avoid the interrogator timeout issue, you can also limit the file number to 250,000 files in a directory. Since each directory has its own timeout, having fewer files in a single directory ensures efficient operation. Splitting large directories into many small ones also helps resolve the interrogator timeout issues. For example, 1,000,000 files that are equally distributed into 10 directories have fewer risks of timeouts than if they are in one directory.

Table 23: Connections volumes: Add Volume dialog box options for primary volumes		
Connections: Add Volume dialog box options	**Action**	**Notes**
Server Type	Select **Connections** in the Server type list.	
Server	Enter the fully qualified domain name of the server from which the volume is available for mounting.	

Table 23: Connections volumes: Add Volume dialog box options for primary volumes (continued)

Connections: Add Volume dialog box options	Action	Notes
Class name	Enter `deepfile.fs.template.impl.ibmconnections.ibmconnectionsconn.IBMConnections`	Required
Repository name	Enter `deepfile.fs.template.impl.ibmconnections.ibmconnectionsconn`	Required
Option string		
Connect as	Enter the user name of the account that is set up with admin and search-admin privileges on the Connections server.	
Password	Enter the password of the account that is set up with admin and search-admin privileges on the Connections server.	
Volume	Enter any name.	
Initial Directory		
Index options	Select either or both of the **Index options** check boxes. • **Include system metadata for data objects within containers.** • **Include content tagging and full-text index.**	These options are selected by default.
Volume	In the **Volume** text box, enter the name or names of the volume to be mounted.	
Validation	To validate volume accessibility, select **Validation**.	When selected (the default state), IBM StoredIQ tests to see whether the volume can be accessed.
Include directories	In **Include directories**, specify a regular expression for included directories for each harvest (if it was specified).	These directories are defined as sets of "first node" directories, relative to the specified (or implied) starting directory that is considered part of the logical volume.

Table 23: Connections volumes: Add Volume dialog box options for primary volumes (continued)		
Connections: Add Volume dialog box options	**Action**	**Notes**
Start directory	In **Start directory**, designate a start directory for the harvest. The start directory involves volume partitioning to break up a large volume. If an initial directory is defined, the start directory must be underneath the initial directory. In the case of directories E-H, E is the start directory and H is the end directory.	
End directory	In **End directory**, determine the end directory for the harvest. The end directory is also part of volume partitioning and is the last directory harvested.	
Access Times	In **Access times**, select one of these options: • **Reset access times but do not synchronize them.** (This setting is the default setting.) • **Do not reset or synchronize access times.** • **Reset and synchronize access times on incremental harvests.**	

Table 23: Connections volumes: Add Volume dialog box options for primary volumes (continued)		
Connections: Add Volume dialog box options	**Action**	**Notes**
Constraints	In **Constraints**, select one of these options: • **Only use __ connection process (es)**: Specify a limit for the number of harvest connections to this volume. If the server is also being accessed for attribute and full-text searches, you might want to regulate the load on the server by limiting the harvester processes. The maximum number of harvest processes is automatically shown. This maximum number is set on the system configuration tab. • **Control the number of parallel data object reads**: Designate the number of parallel data object reads. • **Scope harvests on these volumes by extension**: Include or exclude data objects that are based on extension.	

Configuring SMB properties

Depending on the configuration of your SMB server, you might need to change the SMB settings in IBM StoredIQ Data Server to make the settings match.

To change the SMB settings on the data server to which you want to add the volume:

1. Using an SSH tool, log in to the data server as root.
2. Create a `jcifs.properties` file in the `/usr/local/siqsmb` folder by copying the original properties file. Use this command:

```
cp /usr/local/siqsmb/lib/jcifs.properties
 /usr/local/siqsmb/jcifs.properties
```

The properties file contains a set of SMB properties that you can use to adjust your SMB configuration. For information about additional properties for further configuration, contact IBM Support.

`jcifs.smb.client.ipcSigningEnforced`
Determines whether client signing for interprocess communication (IPC) connections is enforced. The default value is true.

This means that, although the data server is configured with signing not being required and not being supported, IPC connections to the SMB server still require signing by default. If the SMB servers does not support signing for IPC connections, the data server cannot connect to that server during volume creation unless you set this property to false.

Changed setting example: `jcifs.smb.client.ipcSigningEnforced=false`

jcifs.smb.client.listCount
Determines the maximum number of directories and files to be returned with each request of the TRANS2_FIND_FIRST/NEXT2 operation. The default value is 200.

Depending on the characteristics of the files and directories on the SMB server, you might want to adjust this value for performance reasons. For example, on higher latency networks a lower value can result in better performance.

Changed setting example: `jcifs.smb.client.listCount=300`

jcifs.smb.client.signingPreferred
Enables SMB signing if available. The default value is false.

If the SMB server requires SMB signing, you must set this property to true to have the data server as a JCIFS client negotiate SMB signing with that SMB server. If the SMB server does not require SMB signing but supports it, you can set this property to true for signing to occur. Otherwise, SMB signing is disabled.

Changed setting example: `jcifs.smb.client.signingPreferred=true`

jcifs.smb.client.signingEnforced
Determines whether client signing in general is enforced. The default value is false.

If the SMB server does not require and does not support signing, setting this property to true causes the connection to the SMB server to fail. Only set this property to true if one of these security policies is set on the SMB server:

- Microsoft network server: Digitally sign communication (always)
- Microsoft network server: Digitally sign communication (if client agrees)

Changed setting example: `jcifs.smb.client.signingEnforced=true`

jcifs.smb.client.enableSMB2
Enables SMB2 support. The default value is true.

Changed setting example: `jcifs.smb.client.enableSMB2=false`

jcifs.smb.client.disableSMB1
Disables SMB1 support. The default value is false.

Changed setting example: `jcifs.smb.client.disableSMB1=true`

jcifs.smb.client.dfs.disabled
Disables Distributed File System (DFS) referrals. The default value is false.

In non-domain environments, you might want to set this property to true to disable domain-based DFS referrals. Domain-based DFS referrals normally run when the data server as a JCIFS client first tries to resolve a path. In non-domain environments, these referrals time out causing a long startup delay.

Changed setting example: `jcifs.smb.client.dfs.disabled=true`

jcifs.smb.client.connTimeout
Determines the connection timeout, that is the time period in milliseconds that the client waits to connect to a server. The default value is 35000.

Changed setting example: `jcifs.smb.client.connTimeout=70000`

jcifs.smb.client.soTimeout
Determines the socket timeout, that is the time period in milliseconds after which sockets are closed if there is no activity. The default value is 35000.

Changed setting example: `jcifs.smb.client.soTimeout =70000`

jcifs.smb.client.responseTimeout
Determines the timeout for SBM responses, that is the time period in milliseconds that the client waits for the server to respond to a request. The default value is 30000.

Changed setting example: `jcifs.smb.client.responseTimeout=60000`

`jcifs.smb.client.sessionTimeout`
Determines the timeout for SMB sessions, that is the time period in milliseconds after which the session is closed if there is not activity. The default value is 35000.

Changed setting example: `jcifs.smb.client.sessionTimeout=70000`

3. Edit the `/usr/local/siqsmb/jcifs.properties` file.
4. Locate the property that you want to change, uncomment it, and set the appropriate value.
5. Restart services using this command: **service deepfiler restart**
6. Exit the data server.

Adding an SMB1 server as primary volume

In IBM StoredIQ, SMB1 and SMB2 are enabled by default, with the first choice being SMB2 connections. If the CIFS server that you add as a primary volume does not support SMB2, SMB1 is used. However, if the server supporting SMB1 only does also not provide client signing, you must disable IPC client signing on the data server on which the respective CIFS volume is defined by setting the `jcifs.smb.client.ipcSigningEnforced` property to false:

```
jcifs.smb.client.ipcSigningEnforced=false
```

Configuring Exchange 2007 Client Access Server support

The system supports the harvest of multiple Client Access Servers (CAS) when you configure Exchange 2007 primary volumes. This feature does not support redirection to other CAS/Exchange clusters or autodiscovery protocol.

1. Go to **Administration > Data sources > Volumes > Primary > Add primary volumes**.
2. In the **Server type** list, select **Exchange**.
3. In the **Version** list, select **2007**.
4. In the **Server** text box, type the name of the Exchange server. This server must be load-balanced at the IP or DNS level.
5. In the **Mailbox server:** text box, enter the name of one or more mailbox servers, which are separated by a comma and a space.
6. Complete the remaining fields for the primary volume, and then click **OK**.

Adding Domino as a primary volume

Domino volumes can be added as primary volumes.

1. Add a Lotus Notes user by uploading its **user ID** file in Lotus Notes User Administration on the **Administration > Configuration** tab.
   - If you want to harvest a user's mailbox, add the **user ID** file for that user.
   - If you want to harvest multiple mailboxes within one volume definition, add the administrator's ID file.
   - If the mailboxes have encrypted emails or NSFs, then you need each user's **user ID** file to decrypt a user's data.
2. Point the volume to the Domino server. If a single mailbox must be harvested, set the initial directory to be the path to the mailbox on the Domino server, such as `mail\USERNAME`.
3. To harvest mailboxes, select the **Harvest mailboxes** option, which obtains the list of all known Domino users and their NSFs. It then harvests those mailboxes unless it was pointed to a single mailbox by using the initial directory.
4. To harvest all mail journals, select the **Harvest mail journals** option.
5. To harvest all mail applications, select the **Harvest all applications** option, which looks at all NSFs, including mail journals, on the Domino server.

Special note about SharePoint volumes

Certain fields must be configured when SharePoint volumes are added.

IBM StoredIQ supports the entire sites portion of a Sharepoint URL for the volume `/sites/main_site/sub_site` in the **Volume** field when you add a SharePoint volume. However, if the SharePoint volume URL contains spaces, then you must also use the **Server**, **Volume**, and **Initial directory** fields in the Add volume dialog box in addition to the required fields **Server type**, **Server**, **Connect as**, and **Password**. For example, the SharePoint volume with the URL `http://shpt2010.reglab5.local/sitestest/autoteamsite1/Attribute Harvest WikiPages Library/` would require the fields in the following table because of the spaces in the URL.

Table 24: SharePoint volumes as primary volumes: Fields and examples

Primary volume field	Example
Server	`shpt2010.reglab5.local`
Volume	`/sitestest/autoteamsite1`
Initial directory	`Attribute Harvest Wiki Pages Library`

Performance conditions for using versions

When you add a primary volume, you define the volume by setting certain properties. If a SharePoint volume is added, you have the option of indexing different versions of data objects on that volume.

Since most versions of any object share full-text content and attributes, the effort in processing them and maintaining an updated context for the version history of an object in the index is duplicated. Additionally, if you enable version feature on a SharePoint volume, the API itself causes extra overhead in fetching data and metadata for older versions.

- For each object, an extra API call must be made to get a list of all its versions.
- To fetch attributes for the older versions of an object, an API call must be made for each attribute that needs to be indexed.

Limitations of SharePoint volumes

For SharePoint volumes, these limitations and specific warnings must be carefully considered.

- If data that is essential to the functioning of the SharePoint site as an application is deleted, the site might become unusable. For example, if you delete stylesheets and forms, you might get errors when you try to display certain pages on the site.
- It is possible to delete objects that are normally not visible. For example, documents that are filtered by views might not be visible within SharePoint. Regardless of their visibility, these data objects are indexed during harvests and can appear in infosets when responsive.
- The **delete** action is supported only for files that are held in document libraries. Other SharePoint object types can be present in an infoset, and they are audited as an unsupported operation. No folders, sites, or document libraries are deleted.
- Delete occurs at the system-file level. It means that the deletion of an older version or of a contained object results in deletion of all versions and the containing file and all of the objects it contains. The volume index is updated to reflect this change.
- The **delete** action can prevent the deletion of items that were accessed or modified since the previous harvest. For SharePoint volumes, this option is ignored. Recently accessed or modified files are deleted.
- It is not possible to delete files that are currently checked out.

Configuring FileNet

By providing the configuration values for a FileNet domain, you are supplying the values that are needed to bootstrap into a domain.

Within IBM StoredIQ Data Server, the FileNet domain must be configured prior to any FileNet volumes being created.

Note: With regards to object storage for FileNet cluster set ups, use other storage mechanisms than database (BLOB) storage. Additionally, files larger than 100MB should not be written to database storage.

1. Go to **Administration > Data sources > Specify Servers > FileNet domain configurations**.
2. Click **Add new FileNet domain configuration**
   The FileNet domain configuration editor page appears.
3. In the FileNet domain configuration editor page, configure these fields:
   a) In the **Configuration name** text box, enter the configuration name for this server.
   b) In the **Server name** text box, enter the server name.
   c) In the **Connection** list, select the connection type.
   d) In the **Port** text box, enter the port number.
   e) In the **Path** text box, enter the path for this server.
   f) In the **Stanza** text box, enter the stanza information for this server.
4. Click **OK** to save your changes.

Configuring Chatter messages

Within Chatter, the default administrator profile does not have the **Manage Chatter Messages** permission, but the appropriate permissions are required to harvest private messages.

A user must have certain administrative permissions when that user account is used in the **Connect as** text box in Chatter. When you set up a Chatter user account to harvest and run actions against Chatter, you must use an account with the built-in system administrator profile. In general, however, these administrative permissions must be assigned to the account you use:

- API enabled
- Manager Chatter Messages (required if you want to harvest Chatter Private Messages)
- Manage Users
- Moderate Chatter
- View All Data
- For Chatter administrators who use the **Auth token** option, see how to set up a sandbox account.

Adding a Documentum server as a data source

A Documentum server can be added as a data source and used as any other primary volume.

1. Using an SSH tool, turn on the Documentum license with these commands:
   a) `psql -U dfuser -d dfdata -c "update productlicensing set pl_isactive = true where pl_product = 'documentum';"`
   b) `service deepfiler restart`
2. In a browser, log in to the IBM StoredIQ data server.
3. Resolve the server name. Click **Administration** > **Configuration** > **Network settings** > **Server name resolution**.
   a) Enter the Doc broker settings. In the **Host** area, enter the Documentum host name, such as `DOCUMENTUMTEST1.local.com:1399`
      If you have more than one host, enter a single host per line. IP addresses can also be used.
   b) Click **OK**.

c) Using an SSH tool, connect to your target data server and edit `/etc/hosts`. Type `<ip of dataserver> <hostname entered in step 4a>`

Note: Remove the < > characters when entering the lines. Use one entry line per hostname.

4. Restart services by using either of these methods:
   a) Click **Administration** > **Dashboard** > **Controller**. Scroll to the bottom of the page and click **Restart services**.
   b) Using an SSH tool, enter `service deepfiler restart`
5. Add Documentum as a primary volume.
   a) Click **Administration > Data sources > Volumes > Add primary volumes**. The **Add volume** dialog box appears.
   b) In the **Server type** list, select **Documentum**.
   c) In the **Doc base** text box, enter the doc base name. Note: Obtain this information from the customer.
   d) In the **Connect as** text box, enter the connection name. Note: Obtain this information from the customer.
   e) In the **Password** text box, enter the password for the Documentum server. Note: Obtain this information from the customer.
   f) In the **Volume** text box, enter the name of the Documentum volume. Note: Obtain this information from the customer.
   g) Optional: Check the **Harvest all document versions** box if you want to harvest all document versions.
   h) In the **Initial directory** text box, enter the name of the initial directory. Note: Obtain this information from the customer.
   i) The Index options are preselected.
   j) Click **OK**.

Documentum is now available as a data source within IBM StoredIQ.

IBM Content Manager attributes

In the `SiqDocument` item type, various attributes are increased when you run copy to IBM Content Manager.

In the `SiqDocument` item type, the length of the following attributes is increased 128 - 256 bytes when you run copy to IBM Content Manager:

- `SiqServer`
- `SiqShare`
- `SiqInitialDirectory`
- `SiqFileName`
- `SiqContainerPath`
- `SiqOwner`

This change is handled automatically if you do not already have an `SiqDocument` item type in your IBM Content Manager server. However, if this item type exists, it must be recreated with the new attribute lengths for this change to take effect.

Note: If you run a working CopyTo IBM Content Manager without issues or if you know that your attribute lengths are not greater than 128 in length, then you can defer this action as you did not encounter the attribute length issue.

Note: The attribute length is in bytes. The number of actual characters this length holds varies based on the database code page that is used. For example, if ASCII is used, then the number of characters is equal to the number of bytes. If UTF-8 is used, the number of bytes per character varies depending on the

characters. Without this change, you see errors if the source attributes for a copy to IBM Content Manager exceed 128 bytes. If you see these errors, you need to take the following actions.

To extend the length of these attributes, take the following actions:

- If the `SiqDocument` item type does not exist in the IBM Content Manager server, create a new IBM Content Manager volume with a CopyTo option to select `SiqDocument`. It creates the item type and its attributes with the correct lengths.
- If `SiqDocument` item type exists in the IBM Content Manager server and you need to fix the attribute length problem, then the administrator must delete or drop the `SiqDocument` item type and recreate or update the IBM Content Manager volume that is used for copy. It automatically creates the item types and attributes desired.

  Note: Take a backup of the database before you drop and recreate `SiqDocument` item type. When you drop the item type, you permanently lose all the items (documents) stored in it. If the source documents are still available, you can run copy again to copy the data back into this item type. If no items exist in the item type, then it is not an issue.

  To drop the `SiqDocument` item type,

  1. Delete all the items in the `SiqDocument`. This delete is permanent and you lose all of the data.
  2. Delete the `SiqDocument` item type.
  3. Delete all the attributes that belong to the `SiqDocument`

If the item type exists and contains data that you need to keep, and you need to extend these attributes, this process is possible through direct database manipulation. However, this process is not supported and issues that derive from it cannot be covered by IBM support. If you want this process, services must be employed to make these database changes.

Creating retention volumes

Retention volumes store data objects that are placed under retention. It means that the object is retained.

1. Configure your retention servers.
2. Create management or retention classes, or, if you are using Centera retention servers, create Centera pools.
3. If you are using Enterprise Vault, ensure that you defined Enterprise Vault Sites. You might also want to define Retention Categories on the Enterprise Vault server.
4. Create retention volumes.

Adding a retention volume

Retention volumes can be added and configured. Applicable volume types include Enterprise Vault, CIFS (Windows platforms), NFS v3, Centera, and Hitachi.

1. Go to **Administration > Data sources > Volumes**, and then click **Retention**.
2. Depending on the type of retention server you are adding, complete the fields as described in the appropriate table.

   Individual tables describe the Enterprise Vault, CIFS (Windows platforms), NFS v3, Centera, and Hitachi retention-volume options.
3. Click **OK** to save the volume.

   Note: Case-sensitivity rules apply. Red asterisks within the user interface denote required fields.

Table 25: Enterprise Vault retention volume: fields, required actions, and notes

Field name	Required action	Special notes
Server type	In the **Server type** list, select the server type.	

Table 25: Enterprise Vault retention volume: fields, required actions, and notes (continued)		
Field name	**Required action**	**Special notes**
Site	In the **Site** list, select the appropriate site.	
Vault store	In the **Vault store** list, select the wanted vault store for the volume.	
Matter archive name	In the **Matter archive name** text box, enter the name of the matter archive.	
Description	In the **Description** text box, enter a description.	
Default retention category	In the **Default retention category** list, select the default retention category.	
Retention category override	Select the **Allow retention category** to be overridden on policy.	
Constraints	In **Constraints**, select either or both of these options: • **Only use __ connection process(es)**: Specify a limit for the number of harvest connections to this volume. If the server is also being accessed for attribute and full-text searches, you might want to regulate the load on the server by limiting the harvester processes. The maximum number of harvest processes is automatically shown. This maximum number is set on the system configuration tab. • **Control the number of parallel data object reads**: Designate the number of parallel data object reads.	

Table 26: CIFS (Windows platforms) retention volume: fields, required actions, and notes		
Field name	**Required action**	**Special notes**
Server type	In the **Server type** list, select the server type.	

Table 26: CIFS (Windows platforms) retention volume: fields, required actions, and notes (continued)		
Field name	**Required action**	**Special notes**
Server	In the **Server** text box, assign the server a name.	
Connect as	In the **Connect as** text box, enter the login ID.	
Password	In the **Password** text box, enter the password for the login ID.	
Volume	In the **Volume** text box, enter the name or names of the volume to be mounted.	
Index options	Select either or both of the **Index options** check boxes. • **Include system metadata for data objects within containers.** • **Include content tagging and full-text index.**	These options are selected by default.
Constraints	In **Constraints**, select either or both of these options: • **Only use __ connection process(es)**: Specify a limit for the number of harvest connections to this volume. If the server is also being accessed for attribute and full-text searches, you might want to regulate the load on the server by limiting the harvester processes. The maximum number of harvest processes is automatically shown. This maximum number is set on the system configuration tab. • **Control the number of parallel data object reads**: Designate the number of parallel data object reads.	

Table 27: NFS v3 retention volume: fields, required actions, and notes		
Field name	**Required action**	**Special notes**
Server type	In the **Server type** list, select the server type.	

Table 27: NFS v3 retention volume: fields, required actions, and notes (continued)		
Field name	**Required action**	**Special notes**
Server	In the **Server** text box, assign the server a name.	
Volume	In the **Volume** text box, enter the name or names of the volume to be mounted.	
Index options	Select either or both of the **Index options** check boxes. • **Include system metadata for data objects within containers.** • **Include content tagging and full-text index.**	These options are selected by default.
Constraints	In **Constraints**, select either or both of these options: • **Only use __ connection process(es)**: Specify a limit for the number of harvest connections to this volume. If the server is also being accessed for attribute and full-text searches, you might want to regulate the load on the server by limiting the harvester processes. The maximum number of harvest processes is automatically shown. This maximum number is set on the system configuration tab. • **Control the number of parallel data object reads**: Designate the number of parallel data object reads.	

Table 28: Centera retention volume: fields, required actions, and notes		
Field name	**Required action**	**Special notes**
Server type	In the **Server type** list, select the server type.	
Pool	In the **Pool** list, select the StoredIQ pool profile name to provide access to a specific Centera pool.	

Table 28: Centera retention volume: fields, required actions, and notes (continued)		
Field name	**Required action**	**Special notes**
Volume	In the **Volume** text box, enter the name or names of the volume to be mounted.	
Index options	Select either or both of the **Index options** check boxes. • **Include system metadata for data objects within containers.** • **Include content tagging and full-text index.**	These options are selected by default.
Constraints	In **Constraints**, select either or both of these options: • **Only use __ connection process(es)**: Specify a limit for the number of harvest connections to this volume. If the server is also being accessed for attribute and full-text searches, you might want to regulate the load on the server by limiting the harvester processes. The maximum number of harvest processes is automatically shown. This maximum number is set on the system configuration tab. • **Control the number of parallel data object reads**: Designate the number of parallel data object reads.	

Table 29: Hitachi retention volume: fields, required actions, and notes		
Field name	**Required action**	**Special notes**
Server type	In the **Server type** list, select the server type.	
Server	In the **Server** text box, assign the server a name.	
Volume	In the **Volume** text box, enter the name or names of the volume to be mounted.	

Table 29: Hitachi retention volume: fields, required actions, and notes (continued)		
Field name	**Required action**	**Special notes**
Index options	Select either or both of the **Index options** check boxes. • **Include system metadata for data objects within containers.** • **Include content tagging and full-text index.**	These options are selected by default.
Constraints	In **Constraints**, select either or both of these options: • **Only use __ connection process(es)**: Specify a limit for the number of harvest connections to this volume. If the server is also being accessed for attribute and full-text searches, you might want to regulate the load on the server by limiting the harvester processes. The maximum number of harvest processes is automatically shown. This maximum number is set on the system configuration tab. • **Control the number of parallel data object reads**: Designate the number of parallel data object reads.	

Configuring Enterprise Vault retention volumes

Enterprise Vault retention volumes can be configured on an as-needed basis.

1. Go to **Administration > Data sources > Volumes > Retention**.
2. Click **Add retention volumes**.
3. In the **Server type** list, select **Enterprise Vault**.
4. In then**Site** list, select the Enterprise Vault site you created. See Enterprise Vault.
5. Enter the information described in Creating Retention Volumes, based on your server type.
6. Click **OK** to save the volume.
7. Select one of the following:
   - **Add another volume on the same server**
   - **Add another volume on a different server**
   - **Finished adding volumes**

Creating system volumes

System volumes support volume export and import. When you export a volume, data is stored on the system volume. When you import a volume, data is imported from the system volume.

1. Go to **Administration > Data sources > Specify volumes > Volumes**.
2. Select the **System** tab, and then click **Add system volumes**.
3. Enter the information described in the table below, and then click **OK** to save the volume.

   Note: Case-sensitivity rules apply. Red asterisks within the user interface denote required fields.

Table 30: System volume fields, descriptions, and applicable volume types

Field name	Required action	Special notes	Applicable volume type
Server type	Using the **Server type** list, select the type of server.		• CIFS (Windows platforms) • NFS v2, v3
Server	In the **Server** text box, enter the name of the server where the volume is available for mounting.		• CIFS (Windows platforms) • NFS v2, v3
Connect as	In the **Connect as** text box, enter the logon ID used to connect and mount the defined volume.		• CIFS (Windows platforms)
Password	In the **Password** text box, enter the password used to connect and mount the defined volume.		• CIFS (Windows platforms)
Volume	In the **Volume** text box, enter the name of the volume to be mounted.		• CIFS (Windows platforms) • NFS v2, v3

Table 30: System volume fields, descriptions, and applicable volume types (continued)			
Field name	**Required action**	**Special notes**	**Applicable volume type**
Constraints	To utilize **Constraints**, select this option: • **Only use __ connection process (es)**: Specify a limit for the number of harvest connections to this volume. If the server is also being accessed for attribute and full-text searches, you may want to regulate the load on the server by limiting the harvester processes. The maximum number of harvest processes is automatically shown. This maximum number is set on the system **Configuration** tab.		• CIFS (Windows platforms) • NFS v2, v3

Export and import of volume data

Metadata and full-text indexed data can be collected or exported from separate locations, such as data servers in various offices of the enterprise. When the data is available, it can be imported to a single location, such as a headquarter's office data server, where selected files might be retained.

Only primary and retention volume data can be exported or imported with the export and import feature. Discovery export and system volumes cannot be imported or exported. The target location of an export or the source location of an import is always the IBM StoredIQ system volume.

Export and import volume processes can be run as jobs in the background. These jobs are placed into their prospective queues, and they are run sequentially. When one job completes, the next one automatically starts. These jobs can be canceled at any time while they are running. Canceling one import or export job also cancels all the jobs that come after the one canceled. Because the export jobs and import jobs are in separate queues, canceling one type of job does not cancel jobs in the other queue. The jobs cannot be restarted.

Exporting volume data to a system volume

When volume data is exported to a system volume, the export process creates two files: a binary file and a metadata file, which contains the exported data.

The export process creates two files: a binary file and a metadata file, which contains the exported data. These files' names contain the following information:

- Data server name and IP address
- Volume and server names
- Time stamp

Note the following information:

- The exported data consists of data from the selected volume and any related information that describes that data except for volume-specific audits.
- The exported data must be made available to the import data server before it can be imported. It might require you to physically move the exported data to the system volume of the import data server.
- Licenses on the import appliances are enabled automatically if a feature of the imported volume requires it (such as Centera or Exchange licenses).

1. Go to **DSAdmin > Administration > Data sources > Volumes**.
2. Select a volume of data to export from the list of volumes by clicking the **Discovery export** link in the far right column.
3. Complete the **Export volumes details** window, described in this table.
4. Click **OK**. A dialog box appears to show that the data is being exported.
5. To monitor the export progress, click the **Dashboard** link. To cancel the export process, under the Jobs in progress section of the dashboard, click **Stop this job**. Alternately, click **OK** to return to the **Volumes** page.

   Note: The job cannot be restarted.

Option	Description
Server	The name of the server where the data is.
Volume	The name of the volume where the data is.
Export path (on system volume)	This path is where to save the data on the system volume. The default path is `/exports`. You can edit the export path. The specified location is automatically created if necessary.
Description (optional)	Enter a description of the exported data.
Export full-text index	On a data server of the type DataServer - Classic, you can select this option to export the volume's full-text index. This option is available only if the volume has a full-text index. On a data server of the type DataServer - Distributed, this option is always shown and is checked by default because the full-text index is a part of the Elasticsearch index in general. You cannot change the setting.

Importing volume data from a system volume

An imported volume looks, acts, and is, just like a volume that is originally defined and harvested on the data server with a few key differences.

For imported volumes, any action or relationship that is valid for a non-imported volume is valid for an imported volume, with a few exceptions:

- Logs and audit trails that capture the activity on the volume before the import is not available. However, the import itself is audited.
- The imported volume can be reharvested if the appliance has the proper network access and rights to the original source server and volume.
- The imported volume can be reharvested if the data server has the proper network access and rights to the original source server and volume.
- The data viewer works only if the appliance has the proper network access and rights to the source server and volume. You must have access and permission on export servers and volumes if the file you want to view was migrated to a secondary server at the time of the export.

Notes:

- When a volume with a licensed feature is imported into a data server that does not use licensing, the license is imported along with the volume. To see the licensed features, users need to log out and then log back in to the data server.
- Imports can only occur between the same type of data server. A volume export from a data server of the type DataServer - Classic cannot be imported into a data server of the type DataServer - Distributed, and vice versa.
- If the volume that you want to import into a data server of the type DataServer - Distributed already exists, delete the volume from the data server before importing it again.

1. Make sure that exported data file is present in the system volume of the import appliance.
2. Go to **Administration > Data sources > Volumes**, and then click either the **Primary** or **Retention** tab.
3. Click the **Import volume** link at the top of either the primary or the retention volume lists. The **Import volumes** page appears, listing all of the volumes available for import. By default, the data server searches for available volumes to import in the `/imports` directory of the system volume. If you placed the exported data to another path, click **Change path** and enter the appropriate path.
4. Click **OK**. The Import volumes page now displays the following information about the imported volumes. See the following table.
5. From the list of volumes, select a volume to import by clicking the **Import** link in the last column on the right.
6. Select the **Import full-text index** check box to import the selected volume's full-text index.

   On a data server of the type DataServer - Classic, this option is active only if the volume has a full-text index.

   On a data server of the type DataServer - Distributed, this option is always shown and is checked by default because the full-text index is a part of the Elasticsearch index in general. You cannot change the setting.
7. Select the **Overwrite existing volume** check box to replace the existing data of the volume with the imported data.
8. Click **OK**. A dialog box appears to show that the volume is being imported. If the volume exists and unless the **Overwrite existing volume** option is selected. **OK** is not enabled .
9. To view import progress, click the **Dashboard** link in the dialog box. To cancel the import process, under the **Jobs in progress** section of the dashboard, click **Stop this job**, or click **OK** to return to the **Manage volumes** page.

Option	Description
Server and volume	The server name and volume name where the data physically is.
Description	A description added when the volume was exported.
Volume type	The types of volume, which are Exchange, SharePoint, and other types.
Category	The category of volume, which is Primary or Retention.
Exported from	The server name and IP address of the server from which the data was exported.
Export date	The day and time the data was exported.
Total data objects	The total number of data objects that are exported for the volume.
Contains full-text index	Indicates whether the full-text index option was chosen when the data was exported. For the index type **distributed**, this field will always show the value **Yes**.
Index type	Indicates which type of data server generated the export. You can import only volumes with an index type that matches the type of data server you are working with.

Creating discovery export volumes

Discovery export volumes contain the data produced from a policy, which is kept so that it can be exported as a load file and uploaded into a legal review tool. Administrators can also configure discovery export volumes for managing harvest results from cycles of a discovery export policy.

1. Go to **Administration > Data sources > Specify volumes**, and then click **Volumes**.
2. Click **Discovery export**, and then click **Add discovery export volumes**.
3. Enter the information described in the table below, and then click **OK** to save the volume.

   Note: Case-sensitivity rules for each server type apply. Red asterisks within the user interface denote required fields.

Table 31: Discovery export volumes: fields, required actions, and applicable volume types.

Field	Required action	Applicable volume type
Server type	Using the **Server type** list, select the type of server.	• CIFS (Windows platforms) • NFS v2, v3
Server	In the **Server** text box, enter the name of the server where the volume is available for mounting.	• CIFS (Windows platforms) • NFS v2, v3
Connect as	In the **Connect as** text box, enter the logon ID used to connect and mount the defined volume.	• CIFS (Windows platforms)
Password	In the **Password** text box, enter the password used to connect and mount the defined volume.	• CIFS (Windows platforms)
Volume	In the **Volume** text box, enter the name of the volume to be mounted.	• CIFS (Windows platforms) • NFS v2, v3
Constraints	To utilize **Constraints**, select this option: • **Only use __ connection process (es)**: Specify a limit for the number of harvest connections to this volume. If the server is also being accessed for attribute and full-text searches, you may want to regulate the load on the server by limiting the harvester processes. The maximum number of harvest processes is automatically shown. This maximum number is set on the system **Configuration** tab.	• CIFS (Windows platforms) • NFS v2, v3

Deleting volumes

Administrators can delete volumes from the list of available data sources, provided that the data server is connected to the gateway.

Note the following regarding deleted volumes:

- Deleted volumes are removed from target sets.
- Deleted volumes are removed from all volume lists, both from IBM StoredIQ and IBM StoredIQ Administrator.
- Within created jobs, steps that reference deleted volumes are implicitly removed, meaning that a job might contain no steps. The job itself is not deleted.
- Applicable object counts and sizes within IBM StoredIQ Administrator adjust automatically.
- Object counts and sizes within user infosets remain the same. Remember, those user infosets were created at a specific point in time when this data source was still available.
- Users who explore a specific data source and any generated reports no longer reference the deleted volume.
- No exceptions are raised on previously run actions. Instead, the data is no longer available. For example, if an infoset is copied that contained data objects from a volume that was 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 checks in, it sees that change in status and uninstalls itself.

Note: If retention volumes such as Centera, Hitachi contain data, they cannot be deleted as IBM StoredIQ is the source of record. Instead, you see the **Under Management** link.

1. Go to **Administration > Data sources > Specify volumes > Volumes**.
2. Click the tab of the volume type that you want to delete: **Primary**, **Retention**, **System**, or **Discovery export**.
3. Click **Delete**, and in the confirmation dialog window, click **OK**.
   The volume is deleted, removing it from the list of available volumes.

Action limitations for volume types

Volume types have different action limitations.

IBM StoredIQ imposes some action limitations on volume types, which are identified in this table.

Table 32: Action limitations for volume types		
Action type	**Source limitations**	**Other restrictions**
Copy from	• Centera • Hitachi • CIFS • NFS • Box • HDFS • Documentum • SharePoint • Exchange • Enterprise Vault (Discovery Accelerator) • NewsGator • OpenText Livelink • FileNet • Jive • Chatter • Domino • CMIS • IBM Content Manager • Desktop • Connections • OneDrive	Copy from SharePoint Online supports copy to CIFS, FileNet, and SharePoint 2007/2010/2013/2016/Online target sets. Copy from Box supports copying files to CIFS, NFS, FileNet, and Box target sets. Copy from Connections supports copying files to CIFS targets. Copy from OneDrive for Business supports copying files to CIFS and NFS shares.

Table 32: Action limitations for volume types (continued)

Action type	Source limitations	Other restrictions
Copy to	Primary Volume • Box • CIFS • NFS • HDFS • SharePoint • Documentum • FileNet • IBM Content Manager • CMIS	**Note:** For copy policies that target SharePoint volumes, the destination can write documents into an existing custom library. Two known SharePoint library configuration requirements are as follows. • For the **Require documents to be checked out before they can be edited** option, select **No**. This option is found within **Versioning settings**. • Each column that is marked as **Required** must specify a default value. In the **Author: Edit Default Value** dialog box, select the **Use this default value** option, and then supply a default value. This option is found within **Common default value settings**. Both configuration options are available within the Library settings of SharePoint.
Copy to (Box volume)	• CIFS • NFS • SharePoint	Box volumes can be added only from IBM StoredIQ Administrator, not from IBM StoredIQ Data Server. **Preserve owners** is only supported for SharePoint and CIFS source volumes. **Map permissions** is only supported for CIFS source volumes.
Copy to (OneDrive volume)	• CIFS • NFS	OneDrive volume can be added only from IBM StoredIQ Administrator, not from IBM StoredIQ Data Server. For copy policies that target OneDrive volumes, the destination folder needs to be set to the name of an existing drive at least. It cannot be left blank.
Copy to (retention)	• Centera • CIFS • NFS • Hitachi HCAP • Symantec Enterprise Vault	

Table 32: Action limitations for volume types (continued)

Action type	Source limitations	Other restrictions
Copy to SharePoint Online	• CIFS • FileNet • SharePoint 2007/2010/2013/2016/Online	**Note:** The Copy to SharePoint Online function does not include the ability to update some file system metadata or custom metadata. For example, the target file's Create Time and Modified Time reflect the time of the copy, not the source file's Create Time and Modified Time.
Delete	• Box • Centera • CIFS • Hitachi • NFS • HDFS • Documentum • SharePoint • Desktop	
Discovery export from	• Box • Centera • Hitachi • CIFS • NFS • HDFS • Documentum • SharePoint • Exchange • Enterprise Vault (Discovery Accelerator) • NewsGator • OpenText Livelink • FileNet • Jive • CMIS • IBM Content Manager • Desktop • Connections • OneDrive for Business	
Discovery export to	• CIFS • NFS	Considered discovery export volumes (category) not harvested

Table 32: Action limitations for volume types (continued)		
Action type	**Source limitations**	**Other restrictions**
Move from	• Box • Centera • CIFS • Hitachi • NFS • Documentum • Desktop	Move from Box supports copying files to CIFS, NFS, FileNet, and Box target sets.
Modify security	• CIFS • NFS	
Move to	• Centera • CIFS • Hitachi • NFS • Documentum • SharePoint • FileNet	

Volume limitations for migrations

The IBM StoredIQ supported source-infoset members for migrations to SharePoint targets have limitations.

For migrations that preserve version hierarchies, IBM StoredIQ preserves them at the file system level. If container-member objects are present in the source infoset, they are copied to new version chains, independent of the source-container version. Therefore, member-version numbers might not match the parent-container version.

Data harvesting

Harvesting or indexing is the process or task by which IBM StoredIQ examines and classifies data in your network.

Running a **Harvest every volume** job indexes all data objects on all volumes.

- A full harvest can be run on every volume or on individual volumes.
- An incremental harvest only harvests the changes on the requested volumes

These options are selected when you create a job for the harvest. A harvest must be run before you can start searching for data objects or textual content. An Administrator initiates a harvest by including a harvest step in a job.

Most harvesting parameters are selected from the **Configuration** subtab. You can specify the number of processes to use during a harvest, whether a harvest must continue where it left off if it was interrupted and many other parameters. Several standard harvesting-related jobs are provided in the system.

Harvesting with and without post-processing

You can separate harvesting activities into two steps: the initial harvest and harvest post-processing. The separation of tasks gives Administrators the flexibility to schedule the harvest or the post-process loading to run at times that do not affect system performance for system users. These users might, for example, be running queries. Examples of post-harvest activities are as follows:

- Loading all metadata for a volume.
- Computing all tags that are registered to a particular volume.
- Generating all reports for that volume.
- If configured, updating tags, and creating explorers in the harvest job.

Incremental harvests

Harvesting volumes takes time and taxes your organization's resources. You can maintain the accuracy of the metadata repository quickly and easily with incremental harvests. With both of these features, you can ensure that the vocabulary for all volumes is consistent and up to date. When you harvest a volume, you can speed up subsequent harvests by only harvesting for data objects that were changed or are new. An incremental harvest indexes new, modified, and removed data objects on your volumes or file servers. Because the harvests are incremental, it takes less time to update the metadata repository with the additional advantage of putting a lighter load on your systems than the original harvests.

Note: Harvesting NewsGator Volumes: Since NewsGator objects are just events in a stream, an incremental harvest of a NewsGator volume fetches only new events that were added since the last harvest. To cover gaps due to exceptions or to pick up deleted events, a full harvest might be required.

Harvest of properties and libraries

When harvesting private information, SharePoint volumes must use administrative roles for mounting permission.

Note: Administrative permissions are required to harvest personal information, libraries, and objects that are not designated as being visible to **Everyone** for user profiles.

The SharePoint volume needs to be mounted with administrative permissions. If the harvest is conducted without administrative permissions, then any of the user profile's properties that were marked as visible to the category other than **Everyone** is not visible in results. To harvest users' personal documents and information, volumes that are mounted without administrative permissions must use credentials that

have full control on all SharePoint site collections. These collections are hosted by the user profile service application.

To override this restriction, see http://technet.microsoft.com/en-us/library/ee721057.aspx.

Lightweight harvest parameter settings

To conduct a lightweight harvest, certain configuration changes can be made.

With IBM StoredIQ, you can conduct many types of harvests, depending on your data needs. While in-depth harvests are common, instances exist, where you need an overview of the data and a systemwide picture of files' types and sizes. For example, at the beginning of a deployment, you might want to obtain a high-level view of a substantial amount of data. It helps make better decisions about how you want to handle harvesting or other policies in the future. The following section provides possible system configurations for the system to process the volumes' data in the quickest manner possible.

Determining volume configuration settings

To conduct a lightweight harvest, you can make various configuration changes.

Volume Details: When you configure data sources for a lightweight harvest, you do not need to include content tagging and full-text indexes. By clearing this option, the system indexes the files' metadata, not the entire content of those files. The system can then run and complete harvests quickly. You can obtain much information about file types, the number of files, the age of the files, file ownership, and other information.

1. Go to **Administration > Data sources > Specify volumes > Volumes**.
2. On the **Primary volume list** page, click **Add primary volumes** or **Edit** to edit an existing volume.
3. Verify that all of the **Index options** check boxes are cleared (some are selected by default).
4. Optional: Edit the advanced details.

   In some cases, you might want to reduce the weight of a full-text harvest. In these instances, you can adjust the processing that is involved with the various harvest configuration controls.

   Within volume configuration, the advanced settings are used to control what is harvested within the volume. By harvesting only the directory structures that you are interested in, you can exercise some control over the harvest's weight.

   Click **Show Advanced Details**.

   - **Include Directory**: If you want to harvest a subtree of the volume rather than the whole volume, then you can enter the directory here. It eliminates the harvest of objects that are not relevant to your project.
   - **Start Directory** and **End Directory**: You can select a beginning and end range of directories that are harvested. Enter the start and end directories.
   - **Constraints**: You can limit the files that are harvested through connection processes, parallel data objects, or scoping harvests by extension. For example, with the **Scope harvest on these volumes by extension** setting, you can limit the files that you harvest by using a set of extensions. If you want to harvest only Microsoft Office files, you can constrain the harvest to .DOC, .XLS, and .PPT files.
5. Click **OK** and then restart services.

Determining harvester configuration settings

When you conduct a lightweight harvest, you can make certain harvester configuration changes.

1. Determine **Skip Content Processing** settings.

   Note: This setting is relevant only for full-text harvests.

   You might have many files that are types for which you do not need the contents such as .EXE files. In these instances, you can add these file types to the list of files for which the content is not processed. There are two points to consider when to skip content processing:

- You do not spend time harvesting unnecessary objects, which can be beneficial from a time-saving perspective.
- Later, you have the option of viewing the content of the skipped files. It creates more work, reharvesting these skipped files.

2. Determine which **Locations to ignore**.

   There might be instances where large quantities of data are contained in subdirectories, and that data is not relevant to your harvest strategy. For example, you might have a directory with a tree of source code or software archive that is not used as a companywide resource. In these cases, you can eliminate these directories from the harvests by adding the directory to the **Locations to ignore**. These locations are not specific to a volume, but can instead be used for common directories across volumes.

3. Determine **Limits**.

   - **Maximum data object size**: This setting is only relevant for full-text harvests. In cases many large files, you might want to eliminate processing those files by setting the **Maximum data object size** to a smaller number. The default value is 1,000,000,000. You can still collect the metadata on the large files, so you can search for them and determine which files were missed due to the setting of this parameter.

4. Determine **Binary Processing**.

   If the standard processing cannot index the contents of a file, binary processing is extra processing that can be conducted. For lightweight harvests, the **Run binary processing when text processing fails** check box must be cleared as this setting is only relevant for full-text harvests.

Determining full-text settings

When you conduct a lightweight harvest, you can make full-text-index setting configuration changes.

1. Determine **Limits**.

   Limit the length of words to be harvested by selecting the **Limit the length of words index to __ characters** option. The default value is 50, but you can reduce this number to reduce the quantity of indexed words.

2. Determine **Numbers**.

   If there are large quantities of spreadsheet files, you can control what numbers are indexed by the system.

Determining hash settings

When you conduct a lightweight harvest, you can make certain hash-setting configuration changes.

Before you change any settings, review the information in “Configuring hash settings” on page 22.

- Determine the hash settings.

  A file hash is a unique, calculated number that is based on the content of the file. By selecting **Partial data object content**, you reduce the processing to create the hash. However, the two different data objects might create the same hash. It is a small but potential risk and is only relevant for full-text harvests.

Job configuration

Jobs have different functions, and depending on the job type that is chosen, different system outcomes occur. Several jobs that are ready for use are also included.

Jobs start tasks such as harvests. They can be run at the time of creation or scheduled to run at a designated future time and at regular intervals. Jobs consist of either a single step or a series of steps. The actions available at each step depend on the type of job that is being created. Several jobs are ready for use. They are described in this table.

Table 33: Out-of-the-box job types

Job	Description
Centera deleted files synchronizer	An unscheduled, one-step job that synchronizes the deleted Centera files. This job is in the **Library/Jobs** folder.
CIFS/NFS retention volume deleted files synchronizer	An unscheduled, one-step job that harvests the Windows Share/NFS retention volumes, looking for files that require removal because the physical file was deleted from the retention file system. This job is in the **Library/Jobs** folder.
Database Compactor	A scheduled job that helps to limit "bloat" (unnecessary storage usage) in the database. While this job runs, it must have exclusive, uninterrupted access to the database. Administrators can override this job by logging in and then proceed to use the system. This job is in the **Library/Jobs** folder.
Harvest every volume	An unscheduled, one-step job that harvests all primary and retention volumes. This job is in the **Workspace/Templates/Jobs** folder.
Hitachi deleted files synchronizer	An unscheduled, one-step job that harvests the Hitachi volumes, looking for files that require removal because the physical file was deleted from the file system. This job is in the **Library/Jobs** folder.
System maintenance and cleanup	A multistep job in the **Library/Jobs** folder. The system is configured to run a system maintenance and clean up job once a day and includes these items: • Email users about reports • Email Administrators about reports • Delete old reports • Delete old harvests • Load indexes • Optimize full-text indexes
Update Age Explorers	A one-step job in the **Library/Jobs** folder that recalculates these items: • Owner Explorer data for Access Date • Owner Explorer data for Modified Date • Created Date (API only) values

Creating a job

Creating a job is the first step to work with different job types.

1. From the **Folders** tab > **Workspace** folder, select **New > Job**.
2. Enter a unique job name.
3. In the **Save in:** list, select the appropriate folder, and the job is created.

4. Click **OK**. If you want to view the job and add steps, click **Yes**.
5. On the View job page, click **Add step** and select a step type from the list.
6. **For Run harvest** jobs, on the Specify harvest and load options page, configure the following options:
   - **Harvest these volumes**: Select a volume from the list.
   - **Harvest type**: Specify the type of harvest:
     - **Run a full harvest**: All data objects on this volume are indexed.
     - **Run an incremental harvest** (default): Only files or data objects that changed since the last harvest is indexed.
   - **Harvest and load scheduling**: You can separate harvest and load processes to limit resource use. Select one of the options:
     - Load indexes when harvest completes
     - **Load indexes with next nightly system services job**: This option delays the index loading to run with the next system-services job after the harvest is completed. The system-services job is scheduled to run at midnight by default.
   - **Run harvest only**: Select this option if you plan to load harvested data into indexes later.
   - **Load indexes only**: Select this option to load previously harvested data into indexes.
7. **Harvest sampling**: Select this option if you want to limit the harvest to a smaller sample. This option skips every second, third, 10th, or other number data object as entered in the text box.
8. **Harvest limits**: Limit the harvest by time or total number of data objects. Enter the number of minutes or number of data objects.
9. Click **OK**.

Creating a job to discover retention volumes

Jobs can be created with various types of steps, including how to discover CIFS retention volumes.

The Discovery Retention volume job link appears only if a retention volume is configured on the data server.

1. From the **Folders** tab > **Workspace** folder, select **New > Job**.
2. Enter a unique job name.
3. In the **Save in:** list, select the appropriate folder, and the job is created.
4. Click **OK**. If you want to view the job and add steps, click **Yes**.
5. On the View job page, click **Add step**.
6. Select **Discover Retention volumes**.
7. In the **Discover Retention volume** list, select the retention volume to be used for this job.
8. Enter the duration that the harvest runs in **Run harvest for __**, and then select **Minutes** or **Hours**.
9. Enter the number of data objects to be harvested in **Only harvest __data objects**.
10. Click **OK**.

Editing a job

Existing, created jobs can be edited.

1. From the **Folders** tab > **Workspace** folder, click the job that you want to edit. The **Job details** page opens.
2. Click **Edit job details**. The Edit job details dialog box appears. You to specify the time, date, and frequency for the job to run.

- In the **Time:** field, enter the time that the job must start, or click **Now** to populate the time field with the current time. If you did not specify all of the job steps, you might want to add some time.
- In the **Date:** field, enter the date on which to run the job, or click **Today** to populate the date field with the current date.

3. Using the options in the **Frequency** field, specify how often the job must run. If you select **None**, the job runs once, at the time and date provided.
4. Click **OK**.
5. To edit the job steps:
   a) Add a step to by job by clicking **Add step**.
   b) Edit an existing step by clicking **Edit**.
   c) Remove an existing step by clicking **Remove**.
   d) Change the order of existing steps by clicking **Move up** or **Move down** icons.
6. Click **OK**.

Starting a job

Jobs can be started on an as-needed basis.

1. To start a job, do either of the following actions:
   - From the **Folders** tab > **Workspace** folder, click the name of the job you want to start, and in the Job details page, click **Start job**.
   - From the **Folders** tab > **Workspace** folder, right-click the job and select **Start job**.
2. In the **Schedule** area, a started job is displayed as **This job is running now**.
3. Click **OK** to return to the **Folders** tab. The started job displays **Running** in the **Status** column.

Saving a job

Jobs can be saved on an as-needed basis for future use.

1. From the **Folders** tab > **Workspace** folder, click the name of the job you would like to save. The Job details page opens.
2. Click **Save as**, and the Save job as dialog box appears.
3. In the **Job name** text box, verify that the job's name is correct. If it is saved in the same folder as another job, each job must have a unique name.
4. In the **Save in** list, select the folder in which you would like to save the job.
5. Click **OK** to close the Save job as dialog box and to return to the **Folders** tab.

Running a predefined job

Predefined jobs can be run on an as-needed basis.

1. Go to **Folders > Library**.
2. Click **Jobs** folder to open the job list.
3. Click the predefined job that you want to edit.
4. To set the schedule, click **Edit job details**, completed the changes, and then click **OK**.
5. Alternately, click **Start job** (selected jobs only) in the bottom right area of the pane to start the job immediately.

Deleting a job

Jobs can be deleted when they are no longer needed.

1. From the **Folders** tab > **Workspace** folder, select **Filter by... Jobs**.
2. Click the job name to open the job details.
3. Click **Delete** in the lower-left hand corner of the screen. Click **OK**.

Monitoring processing

You can track the system's processing on your harvest/policy and discovery export tasks with the **View cache details** feature. The appliance gathers data in increments and caches the data as it gathers it. If a collection is interrupted, the appliance can resume collection at the point that it was interrupted, instead of starting over from the beginning of the task.

From **Administration > Dashboard**, in the **Appliance status** pane, click **View cache details**. The View cache details page appears. To see harvest/policy progress, click the **Volume cache** tab. Or, to see discovery export job progress, click the **Discovery export cache** tab.

Note: Information for a job is only available while the job is running. After a task is completed, the job disappears from the list.

Table 34: Harvest/Volume cache details: Fields, descriptions, and values

Field	Description	Value
Name	The name of the volume that is being harvested.	
Start date	The time that the job started.	
Type	The type of job that is run.	• Copy • Harvest - full • Harvest - incremental
State	The status of the process.	• Caching: The volume cache is being created/updated by a harvest or policy • Cached: Creation or update of volume cache is complete (harvest only) • Loading: Volume cache contents are being successfully loaded into the volume cluster
Full-text	It indicates whether a full-text harvest is being conducted.	Yes or No
View audit link details	Link to the harvest/policy audit page.	

Table 35: Discovery export cache details: Fields, descriptions, and values

Field	Description	Value
Name	The name of the volume that is being processed.	

Table 35: Discovery export cache details: Fields, descriptions, and values (continued)		
Field	**Description**	**Value**
Start date	The starting date/time for the process.	
Type	Type of file that is being prepared for export.	Discovery export
State	The status of the discovery export job.	• Aborted: Discovery export policy was canceled or deleted by the user. • Caching: The volume cache is being created/updated by a harvest or policy. • Cached: Creation or update of volume cache is complete (harvest only). • Loading: Volume cache contents are being successfully loaded into the volume cluster.
Full-text	Whether a full-text harvest is being conducted.	Yes or No

Deleting a volume cache

Volume caches can be deleted as needed.

1. From the volume cache or discovery export cache list, select the check box of the cache you want to delete.
2. Click **Delete**. A confirmation dialog appears. Click **OK**.

Determining whether a harvest is stuck

The speed of a harvest depends on volume size and processing speed; however, harvests do occasionally become stuck and are unable to complete successfully. Use the procedures that are outlined here to troubleshoot the harvest process.

1. Click **Administration > Dashboard > Jobs in Progress** to verify that your job continues to run.
2. In **Jobs in Progress**, note the **Total data objects encountered** number.
3. Wait 15 minutes, leaving the harvest to continue to run.
4. Note the new value **Total data objects encountered**, and then compare it to that value denoted previously.
5. Answer the questions in this table.

Option	**Description**
Question	Action
Question 1: Is the Total data object encountered counter increasing?	• **Yes**: If the number of encountered data objects continues to increase, then the harvest is running correctly. • **No**: If the number of encountered objects remains the same, then go to Question 2.
Question 2: Is the load average up?	To view load averages, on **Appliance status > About appliance > View details > System services**, look at the load averages in the **Basic system information** area.

Option	Description
	• **Yes**: If the load averages number is up, the harvest might be stuck. Call technical support to report that the harvest is stuck on files. • **No**: The job is not really running. It means that the job must be restarted. Go to Question 3.
Question 3: Did the job complete on the second pass?	• **Yes**: If the job completed successfully after it was restarted, then the harvest is not stuck. • **No**: The job did not complete successfully. Call technical support to report a job that does not complete.

Desktop collection

When you configure desktop settings, you are enabling or disabling encryption within IBM StoredIQ. The IBM StoredIQ Desktop Data Collector (desktop client or client) enables desktops as a volume type or data source, allowing them to be used just as other types of added data sources. The IBM StoredIQ Desktop Data Collector is provided as a standard MSI file and is installed according to the typical method (such as Microsoft Systems Management Service (SMS)) used within your organization. The IBM StoredIQ Desktop Data Collector can collect PSTs, compressed files, and other data objects and can remove itself when its work is completed.

After the desktop client is installed on a desktop, you connect and register it with the data server. That desktop is available as a data source within the list of primary volumes. Additionally, while the snippet support and the Step-up Snippet action are supported by IBM StoredIQ Desktop Data Collector, a desktop cannot be the target or destination of an action.

IBM StoredIQ Desktop Data Collector client installation

The IBM StoredIQ Desktop Data Collector agent works with the following operating systems:

- Windows 7 32- and 64-bit
- Windows 8 32- and 64-bit
- Windows 10 32- and 64-bit
- Windows Server 2003, 2008, 2012, 2016

Note: For Desktop Agent with Windows Vista SP2 or Windows Server 2008 SP2, you must use Service Pack 2 and https://support.microsoft.com/kb/2763674/en-us. It is a required Microsoft update.

Installation requires administrative privileges on the desktop. Before you use the IBM StoredIQ Desktop Data Collector, you need to notify users that desktop collection is going to be conducted and make them aware of the following items:

- The desktop must be connected over the network during data collection. If the connection is interrupted, IBM StoredIQ Desktop Data Collector resumes its work from the point at which it stopped.
- Users might notice a slight change in performance speed, but that they can continue working normally. Desktop collection does not interfere with work processes.
- That certain actions can be taken from the tray icon: **Right-click for About**, **Restart**, **Status**, and **Email Logs** (which packages logs in to single file and starts the email client so that the user can mail them to the IBM StoredIQ administrator).

All communications are outbound from the client. The appliance never pushes data or requests to the desktop. The IBM StoredIQ Desktop Data Collector pings the appliance about every 60 seconds, and the **Last Known Contact** time statistic is updated approximately every 30 minutes. Additionally, the IBM StoredIQ Desktop Data Collector checks for task assignments every 5 minutes.

One can download the installer application from the application in the **Configuration** tab. Also, the Administrator can temporarily disable the client service on all desktops that are registered to the data server from the **Configuration** tab.

IBM StoredIQ Desktop Data Collector installation methods

During installation, the host name and IP address of the IBM StoredIQ must be supplied. If the installation is conducted manually by users, you must provide this information to them using email, a text file, or another method.

The IBM StoredIQ Desktop Data Collector can be installed with the following methods.

Mass distribution method (SMS)

The appliance ID is part of the distribution configuration. This method supports passing installation arguments as MSI properties.

- Required
  - **SERVERACTIONNODEADDRESS** IP address or host name for the Action node. When the installation is not silent, the user is prompted for IP address or host name. The default is the value of this argument. This field must be entered accurately or manual correction is required in the desktop configuration file.
- Optional
  - **SERVERACTIONNODEPORT** Port number for the Agent on the Action node. Defaults to 21000, and can be only changed when the agent connects on a different port that is then mapped to 21000.
  - **NOTRAYICON** Specifies whether the agent displays the IBM StoredIQ Desktop Data Collector tray icon while it is running. Changing the setting to **1** forces the agent to run silently and not display a tray icon.
  - **SERVERACTIONNODEADDRESS** IP address or host name for the Action node. When the installation is not silent, the user is prompted for IP address or host name. The default is the value of this argument. This field must be entered accurately or manual correction is required in the desktop config file.
  - **Emailing links** Send a link within an email such as `file:\\g:\group\install\Client-install.vbs`. The link can be to any executable file format such as .BAT, .VBS, or .MSI. The .BAT/.VBS formats can be used to pass client arguments to an .MSI file. The user who clicks the link must have administrative privileges.
  - **NT Logon Script** A .BAT file or .VBS script starts msiexec. Examples are given here:
    - `/i`: Install
    - `/x {7E9E08F1-571B-4888-AC08-CEA8A076F5F9}`: Uninstall the agent. The product code must be present.
    - `/quiet`: install/uninstall runs silently. When you specify this option, SERVERACTIONNODEADDRESS must be supplied as an argument.

```
Set WshShell = CreateObject("WScript.Shell")

WshShell.Run "%windir%\System32\msiexec.exe /i G:\group\install\
desktopclient.msi
NOTRAYICON=0 SERVERACTIONNODEADDRESS=clust017.test.local /q"Set WshShell
= CreateObject("WScript.Shell")

Set WshShell = Nothing
```

- **MSI**

`Batch file`

```
msiexec /i G:\group\install\desktopclient.msi NOTRAYICON=1
            SERVERACTIONNODEADDRESS=clust017.test.local /q
```

Configuring the IBM StoredIQ Desktop Data Collector collection

Desktop collection configuration includes the IBM StoredIQ Desktop Data Collector installer and Encrypted File System.

For procedural information on downloading the IBM StoredIQ Desktop Data Collector installer, see “Downloading the IBM StoredIQ Desktop Data Collector installer from the application” on page 24.

1. Complete the procedure that is outlined in Using the Encrypted File-System Recovery Agent.
2. Restart the service and begin collection, noting the following points:

- If the computer is not part of a domain and it is running any version of Windows earlier than 7.0, then the user name must be the user name.
- If the computer is not part of a domain and running Windows 7.0 or later, then the user name must be the name of the PC and the domain.

Delete policy: special notes

When you use the delete policy, certain files cannot be deleted.

When you use the IBM StoredIQ Desktop Data Collector to delete IBM StoredIQ Desktop Data Collector files from a desktop, they are removed permanently. They are not transferred to the appliance or backed up to any other location. You must carefully review the infoset of affected data objects before you take a delete action. Your organization can use custom applications or other files that you might not want to delete. In reviewing the returned list, do not allow the following files to be deleted.

- Anything under this directory: `c:\Windows`
- Anything under this directory: `Documents and Settings`, with these extensions:

```
c:\Documents and Settings\<username>\UserData\ and extension *.xml
```

```
c:\Documents and Settings\<username>\Cookies\ and extension *.txt
```

```
c:\Documents and Settings\<username>\Start Menu\Programs\ and extension *.lnk
```

- Executable files
  - *.dll
  - *.exe
  - *.ocx
- Drivers
  - *.sys
  - *.inf
  - *.pnf
- Installers
  - *.msi
  - *.mst
- Important data files
  - *.dat
  - *.ini
  - *.old
  - *.cat
- These file names
  - `desktop.ini`
  - `ntuser.dat`
  - `index.dat`
  - `ntuser.pol`
  - `ntuser.dat.log`

Folders

Both conceptual and procedural information about folders and their usage is described as follows.

The **Folders** tab displays two types of folders: **Library** and **Workspace**.

- **Library**. The **Library** folder contains the **Jobs** folder.

  Note: This folder cannot be renamed, moved, or deleted.
- **Workspace**. The **Workspace** folder is a custom folder that reflects your use of the system. By default, it contains a folder entitled **Templates**. If you are using the system for IT purposes, you might want to create folders for each locale or function.

  Note: These folders can be renamed, moved, or deleted, and you also have the options of setting folder security.

Creating a folder

Folders are created from the **Folders** tab.

1. From the **Folders** tab, select **New > New Folder**. The Create new folder dialog appears.
2. In the **Name:** field, give a name that represents the folder's purpose (legal matter, local, business unit, or the like).
3. In the **Description:** field, type a description for the folder.
4. In the **Create in:** field, use the list to select a place for the folder. All custom folders must be placed in the **Workspace** or a **Workspace** subdirectory; you cannot add folders to the Library folder.
5. Click **OK**. If you want to open the folder, click **OK** in the dialog that appears.

Deleting a folder

When you delete folders, you can delete only empty folders.

1. From within the **Folders** tab, **Workspace** folder, do either of the following actions.
   - Select the check box next to the folder you want to delete, and in the **Actions** list, select **Delete**.
   - Right-click on a folder name and select **Delete**.
2. In the confirmation box that appears, click **OK**.

Moving a folder

Folders can be moved from one location to another.

1. From within the **Folders** tab, **Workspace** folder, do either of the following actions.
   - Select the check box next to the folder you want to move, and in the **Actions** list, select **Move**.
   - Right-click on the folder name and select **Move**.
2. Click **OK**, and in the Move items dialog, select the new location from the list.
3. Click **OK**.

Renaming a folder

Folders can be renamed as needed.

1. Right-click on the folder name and select **Rename**.

2. In the Rename folder dialog, change the **Name** and the **Description**.
3. Click **OK**.

Copying items to different folders

Items can be copied from one folder to another.

1. From within **Folders > Workspace**, right-click the item that you want to copy, and then select **Copy**.
2. In the Copy dialog, assign a new name (if appropriate) in the **Name** field. You cannot reuse the same name for an item within a single folder.
3. In the **Description** field, type a description.
4. In the **Save in** list, choose the location for the copied item, and then click **Save**.

Saving items into different folders

You can rename jobs and save them into other folders.

1. From the item's editor pane, click **Save as**.
2. In the **Save [item] as...** pane, type a name in the **[Item] name:** field.
3. In the **Description:** field, type a description.
4. In the **Save in:** field, use the list to select a location for the item.
5. Click **Save**.

Filtering items within the folder view

Items can be filtered within the Folder view.

1. From within **Folders**, click **Filter by...**
2. Select the component that you want to display.

Audits and logs

The following section describes the audit and log categories in the system, including descriptions of the various audit types and how to view and download details.

Harvest audits

Harvest audits provide a summary of the harvest, including status, date, duration, average harvest speed, and average data object size. They can be viewed in two ways: by volume name or by the date and time of the last harvest.

Data objects can be skipped during a harvest for various reasons such as the unavailable object or a selected user option that excludes the data object from the harvest. The **Harvest details** page lists all skipped data objects that are based on file system metadata level and content level.

All skipped harvest-audit data and other files that are not processed can be downloaded for analysis.

Table 36: Harvest audit by volume: Fields and descriptions

Harvest audit by volume field	**Description**
Server	The server name.
Volume	The volume name.
Harvest type	The type of harvest that is conducted: Full Harvest, ACL only, or Incremental.
Last harvested	The date and time of the last harvest.
Total system data objects	The total number of system data objects encountered.
Data objects fully processed	The number of data objects that were fully processed.
Data objects previously processed	The number of data objects that were previously processed.
Processing exceptions	The number of exceptions that are produced during processing.
Binary processed	The number of processed binary files.
Harvest duration	The length of time of the harvest's duration.
Status	The harvest's status: **Complete** or **Incomplete**.
Average harvest speed	The average harvest speed, which is given in terms of data objects that are processed per second.
Average data object size	The average size of encountered data objects.

Table 37: Harvest audit by time: Fields and descriptions

Harvest audit by time field	**Description**
Harvest start	The time and date at which the harvest was started.
Harvest type	The type of harvest that is conducted: **Full Harvest**, **ACL only**, or **Incremental**.

Table 37: Harvest audit by time: Fields and descriptions (continued)	
Harvest audit by time field	**Description**
Total system data objects	The total number of system data objects that were found.
Data objects fully processed	The total number of system data objects that were fully processed.
Data objects previously processed	The total number of system data objects that were previously processed.
Processing exceptions	The total number of encountered processing exceptions.
Binary processed	The total number of processed binary files.
Harvest duration	The length of time of the harvest's duration.
Status	The harvest's status: **Complete** or **Incomplete**.
Average harvest speed	The average harvest speed, which is given in terms of data objects that are processed per second.
Average data object size	The average size of encountered data objects.

Table 38: Harvest overview summary options: Fields and descriptions	
Harvest overview summary options field	**Description**
Harvest type	The type of harvest: **Full Harvest**, **ACL only**, or **Incremental**.
Harvest status	The harvest's status. Options are **Complete** or **Incomplete**.
Harvest date	The date and time of the harvest.
Harvest duration	This duration is the length of time of the harvest's duration.
Average harvest speed	The average harvest speed, which is given in terms of data objects that are processed per second.
Average data object size	The average size of encountered data objects.

Table 39: Harvest overview results options: Fields and descriptions	
Harvest overview results options field	**Description**
Total system data objects	The total number of system data objects that were found.
Total contained data objects	The total number of contained data objects.
Total data objects	The total number of encountered data objects.

Table 40: Harvest overview detailed results: Fields and descriptions	
Harvest overview detailed results field	**Description**
Skipped - previously processed	The number of skipped objects that were previously processed.

Table 40: Harvest overview detailed results: Fields and descriptions (continued)	
Harvest overview detailed results field	**Description**
Fully processed	The number of fully processed data objects.
Skipped - cannot access data object	The number of data objects that were skipped as they might not be accessed.
Skipped - user configuration	The number of data objects that were skipped because of their user configuration.
Skipped directories	The number of data objects in skipped directories.
Content skipped - user configuration	The number of data objects where the content was skipped due to user configuration.
Content type known, partial processing complete	The number of data objects for which the content type is known and partial processing is complete.
Content type known, but error processing content	The number of data objects for which the content type is known, but an error was produced while processing content.
Content type known, but cannot extract content	The number of data objects for which the content type is known, but the content might not be extracted.
Content type unknown, not processed	The number of data objects for which the content type is unknown and is not processed.
Binary text extracted, full processing complete	The number of data objects for which the binary text is extracted and full processing is completed.
Binary text extracted, partial processing complete	The number of data objects for which the binary text is extracted and partial processing is completed.
Error processing binary content	The number of data objects for which an error was produced while binary content is processed.
Total	The total number of data objects.

Viewing harvest audits

Harvest audits can be viewed from the **Audit** tab.

1. Go to **Audit > Harvests > View all harvests**. The **Harvest audit by volume** page opens, which lists recent harvests and includes details about them.
2. In the **Volume** column, click the volume name link to see the harvest audit by time page for that particular volume. The harvest audit by time page lists all recent harvests for the chosen volume and includes details about each harvest.
3. In the **Harvest start** column, click the harvest start time link to see the harvest overview page for the volume. You can also access the page by clicking the **Last harvested time** link on the **Harvest audit by volume** page. The **Harvest overview** page provides the following options.
   - **Summary**: Harvest type, status, date and time, duration, average harvest speed, and average data object size.
   - **Results**: Total system data objects, total contained data objects, total data objects.
   - **Detailed results**: Skipped - previously processed; fully processed; skipped - cannot access data object; skipped - user configuration; skipped directories; content skipped - user configuration; content type known, partial processing complete; content type known, but error processing content; content type known, but cannot extract content; content type unknown, not processed;

binary text extracted, full processing complete; binary text extracted, partial processing complete; error processing binary content; error-gathering ACLs; and total.

4. To view details on data objects, click the link next to the data objects under **Detailed results**.
   - With the exceptions of skipped - previously processed, fully processed, and the total, all other results with more than zero results have links that you can view and download results.
   - The skipped data object list includes object name, path, and reason skipped. Data objects can be skipped at the file system metadata level or at the content level. Data objects skipped at the content level are based on attributes that are associated with the data object or its contents. Skipped Data Objects Results Details provides details about skipped data objects.

   If data objects were not harvested, you might want to download the data object's harvest audit list details for further analysis.

Downloading harvest list details

Harvest list details can be downloaded in a .CSV format.

1. From the **Harvest details** page, click the active link next to the data objects under **Detailed results**. A page named for the detailed result chosen (such as Skipped - user configuration or binary text extracted, full processing complete) appears.
2. Click the **Download list in CSV format** link on the upper left side of the page. A dialog informs you that the results are being prepared for download.
3. Click **OK**. A new dialog appears, prompting you to save the open or save the .CSV file. Information in the downloaded CSV file includes:
   - Object name
   - System path
   - Container path
   - Message explaining why data object was skipped
   - Server name
   - Volume name

Import audits

Volume-import audits provide information about the volume import. This information includes the number of data objects that are imported, the system that is exported from, the time and date of the volume import, whether the imported volume overwrote an existing volume, and status. The volume name links to the **Import details** page.

Table 41: Imports by volumes details: Fields and descriptions

Imports by volumes details field	Description
Volume	The name of the imported volume.
Exported from	The source server of the imported volume.
Import date	The date and time on which the import occurred.
Total data objects imported	The total number of imported data objects.
Overwrite existing	If the import overwrote an existing volume, the status is **Yes**. If the import did not overwrite an existing volume, the status is **No**.
Status	The status of the import: **Complete** or **Incomplete**.

Viewing volume import audit details

Audit details of volume imports can be viewed.

1. Go to **Audit > Imports**, and click **View all imports**. The **Imports by volume** page opens, which lists volume imports and import information.
2. Click a volume name link in the **Volume** column to view the audit details for that particular import.

Event logs

Event logs captures every action that is taken by the system and its users. It documents actions that succeed and fail.

These actions include creating draft and published queries and tags, running policies, publishing queries, deleting objects, configuring settings, and any other action that is taken through the interface. A detailed list of log entries is provided in the event log messages.

You can view event logs for the current day or review saved logs from previous days, and up to 30 days worth of logs can be viewed through the interface. If you select and clear a day of logs, those logs are removed from the system.

Viewing event logs

Event logs can be viewed from the **Dashboard** or from the **Audit** tab.

1. Conduct either of the following actions:
   a) Click **Administration > Dashboard**, and then locate the **Event log** section on the dashboard. The current day's log displays there by default.
   b) Click the **Audit** tab and locate the **Event logs** section.
2. To view a previous day's log on the dashboard, use the **View all event logs** list to select the day for which you want to view an event log.
3. Select a different day from the view event log from the list. This menu displays the event log dates for the past 30 days. Each log is listed by date in YYYY-MM-DD format.

Subscribing to an event

You can subscribe to and be notified of daily event logs.

1. Go to **Audit > Event logs**.
2. Click **View all event logs**, and the **Event log for today** page opens.
3. To the right of the event log to which you want to subscribe, click **Subscribe**. The Edit notification page appears.
4. In **Destination**, select the method by which you want to be notified of this event log. If you select **Email address**, be certain to use commas to separate multiple email addresses.
5. Click **OK**.

   Note: You can also subscribe to an event on the **Dashboard**. In the **Event log** area, click **Subscribe** to the right of the event.

Clearing the current event log

An event log can be cleared from the **Dashboard**.

On the **Administration > Dashboard**, click **Clear** for the current view.

Downloading an event log

Event logs can be downloaded and saved.

1. When you view an event log, click the **Download** link for saving the data to a text file.
2. Select to save the file from the prompt. Enter a name and select a location to save the file.

Policy audits

Policy audits provide a detailed history of the policy. It includes type of action, date last run, start and end dates with times, average speed, total data objects, and data object counts. They can be viewed by name, volume, time, and by discovery export.

Policy audit by name

Table 42: Policy audit by name: Fields and descriptions

Policy audit by name field	Description
Policy name	The policy name.
Policy status	The policy's status.
Number of times executed	The number of times that the policy was run.
Most recent date executed	The date on which the policy was last run.

Policy audit by volume

Table 43: Policy audit by volume: Fields and descriptions

Policy audit by volume field	Description
Volume	The name of the volume on which the policy was run.
Most recent date a policy was executed	The most recent date on which the policy was last run.
Number of policies executed	The number of policies that were run.

Policy audit by time

Table 44: Policy audit by time: Fields and descriptions

Policy audit by time field	Description
Policy name	The policy name.
Policy status	The status of the policy: Complete or Incomplete.
Start	The time at which the policy's execution was started.
End	The time at which the policy's execution was complete.
Success count	The number of processed messages that are classified as a success.
Failure count	The number of processed messages that are classified as a failure.
Warning count	The number of processed messages that are classified as a warning.
Other count	The number of processed messages that are classified as other.
Total data objects	The total number of data objects.
Action type	The type of policy that took place.

Table 44: Policy audit by time: Fields and descriptions (continued)	
Policy audit by time field	**Description**
Ag. actions/second	The average number of actions per second.

Policy audit by discovery export

Table 45: Policy audit by discovery exports: Fields and descriptions	
Policy audit by discovery export field	**Description**
Discovery export name	The name of the discovery export.
Number runs	The number of times the policy ran.
Most recent export status	The status of the most recent discovery export.
Most recent load file status	The status of the most recent load file.
Most recent date executed	The date of the most recent policy execution.

Discovery export runs by discovery export

Table 46: Discovery export runs by discovery export: Fields and descriptions	
Discovery export runs by discovery export field	**Description**
Discovery export run	The name of the discovery export run.
Number of executions	The number of times the run was started.
Success count	The number of processed messages that are classified as a success.
Failure count	The number of processed messages that are classified as a failure.
Warning count	The number of processed messages that are classified as a warning.
Other count	The number of processed messages that are classified as other.
Total data objects	The total number of data objects.
Export status	The status of the export: Complete or Incomplete.
Load file status	The status of the load file: Complete or Incomplete.

Note: A warning in a policy audit trail is a success with the following conditions:

- If you copy an Exchange item such as `re:`, the `re` is copied, not `the:`. It generates a warning.
- The copied file is renamed.
- The file system to which you are copying does not accept characters in the file name.

Viewing policy audit details

Policy audits can be viewed by name, volume, time, or discovery export.

1. Go to **Audit > Policies**, and then click **Name**. The **Policy audit by name** page provides policy name and status, the number of times it was run, and the time and date of the most recent execution.

2. Click a policy name to open the **Policy executions by time** page.
3. Click a policy name to open the **Policy execution results** page.

   Note: To view the list of data objects, click the **[#] data objects** link. To create a report, click **Create XML** or **Create PDF**.

   a) Click **Volume** to open the policy audit by volume page.

   b) Click a volume link to go to the **Policy audit by time** page.

   c) Click **Time** to see **Audit by time** page for the policy.

   d) On the **Policy audit by time** page, click the policy name to open the **Policy execution results** page. To view the list of data objects, click the **[#] data objects** link. To create a report, click **Create XML** or **Create PDF**.

   e) Click **Discovery export**.

   f) On the **Policy audit by discovery export** page, click the discovery export name to open the **Discovery export runs by production** page. The page details further information according to the incremental runs of the policy.

   g) Click a policy name to open the **Policy executions by time** page.

   h) Click a policy name to open the **Policy execution results** page. To view the list of data objects, click the **[#] data objects** link. To create a report, click **Create XML** or **Create PDF**.

   As you review audit results through the pages, you can continue clicking through to review various levels of information, from the volume and policy execution level down to the data objects. To view more policy execution details, click the policy name in the execution summary page, which can be accessed by any of the policy views. As you continue browsing, IBM StoredIQ provides more detailed information such as:

   - Source and destination settings
   - Policy options: Details of the policy action. This section reflects the options that are selected when you create the policy. Most attributes that appear depend upon the type of policy run and the options available in the policy editor.
   - Query (either IBM StoredIQ or user-defined)
   - **View metadata** link: The view metadata page describes security details for source and destination locations of the policy action.

Viewing a policy audit by name

Policy audits can be viewed by their name.

1. Go to **Audit > Policies**.
2. Click **Name**.
   The Policy audit by name page provides policy name and status, the number of times it is run, and the time and date of the most recent execution.
3. Click a policy name to open the Policy execution by time page.

Viewing a policy audit by volume

Policy audits can be viewed by volume.

1. Go to **Audit > Policies**.
2. Click **Volume**.
   The Policy audit by page provides policy name, its most recent time, date of execution, and the number of policies that were run.
3. Click a policy name to open the Policy executions by time page.
4. Click the policy name to open the Policy executions by results page.

Viewing a policy audit by time

Policy audits can be viewed by time.

1. Go to **Audit > Policies**.

2. Click **Time**.
   The policy audit by time provides information about the policy.
3. On the Policy audit by time page, click the policy name to open the Policy execution results page.

Viewing a policy audit by discovery export
Policy audits can be viewed by discovery export.

1. Click **Discovery export**.
2. On the Policy audit by discovery export page, click the discovery export name to open the Discovery export runs by production page. The page details further information according to the incremental runs of the policy.
3. Click a policy name to open the Policy executions by time page.
4. Click a policy name to open the Policy execution results page.

Search audit feature

With the search audit feature, you can search audit trails by entering either by **Policy Details**, **Execution Details**, or **Data Object Details**.

Policy details

Policy audits can be searched with any of these details.

Table 47: Policy audit details: Fields and descriptions

Policy detail field	Description
Audit search by policy details	In this area, select search criteria, define their values, and then add them to the list to search across all audits.
Specify search criteria	In this area, specify the **Policy name**, the **Policy state**, and the **Action type**.
Audit search criteria	In the **Find audits that match** list, select either **Any of the following** or **All of the following**.

Execution details

Policy audits can be searched with any of these execution details.

Table 48: Policy audit execution details: Fields and descriptions

Execution detail field	Description
Audit search by execution detail	In this area, select search criteria, define their values, and then add them to the list to search across all audits.
Specify search criteria	In this area, specify the **Action type, Action status, Action start date, Action end date, Total count, Success count, Failure count, Warning count, Source volume, Destination volume,** or **Query name.**
Audit search criteria	In the **Find audits that match** list, select either **Any of the following** or **All of the following**.

Data object details

Policy audits can be searched with any of these data-object details.

Table 49: Policy audit data object details: Fields and descriptions	
Data object details field	**Description**
Audit search by data object details	In this area, select search criteria, define their values, and then add them to the list to search across all audits.
Specify search criteria	In this area, specify the **Source volume, Destination volume, Source object name, Destination object name, Source system path, Destination system path,** or **Action result.**
Audit search criteria	In the **Find audits that match** list, select either **Any of the following** or **All of the following**.

Saving results from an audit

You can save the results of policy executions into PDF and XML files. The information can be saved as PDF and XML files. The exporting of information appears as a running job on the dashboard until completed.

1. Go to **Audit > Policies**.
2. In the **Browse by** options, click **Time**.
3. Click the policy name.
4. In the **Results** pane, click **Data objects** to see items that were responsive to the policy. To download the material in .CSV, click **CSV**.
5. On the **Policy execution results** page, select **Create PDF** to generate a PDF or **Create XML** to generate an XML file of the results.
6. Access the report through the inbox on the navigation page.

Policy audit messages

A policy audit shows the number of data objects that were processed during the policy execution. Processed data objects are divided into these categories: Success, Warnings, Failures, and Other (discovery export policies only).

Table 50: Types of and reasons for policy audit messages	
Audit message type	**Reason**
Success	• Data object is a duplicate of [object name] • Data object skipped but is loaded in load file. It applies to intermediate and files archives produced during a discovery export policy. • Data object is a duplicate produced in a previous run (discovery export only).

Table 50: Types of and reasons for policy audit messages (continued)

Audit message type	Reason
Warning	• Set directory attributes • Reset time stamps • Set attributes • Set time stamps • Set security descriptor (Windows Share) • Set access modes (Windows Share) • Set owner information • Set group information (NFS) • Set security permissions • Create a link after migration (Windows Share, NFS) • Find template to create a shortcut (Windows Share) • Extract text for the object (Discovery export policy)
Failure	• Failed to create target directory structure • Source does not exist • Failed to find a new name for the incoming object • Target is a directory • File copy failed • Cannot create target • Error copying data to target • Cannot copy due to network errors • Cannot delete source after move • Target disk is full • Source equals target on a copy or move • Insufficient permissions in general to conduct an action • All modify actions failed • File timed out waiting in the pipeline • File under retention; cannot be deleted (retention server) • Data object is a constituent of a container that already encountered failure (discovery export policy)
Other	Data objects are categorized in the other category during a discovery export policy when: • A data object is a member that makes its container responsive. • A data object is a non-responsive member of a container.

Deploying customized web services

This procedure highlights the basic steps that are required to deploy SharePoint custom web services.

This procedure applies only to SharePoint 2010, 2013, or 2016 web services.

1. Obtain the installation package. This package is created for you by IBM StoredIQ.
   Use a utility such as SCP to transfer the installation package. Depending on whether SharePoint 2010, 2013, or 2016 is used, the package located in either `/deepfs/downloads/webservice/sp2010/`, `/deepfs/downloads/webservice/sp2013/`, or `/deepfs/downloads/webservice/sp2016/`.
2. Uninstall an existing instance of the web service.
3. Install the installation package.
4. Verify that the web service is hosted.
5. Configure admin knobs.

Uninstalling an existing instance of the web service

To install an upgrade to a web service, any previous, existing instance must first be uninstalled.

All steps within this procedure are completed on the SharePoint server.

1. Within IIS, click **Sites** and find the website that was created by the previous installation. Right-click that website and remove it.
2. Within IIS, click **Application Pools**, find the web application that was created (it has the same name as the website). Right-click that web application and remove it.
3. In Windows Explorer, go to the folder where the web service was deployed and delete all content within this folder.
4. Reset IIS with the `iisreset` command.

Verifying the location of the hosted web service

This procedure details how to verify the location of the hosted web service and access it in a browser.

1. Within IIS Manager, click **Sites**, and verify that you see the new site name that is listed along with the name that is entered into the installer.
2. Expand **Sites** and verify that you can see the new site name that is listed along with the name that is entered into the installer.
3. Select the newly created site and switch to the **Content View**, which is on the right pane.
4. An SVC file corresponds to the installed web service that is installed. Right-click the SVC file and click **Browse**.
   The web service URL is started in a browser, the address bar of which contains the HTTP location of the web service such as `http://localhost:9000/UpdateItemService.svc`.

Configuration of administration knobs

IBM StoredIQ has various administration knobs that can be customized to deploy web services. Various knobs are described, and usage notes and examples are provided.

Table 51: IBM StoredIQ administration knobs

Admin knob	Notes	Example
cm8_missing _mime_type _error The default value is 1.	If IBM Content Manager does not know about a provided MIME type, an exception is generated. If the user prefers that a generic MIME type application/octet-stream is assigned to the archived content instead of an exception, set this knob to 1.	update ADMINKNOBS SET value = 1 where name = 'cm8_missing_mime_type_error'
global_copy_ ignore_target _vc The default value is zero.	When set to 1, IBM StoredIQ does not automatically harvest the destination volume when it is creating copies.	INSERT INTO adminknobs (name, value, description, valuetype, use)SELECT 'global_copy_ignore_target_vc', 0, 'Global override to turn off auto-harvests on a copy', 'int', 1 WHERE (SELECT name FROM adminknobs WHERE name = 'global_copy_ignore_target_vc') IS NULL
sharepoint_ custom_ webservice _location	Location of custom web-service used to facilitate migration of time stamps and owner information.	INSERT INTO adminknobs (name, value, description, valuetype, use) SELECT 'sharepoint_custom_webservice_location', '', 'Location for custom web-service hosted on SharePoint, eg: 9100:/UpdateItemService.svc', 'str', 2 WHERE (SELECT name FROM adminknobs WHERE name = 'sharepoint_custom_webservice_location') IS NULL
sharepoint_ harvest_docs_ only The default value is zero.	When set to 1, IBM StoredIQ will harvest only document libraries from SharePoint.	INSERT INTO adminknobs (name, value, description, valuetype, use)SELECT 'sharepoint_harvest_docs_only', 0, 'Harvest only Sharepoint document type list objects', 'int', 1 WHERE (SELECT name FROM adminknobs WHERE name = 'sharepoint_harvest_docs_only') IS NULL

Configuring administration knobs

Administration knobs must be configured to point at the installed web service.

Note: It is necessary to restart services after changes to admin knobs.

The value of the admin knob requires that you know the location of the web service. The format of the service location that is entered as the value for the admin knob is port:servicelocation.

1. Verify the location of the web service.

2. Enter the location of the web service as part of the value for the admin knob in the format `port:servicelocation`.
   For example, if the web service URL is `http://localhost:9000/UpdateItemService.svc`, the admin knob value must be set to `9000:/UpdateItemService.svc`. It is done with the following SQL command: `update adminknobs set value='9000:/UpdateItemService.svc' where name='sharepoint_custom_webservice_location';`

Supported file types

The following section provides a comprehensive list of the file types that can be harvested and processed by IBM StoredIQ, organized by name and by category. You can also view SharePoint attributes.

Supported file types by name

All file types by name that is supported by IBM StoredIQ are listed, including category, format, extension, category, and version.

Table 52: Supported file types by name.

Format	Extension	Category	Version
Adobe Acrobat	PDF	graphic	• 2.1 • 3.0-7.0 • Japanese
Adobe FrameMaker Graphics	FMV	graphic	vector/raster through 5.0
Adobe FrameMaker Interchange Format	MIF	word processing	3.0-6.0
Adobe Illustrator		graphic	• Through 7.0 • 9.0
Adobe Photoshop	PSD	graphic	4.0
Ami Draw	SDW	graphic	all
ANSI	TXT	text and markup	7- and 8-bit
ASCII	TXT	text and markup	7- and 8-bit
AutoCAD	DWG	CAD	• 2.5-2.6 • 9.0-14.0 • 2002 • 2004 • 2005
AutoShade Rendering	RND	graphic	2.0
Binary Group 3 Fax		graphic	all
Bitmap	BMP, RLE, ICO, CUR, DIB, WARP	graphic	all
CALS Raster	GP4	graphic	Type I, II
Comma-Separated Values	CSV	spreadsheet	
Computer Graphics Metafile	CGM	graphic	• ANSI • CALS • NIST 3.0

Table 52: Supported file types by name. (continued)

Format	Extension	Category	Version
Corel Clipart	CMX	graphic	5-6
Corel Draw	CDR	graphic	3.x-8.x
Corel Draw (CDR with Tiff header)		graphic	2.x-9.x
Corel Presentations	SHW	presentation	• Through 12.0 • X3
Corel WordPerfect Windows	WPD	word processing	• Through 12.0 • X3
DataEase		Database	4.X
dBase Database		Database	Through 5.0
dBXL		Database	1.3
DEC WPS PLUS	DX	word processing	Through 4.0
DEC WPS PLUS	WPL	word processing	Through 4.1
DisplayWrite (2 and 3)	IP	word processing	all
DisplayWrite (4 and 5)		word processing	Through 2.0
DOS command executable	COM	system	
Dynamic link library files	DLL	system	
EBCDIC		text and markup	all
ENABLE		word processing	• 3.0 • 4.0 • 4.5
ENABLE		Database	• 3.0 • 4.0 • 4.5
ENABLE Spreadsheet	SSF	spreadsheet	• 3.0 • 4.0 • 4.5
Encapsulated Post-Script (raster)	EPS	graphic	TIFF header
Executable files	EXE	system	
First Choice		Database	Through 3.0
First Choice		word processing	Through 3.0
First Choice		spreadsheet	Through 3.0
FoxBase		Database	2.1

Table 52: Supported file types by name. (continued)

Format	Extension	Category	Version
Framework		Database	3.0
Framework		word processing	3.0
Framework		spreadsheet	3.0
GEM Bit Image	IMG	graphic	all
Graphics Interchange Format	GIF	graphic	all
Graphics Environment Manager	GEM VDI	graphic	Bitmap and vector
Gzip	GZ	archive	all
Haansoft Hangul	HWP	word processing	• 1997 • 2002
Harvard Graphics (DOS)		graphic	• 2.x • 3.x
Harvard Graphics (Windows)		graphic	all
Hewlett-Packard Graphics Language	HPGL	graphic	2
HTML	HTM	text and markup	Through 3.0
IBM FFT		text and markup	all
IBM Graphics Data Format	GDF	graphic	1.0
IBM Picture Interchange Format	PIF	graphic	1.0
IBM Revisable Form Text		text and markup	all
IBM Writing Assistant		word processing	1.01
Initial Graphics Exchange Spec	IGES	graphic	5.1
Java class files	CLASS	system	
JPEG (not in TIFF format)	JFIF	graphic	all
JPEG	JPEG	graphic	all
JustSystems Ichitaro	JTD	word processing	• 5.0 • 6.0 • 8.0-13.0 • 2004
JustSystems Write		word processing	Through 3.0
Kodak Flash Pix	FPX	graphic	all

Table 52: Supported file types by name. (continued)

Format	Extension	Category	Version
Kodak Photo CD	PCD	graphic	1.0
Legacy		word processing	Through 1.1
Legato Email Extender	EMX	Email	
Lotus 1-2-3	WK4	spreadsheet	Through 5.0
Lotus 1-2-3 (OS/2)		spreadsheet	Through 2.0
Lotus 1-2-3 Charts	123	spreadsheet	Through 5.0
Lotus 1-2-3 for SmartSuite		spreadsheet	1997-Millennium 9.6
Lotus AMI Pro	SAM	word processing	Through 3.1
Lotus Freelance Graphics	PRZ	presentation	Through Millennium
Lotus Freelance Graphics (OS/2)	PRE	presentation	Through 2.0
Lotus Manuscript		word processing	2.0
Lotus Notes	NSF	Email	
Lotus Pic	PIC	graphic	all
Lotus Snapshot		graphic	all
Lotus Symphony		spreadsheet	• 1.0 • 1.1 • 2.0
Lotus Word Pro	LWP	word processing	1996-9.6
LZA Self Extracting Compress		archive	all
LZH Compress		archive	all
Macintosh PICT1/2	PICT1/PICT1	graphic	Bitmap only
MacPaint	PNTG	graphic	NA
MacWrite II		word processing	1.1
Macromedia Flash	SWF	presentation	text only
MASS-11		word processing	Through 8.0
Micrografx Designer	DRW	graphic	Through 3.1
Micrografx Designer	DSF	graphic	Win95, 6.0
Micrografx Draw	DRW	graphic	Through 4.0
MPEG-1 Audio layer 3	MP3	multimedia	• ID3 metadata only • These files can be harvested, but there is no data in them that can be used in tags.

Table 52: Supported file types by name. (continued)

Format	Extension	Category	Version
MS Access	MDB	Database	Through 2.0
MS Binder		archive	7.0-1997
MS Excel	XLS	spreadsheet	2.2-2007
MS Excel Charts		spreadsheet	2.x-7.0
MS Excel (Macintosh)	XLS	spreadsheet	• 3.0-4.0 • 1998 • 2001 • 2004
MS Excel XML	XLSX	spreadsheet	
MS MultiPlan		spreadsheet	4.0
MS Outlook Express	EML	Email	1997-2003
MS Outlook Form Template	OFT	Email	1997-2003
MS Outlook Message	MSG	Email	all
MS Outlook Offline Folder	OST	Email	1997-2003
MS Outlook Personal Folder	PST	Email	1997-2007
MS PowerPoint (Macintosh)	PPT	presentation	4.0-2004
MS PowerPoint (Windows)	PPT	presentation	3.0-2007
MS PowerPoint XML	PPTX	presentation	
MS Project	MPP	Database	1998-2003
MS Windows XML	DOCX	word processing	
MS Word (Macintosh)	DOC	word processing	• 3.0-4.0 • 1998 • 2001
MS Word (PC)	DOC	word processing	Through 6.0
MS Word (Windows)	DOC	word processing	Through 2007
MS WordPad		word processing	all
MS Works	S30/S40	spreadsheet	Through 2.0
MS Works	WPS	word processing	Through 4.0
MS Works (Macintosh)		word processing	Through 2.0
MS Works Database (Macintosh)		Database	Through 2.0
MS Works Database (PC)		Database	Through 2.0

Table 52: Supported file types by name. (continued)

Format	Extension	Category	Version
MS Works Database (Windows)		Database	Through 4.0
MS Write		word processing	Through 3.0
Mosaic Twin		spreadsheet	2.5
MultiMate 4.0		word processing	Through 4.0
Navy DIF		word processing	all
Nota Bene		word processing	3.0
Novell Perfect Works		word processing	2.0
Novell Perfect Works		spreadsheet	2.0
Novell Perfect Works (Draw)		graphic	2.0
Novell WordPerfect		word processing	Through 6.1
Novell WordPerfect (Macintosh)		word processing	1.02-3.0
Office Writer		word processing	4.0-6.0
OpenOffice Calc	SXC/ODS	spreadsheet	• 1.1 • 2.0
OpenOffice Draw		graphic	• 1.1 • 2.0
OpenOffice Impress	SXI/SXP/ODP	presentation	• 1.1 • 2.0
OpenOffice Writer	SXW/ODT	word processing	• 1.1 • 2.0
OS/2 PMMetafile Graphics	MET	graphic	3.0
Paint Shop Pro 6	PSP	graphic	5.0-6.0
Paradox Database (PC)		Database	Through 4.0
Paradox (Windows)		Database	Through 1.0
PC-File Letter		word processing	Through 5.0
PC-File+Letter		word processing	Through 3.0
PC PaintBrush	PCX, DCX	graphic	all
PFS: Professional Plan		spreadsheet	1.0
PFS: Write		word processing	A, B, C
Portable Bitmap Utilities	PBM	graphic	all
Portable Greymap	PGM	graphic	NA

Table 52: Supported file types by name. (continued)

Format	Extension	Category	Version
Portable Network Graphics	PNG	graphic	1.0
Portable Pixmap Utilities	PPM	graphic	NA
PostScript File	PS	graphic	level II
Professional Write		word processing	Through 2.1
Professional Write Plus		word processing	1.0
Progressive JPEG		graphic	NA
Q &A (database)		Database	Through 2.0
Q & A (DOS)		word processing	2.0
Q & A (Windows)		word processing	2.0
Q & A Write		word processing	3.0
Quattro Pro (DOS)		spreadsheet	Through 5.0
Quattro Pro (Windows)		spreadsheet	• Through 12.0 • X3
R:BASE 5000		Database	Through 3.1
R:BASE (Personal)		Database	1.0
R:BASE System V		Database	1.0
RAR	RAR	archive	
Reflex Database		Database	2.0
Rich Text Format	RTF	text and markup	all
SAMNA Word IV		word processing	
Smart Ware II		Database	1.02
Smart Ware II		word processing	1.02
Smart Ware II		spreadsheet	1.02
Sprint		word processing	1.0
StarOffice Calc	SXC/ODS	spreadsheet	• 5.2 • 6.x • 7.x • 8.0
StarOffice Draw		graphic	• 5.2 • 6.x • 7.x • 8.0

Table 52: Supported file types by name. (continued)

Format	Extension	Category	Version
StarOffice Impress	SXI/SXP/ODP	presentation	• 5.2 • 6.x • 7.x • 8.0
StarOffice Writer	SXW/ODT	word processing	• 5.2 • 6.x • 7.x • 8.0
Sun Raster Image	RS	graphic	NA
Supercalc Spreadsheet		spreadsheet	4.0
Text Mail (MIME)	various	Email	
Total Word		word processing	1.2
Truevision Image	TIFF	graphic	Through 6
Truevision Targa	TGA	graphic	2
Unicode Text	TXT	text and markup	all
UNIX TAR (tape archive)	TAR	archive	NA
UNIX Compressed	Z	archive	NA
UUEncoding	UUE	archive	NA
vCard		word processing	2.1
Visio (preview)		graphic	4
Visio 2003		graphic	• 5 • 2000 • 2002
Volkswriter		word processing	Through 1.0
VP Planner 3D		spreadsheet	1.0
WANG PC		word processing	Through 2.6
WBMP		graphic	NA
Windows Enhanced Metafile	EMF	graphic	NA
Windows Metafile	WMF	graphic	NA
Winzip	ZIP	archive	
WML		text and markup	5.2
WordMARC word processor		word processing	Through Composer
WordPerfect Graphics	WPG, WPG2	graphic	Through 2.0, 7. and 10

Table 52: Supported file types by name. (continued)			
Format	**Extension**	**Category**	**Version**
WordStar		word processing	Through 7.0
WordStar 2000		word processing	Through 3.0
X Bitmap	XBM	graphic	x10
X Dump	XWD	graphic	x10
X Pixmap	XPM	graphic	x10
XML (generic)	XML	text and markup	
XyWrite	XY4	word processing	Through III Plus
Yahoo! IM Archive		archive	
ZIP	ZIP	archive	PKWARE-2.04g

There is a potential for EBCDIC files to be typed incorrectly. It is especially true for raw text formats that do not have embedded header information. In this case, IBM StoredIQ makes a best guess attempt at typing the file.

Supported file types by category

All file types by category that is supported by IBM StoredIQ are listed, including category, format, extension, and version.

Table 53: Supported archive file types by category.		
Format	**Extension**	**Version**
• Gzip	• GZ	• all
• LZA Self-Extracting Comparess		• all
• LZH Compress		• all
• MS Binder		• 7.0-1997
• RAR	• RAR	
• Unix TAR (tape archive)	• TAR	• NA
• Unix Compressed	• Z	• NA
• UUEncoding	• UUE	• NA
• Winzip	• Zip	
• Yahoo! IM Archive		• NA
• ZIP	• ZIP	• PKWARE-2.04g

Table 54: Supported CAD file types by category.

Format	Extension	Version
• AutoCAD	• DWG	• 2.5-2.6 • 9.0-14.0 • 2002 • 2004 • 2005

Table 55: Supported database file types by category.

Format	Extension	Version
• DataEase		• 4.x
• dBase DataBase		• Through 5.0
• dBXL		• 1.3
• ENABLE		• 3.0 • 4.0 • 4.5
• First Choice		• Through 3.0
• FoxBase		• 2.1
• Framework		• 3.0
• MS Access	• MDB	• Through 2.0
• MS Project	• MPP	• Through 2.0
• MS Works Database (Macintosh)		• 2.0
• MS Works Database (PC)		• Through 2.0
• MS Works Database (Windows)		• Through 4.0
• Paradox Database (PC)		• Through 4.0
• Paradox Database (Windows)		• Through 1.0
• Q&A (database)		• Through 2.0
• R:BASE 5000		• Through 3.1
• R:BASE (personal)		• 1.0
• R:BASE System V		• 1.0
• Reflex Database		• 2.0

Table 55: Supported database file types by category. (continued)		
Format	**Extension**	**Version**
• Smart Ware II		• 1.02

Table 56: Supported email file types by category.		
Format	**Extension**	**Version**
• Legato Email Extender	• EMX	
• Lotus Notes	• NSF	
• MS Outlook Express	• EML	• 1997-2003
• MS Outlook Form Template	• OFT	• 1997-2003
• MS Outlook Message	• MSG	• all
• MS Outlook Offline Folder	• OST	• 1997-2003
• MS Outlook Personal Folder	• PST	• 1997-2007
• Text Mail (MIME)	• various	

Table 57: Supported graphic file types by category.		
Format	**Extension**	**Version**
• Adobe Acrobat	• PDF	• 2.1 • 3.0-7.0 • Japanese
• Adobe Framemaker Graphics	• FMV	• vector/raster-5.0
• Adobe Illustrator		• Through 7.0 • 9.0
• Adobe Photoshop	• PSD	• 4.0
• Ami Draw	• SDW	• all
• AutoShade Rendering	• RND	• 2.0
• Binary Group 3 Fax		• all
• Bitmap	• BMP, RLE, ICO, CUR, DIB, WARP	• all
• CALS Raster	• GP4	• Type I, II

Table 57: Supported graphic file types by category. (continued)		
Format	**Extension**	**Version**
• Computer Graphics Metafile	• CGM	• ANSI • CALS • NIST 3.0
• Corel Cliart	• CMX	• 5-6
• Corel Draw	• CDR	• 3.x-8.x
• Corel Draw (CDR with TIFF header		• 2.x-9.x
• Encapsulated Post Script (raster)	• EPS	• TIFF header
• GEM Bit Image	• IMG	• all
• Graphics Interchange Format	• GIF	• all
• Graphics Environment Manager	• GEM VDI	• Bitmap • vector
• Harvard Graphics (DOS)		• 2.x • 3.x
• Harvard Graphics (Windows)		• all
• Hewlett-Packard Graphics Language	• HPGL	• 2
• IBM Graphics Data Format	• GDF	• 1.0
• IBM Picture Interchange Format	• PIF	• 1.0
• JPEG (not in TIFF format)	• JFIF	• all
• JPEG	• JPEG	• all
• Kodak Flash PIX	• FPX	• all
• Kodak Photo CD	• PCD	• 1.0
• Lotus Pic	• PIC	• all
• Lotus Snapshot		• all
• macintosh PICT1/2	• PICT1/PICT2	• Bitmap only
• MacPaint	• PNTG	• NA

Table 57: Supported graphic file types by category. (continued)

Format	Extension	Version
• Micrografx Designer	• DRW	• Through 3.1
• Micrografx Draw	• DRW	• Through 4.0
• Novell Perfect Works (Draw)		• 2.0
• OpenOffice Draw		• 1.1 • 2.0
• OZ/2 PM Metafile Graphics	• MET	• 3.0
• Paint Shop Pro 6	• PSP	• 5.0-6.0
• PC Paintbrush	• PCX, DCX	• all
• Portable Bitmap Utilities	• PBM	• all
• Portable Network Graphics	• PNG	• 1.0
• Portable Pixmap Utilities	• PPM	• NA
• Postscript	• PS	• Level II
• Progressive JPEG		• NA
• StarOffice Draw		• 5.2 • 6.x • 7.x • 8.0
• Sun Raster Image	• RS	• NA
• Truevision Image	• TIFF	• Through 6
• Truevision Targa	• TGA	• 2
• Visio		• 4
• Visio 2003		• 5 • 2000 • 2002
• WBMP		• NA
• Windows Enhanced Metafile	• EMF	• NA
• Windows Metafile	• WMF	• NA

Table 57: Supported graphic file types by category. (continued)		
Format	**Extension**	**Version**
• WordPerfect Graphics	• WPG, WPG2	• Through 2.0 • 7 • 10
• X Bitmap	• XBM	• x10
• XDump	• XWD	• x10
• X Pixmap	• XPM	• x10

Table 58: Supported multimedia file types by category.		
Format	**Extension**	**Version**
• MPEG-1 Audio Layer 3	• MP3	• ID3 metadata only **Note:** These files can be harvested, but there is no data in them that can be used in tags.

Table 59: Supported presentation file types by category.		
Format	**Extension**	**Version**
• Corel Presentations	• SHW	• Through 12.0 • X3
• Lotus Freelance Graphics	• PRZ	• Through Millennium
• Lotus Freelance Graphics (OS/2)	• PRE	• Through 2.0
• Macromedia Flash	• SWF	• text only
• MS PowerPoint (Macintosh)	• PPT	• 4.0-2004
• MS PowerPoint (Windows)	• PPT	• 3.0-2007
• MS PowerPoint XML	• PPTX	
• OpenOffice Impress	• SXI/SXP/ODP	• 1.1 • 2.0
• StarOffice Impress	• SXI/SXP/ODP	• 5.2 • 6.x • 7.x • 8.0

Table 60: Supported spreadsheet file types by category.

Format	Extension	Version
• Comma-Separated Values	• CSV	
• ENABLE Spreadsheet	• SSF	• 3.0 • 4.0 • 4.5
• First Choice		• Through 3.0
• Framework		• 3.0
• Lotus 1-2-3	• WK4	• Through 5.0
• Lotus 1-2-3 (OS/2)	•	• Through 2.0
• Lotus 1-2-3 Charts	• 123	• Through 5.0
• Lotus 1-2-3 for SmartSuite		• 197-9.6
• Lotus Symphony		• 1.0 • 1.1 • 2.0
• MS Excel	• XLS	• 2.2-2007
• MS Excel Charts		• 2.x-7.0
• MS Excel (Macintosh)	• XLS	• 3.0-4.0 • 1998 • 2001 • 2004
• MS Excel XML	• XLSX	
• MS MultiPlan		• 4.0
• MS Works	• S30/S40	• Through 2.0
• Mosaic Twin		• 2.5
• Novell Perfect Works		• 2.0
• OpenOffice Calc	• SXC/ODS	• 1.1 • 2.0
• PFS: Professional Plan		• 1.0
• Quattro Pro (DOS)		• Through 5.0

Table 60: Supported spreadsheet file types by category. (continued)		
Format	**Extension**	**Version**
• Quattro Pro (Windows)		• Through 12.0 • X3
• Smart Ware II		• 1.02
• StarOffice Calc	• SXC/ODS	• 5.2 • 6.x • 7.x • 8.0
• Supercalc Spreadsheet		• 4.0
• VP Planner 3D		• 1.0

Table 61: Supported system file types by category.	
Format	**Extension**
• Executable files	• .EXE
• Dynamic link library files	• .DLL
• Java class files	• .class
• DOS command executables	• .COM

Table 62: Supported text and markup file types by category.		
Format	**Extension**	**Version**
• ANSI	• .TXT	• 7- and 8-bit
• ASCII	• .TXT	• 7- and 8-bit
• EBCDIC		• all
• HTML	• .HTM	• Through 3.0
• IBM FFT		• all
• IBM Revisable Form Text		• all
• Rich Text Format	• RTF	• all
• Unicode Text	• .TXT	• all
• WML		• 5.2
• XML	• .XML	

There is a potential for EBCDIC files to be typed incorrectly. It is especially true for raw text formats that do not have embedded header information. In this case, IBM StoredIQ makes a best guess attempt at typing the file.

Table 63: Supported word-processing file types by category.

Format	Extension	Version
• Adobe FrameMaker Interchange Format	• MIF	• 3.0-6.0
• Corel WordPerfect Windows	• WPD	• Through 12.0 • X3
• DEC WPS PLUS	• DX	• Through 4.0
• DEC WPS PLUS	• WPL	• Through 4.1
• Display Write (2 and 3)	• IP	• all
• Display Write (4 and 5)		• Through 2.0
• ENABLE		• 3.0 • 4.0 • 4.5
• First Choice		• Through 3.0
• Framework		• 3.0
• Haansoft Hangul	• HWP	• 1997 • 2002
• IBM Writing Assistant		• 1.01
• JustSystems Ichitaro	• JTD	• 5.0 • 6.0 • 8.0-13.0 • 2004
• JustSystems Write		• Through 3.0
• Legacy		• Through 1.1
• Lotus AMI Pro	• SAM	• Through 3.1
• Lotus Manuscript		• 2.0
• Lotus Word Pro	• LWP	• 1996-9.6
• MacWrite II		• 1.1
• MASS-11		• Through 8.0

Table 63: Supported word-processing file types by category. (continued)		
Format	**Extension**	**Version**
• MS Windows XML	• DOCX	
• MS Word (Macintosh)	• DOC	• 3.0-4.0 • 1998 • 2001
• MS Word (PC)	• DOC	• Through 6.0
• MS Word (Windows)	• DOC	• Through 2007
• MS WordPad		• all version
• MS Works	• WPS	• Through 4.0
• MS Works (Macintosh)		• Through 2.0
• MS Write		• Through 3.0
• MultiMate 4.0		• Through 4.0
• Navy DIF		• all versions
• Nota Bene		• 3.0
• Novell Perfect Works		• 2.0
• Novell WordPerfect		• Through 6.1
• Novell WordPerfect (Macintosh)		• 1.02-3.0
• Office Writer		• 4.0-6.0
• OpenOffice Writer	• SXW/ODT	• 1.1 • 2.0
• PC-File Letter		• Through 5.0
• PC-File + Letter		• Through 3.0
• PFS Write		• A • B • C
• Professional Write Plus		• Through 2.1
• Q&A (DOS)		• 2.0
• Q&A (Windows)		• 2.0
• Q&A Write		• 3.0

Table 63: Supported word-processing file types by category. (continued)		
Format	**Extension**	**Version**
• SAMNA Word IV		
• Smart Ware II		• 1.02
• Sprint		• 1.0
• StarOffice Writer	• SXW/ODT	• 5.2 • 6.x • 7.x • 8.0
• Total Word		• 1.2

SharePoint supported file types

The following section describes the various SharePoint data object types and their properties that are currently supported by IBM StoredIQ.

Supported SharePoint object types

These types of SharePoint objects are supported:

Table 64: Supported SharePoint object types		
Supported SharePoint object types	**Supported SharePoint object types**	**Supported SharePoint object types**
• Blog posts and comments	• Discussion board	• Calendar
• Tasks	• Project tasks	• Contacts
• Wiki pages	• Issue tracker	• Announcements
• Survey	• Links	• Document libraries
• Picture libraries	• Records center	

Notes regarding SharePoint object types

- **Calendar**: Recurring calendar events are indexed as a single object in IBM StoredIQ. Each recurring calendar event has multiple `Event Date` and `End Date` attribute values, one pair per recurrence. For instance, if there is an event defined for `American Independence Day` and is set to recur yearly, it is indexed with Event Dates `2010-07-04`, `2011-07-04`, `2012-07-04`, and more.
- **Survey**: Only individual responses to a survey are indexed as system-level objects. Each response is a user's feedback to all questions in the survey. Each question in the survey that was answered for a response is indexed as an attribute of the response in the IBM StoredIQ index. The name of the attribute is the string that forms the question while the value is the reply entered. Surveys have no full-text indexable body, and they are always indexed with `size=0`.

Hash computation

The hash of a full-text indexed object is computed with the full-text indexable body of the object. However, in the case of SharePoint list item objects (excluding documents and pictures), the full-text indexable body might be empty or too simplistic. It means that you can easily obtain duplicate items across otherwise two different objects. For this reason, other attributes are included in the hash computation algorithm.

These attributes are included while the hash is computed for the SharePoint data objects, excluding documents and pictures.

Table 65: Attribute summary

Attribute	**Types**
Generic attributes	• Title (SharePoint) • Content type (SharePoint) • Description (SharePoint)
Blog post attributes	• Post category (SharePoint)
Wiki page attributes	• Wiki page comment
Calendar event attributes	• Event category (SharePoint) • Event date (SharePoint) • Event end date (SharePoint) • Event location (SharePoint)
Task or project task attributes	• Task start date (SharePoint) • Task due date (SharePoint) • Task that is assigned to (SharePoint)
Contact attributes	• Contact full name(SharePoint) • Contact email (SharePoint) • Contact job title (SharePoint) • Contact work address (SharePoint) • Contact work phone (SharePoint) • Contact home phone (SharePoint) • Contact mobile phone (SharePoint)
Link attributes	• Link URL (SharePoint)
Survey attributes	• All survey questions and answers in the response are included in the hash.

Supported server platforms and protocols

The following section lists the supported server platforms by volume type and the protocols for supported systems.

Primary volume

A primary volume is storage knowledge workers access to create, read, update, and delete unstructured content. Unstructured content is stored in standard formats such as office documents, text files, system logs, application logs, email, and compressed archives that contain email, documents, or enterprise social media content.

Retention volume

Generally, a retention volume is immutable storage that enforces retention and hold policies. While data is under management, it cannot be modified or deleted, and knowledge workers typically do not access retention storage directly. Specific applications move data to this storage to manage it. This storage typically has its own special API/protocol, although NAS vendors implemented retention/hold features that use standard CIFS/SMB or SMB2 and NFS protocols. 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 are free to 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 that 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 more) to replicate the content and metadata from the retention volume when needed.

Export volume

Export volumes are 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 local documents to be imported into a legal review tool in a format such as standard EDRM or a Concordance-compatible format.

System volume

System volumes are 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.

Supported platforms and protocols by IBM StoredIQ volume type

Platform/ protocol	IBM StoredIQ primary volume	IBM StoredIQ retention volume	IBM StoredIQ export volume	IBM StoredIQ system volume	Notes
Box	x				Box volumes are added through IBM StoredIQ Administrator.
CIFS/SMB or SMB2	x	x	x	x	
CMIS 1.0	x				
Connections	x				
EMC Centera		x			

Platform/ protocol	IBM StoredIQ primary volume	IBM StoredIQ retention volume	IBM StoredIQ export volume	IBM StoredIQ system volume	Notes
EMC Documentum	x				Customer must supply DFC files to enable the connector.
HDFS (HADOOP)	x				
Hitachi HCAP		x			
IBM Content Manager	x				
IBM Domino	x				Email only. This includes IBM Verse.
IBM FileNet	x				
Jive	x				
Microsoft Exchange	x				
Microsoft SharePoint	x				
NewsGator	x				
NFS	x	x	x	x	
OneDrive for Business	x				
OpenText Livelink/ Content Server	x				
Salesforce Chatter	x				
Symantec Discovery Accelerator	x				
Symantec Enterprise Vault		x			

If IBM StoredIQ Desktop Data Collector is installed, Windows desktops can serve as primary volumes. Supported actions for desktops are copy from, move from, export from, and delete. For more information, see “Desktop collection” on page 101.

Supported operations and limitations by platform and protocol

Platform/protocol	Operation type: read	Operation type: write	Operation type: delete	Notes
Box	R	W	D	Box volumes are added through IBM StoredIQ Administrator.
CIFS/SMB or SMB2	R	W	D	
Connections	R			
CMIS 1.0	R	W		The CMIS option references a solution that involves different products. As such, no specific supported version is to cite, and it is not called out in the following tables.
EMC Centera	R	W	D	Metadata and content that are written by IBM StoredIQ or EMC EmailXtender is supported. EmailXtender support is limited to email archived from Exchange.
EMC Documentum	R	W	D	Only **Current** documents in standard cabinets are indexed.
HDFS (HADOOP)	R	W	D	
Hitachi HCAP	R	W	D	
IBM Content Manager	R	W		
IBM Domino	R			Email only. Email is converted to the .MSG format for processing. This includes IBM Verse.
IBM FileNet	R	W		
Jive	R			

Platform/protocol	Operation type: read	Operation type: write	Operation type: delete	Notes
Microsoft Exchange	R			Messages (email), Contacts, Calendar items, Notes, Tasks, and Documents
Microsoft SharePoint	R	W	D	All document versions are optional. SharePoint 2010/2013/2016 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.
NewsGator	R			No API support for Poll responses.
NFS	R	W	D	
OneDrive for Business	R	W		
OpenText Livelink/ Content Server	R			
Salesforce Chatter	R			
Symantec Discovery Accelerator	R			
Symantec Enterprise Vault	R	W	D	The delete action is only supported for retention volumes and only metadata and content that are written by IBM StoredIQ.

Supported CIFS/SMB or SMB2 server platform and protocol versions (read/write/delete)

CIFS/SMB or SMB2 server	Notes
Windows XP, Vista, 7, 8, 10	
Windows Server 2003, 2008, 2008 R2, 2012, 2016	
Samba	
Mac OS X, 10.7, 10.8, 10.9	

Supported Connections server platform and protocol versions (read/write/delete)

Connections server	Notes
Connections 5.0	
Connections 6.0	

Supported EMC Centera server platform and protocol versions (read/write)

EMC Centera server	Notes
Centera API v3.1	

Supported EMC Documentum server platform and protocol versions (read/write)

EMC Documentum server	Notes
Documentum 6.0, 6.5, 6.7	Customer must supply DFC files to enable the connector.
Documentum with Retention Policy Services (RPS) 6.0, 6.5	Customer must supply DFC files to enable the connector.

Supported IBM FileNet server platform and protocol versions (read/write)

FileNet content services	Notes
FileNet 5.2.0.3	

Supported HDFS (HADOOP) server platform and protocol versions (read/write/delete)

HDFS service	Notes
HADOOP v 2.7	

Supported Hitachi HCAP server platform and protocol versions (read/write/delete)

Hitachi HCAP service	Notes
Hitachi Content Platform v1.8	

Supported IBM Content Manager server platform and protocol versions (read/write)

IBM Content Manager 8.4.3 and later	Notes
AIX 5L 5.3, 6.1, 7.1	
DB2 (back-end database)	

IBM Content Manager 8.4.3 and later	Notes
Red Hat Enterprise Linux 4.0, 5.0 SUSE Linux Enterprise Server 9, 10, 11	
Oracle (back-end database)	
Solaris 9, 10	
Windows Server 2003, 2008, 2008 R2	

Supported IBM Domino/Notes server platform and protocol versions (read)

IBM Domino/Notes server	Notes
Domino/Notes 6.x, 7.x, 8.x, 9.x	

Supported Jive server platform and protocol versions (read)

Jive service	Notes
Jive 5.0.2	

Supported Microsoft Exchange server platform and protocol versions (read)

Microsoft Exchange server	Notes
Exchange 2003	WebDAV protocol
Exchange 2007, 2010, 2013, 2016, Online	Exchange web service interface

Supported Microsoft SharePoint server platform and protocol versions (read/write/delete)

Microsoft SharePoint server	Notes
SharePoint 2003	WebDAV protocol
SharePoint 2007, 2010, 2013, 2016, Online	SharePoint web service interface

Supported NewsGator server platform and protocol versions (read)

NewsGator API	Notes
NewsGator Social 2.1.1229	Installed on SharePoint 2010 or later.

Supported NFS server platform and protocol versions (read/write/delete)

NFSv3	Notes
Red Hat Enterprise Linux 5.x, 6.x	
CentOS 5.x, 6.x	
Mac OS X, 10.7, 10.8, 10.9	

Supported OneDrive server platform and protocol versions (read/write)

OneDrive	Notes
graph.microsoft.com	HTTP REST API

Supported OpenText Livelink Enterprise Server platform and protocol versions (read)

Livelink Enterprise Server	Notes
OpenText Livelink Enterprise Server 9.7, 9.7.1	Connector is based on IBM Content Integrator (ICI) version 8.6.

Supported OpenText Content Server platform and protocol versions (read)

Content Server	Notes
Content Server 10.0.0	Connector is based on IBM Content Integrator (ICI) version 8.6.

Supported Salesforce Chatter server platform and protocol versions (read)

Salesforce Chatter service	Notes
Salesforce Partner API v26.0	

Supported Symantec Discovery Accelerator server platform and protocol versions (read)

Symantec Discovery Accelerator server	Notes
Enterprise Vault DCOM API 8.0.3, 9.0, 10.0	Query content and metadata from Discovery Accelerator cases.

Supported Symantec Enterprise Vault server platform and protocol versions (read/write)

Symantec Enterprise Vault server	Notes
Enterprise Vault DCOM API 8.0.3, 9.0, 10.0	

Event log messages

The following section contains a complete listing of all ERROR, INFO, and WARN event-log messages that appear in the **Event Log** of the IBM StoredIQ console.

ERROR event log messages

The following table contains a complete listing of all ERROR event-log messages, reasons for occurrence, sample messages, and any required customer action.

Table 66: ERROR event log messages. This table lists all ERROR event-log messages.

Event number	Reason	Sample message	Required customer action
1001	Harvest was unable to open a socket for listening to child process.	Harvest could not allocate listen port after <number> attempts. Cannot kickstart interrogators. (1001)	Log in to UTIL and restart the application server. Restart the data server. Contact customer support.
9083	Unexpected error while it is exporting a volume.	Exporting volume 'dataserver:/mnt/demo-A' (1357) has failed (9083)	Contact Customer Support.
9086	Unexpected error while it is importing a volume	Importing volume 'dataserver:/mnt/demo-A' (1357) failed (9086)	Contact Customer Support.
15001	No volumes are able to be harvested in a job. For instance, all of the mounts fail due to a network issue.	No volumes harvested. (15001)	Make sure IBM StoredIQ still has appropriate permissions to a volume. Verify that there is network connectivity between the data server and your volume. Contact Customer Support.
15002	Could not mount the volume. Check permissions and network settings.	Error mounting volume <share><start-dir> on server <server-name>. Reported <reason>. (15002)	Make sure that the data server still has appropriate permissions to a volume. Verify that there is network connectivity between the data server and your volume. Contact Customer Support.
15021	Error saving harvest record.	Failed to save HarvestRecord for qa1:auto-A (15021)	Contact Customer Support. This message occurs due to a database error.
17001	Unhandled unrecoverable exception in Centera Discovery.	Centera Harvester fatal failure: <exception description> (17001)	Contact Customer Support.

Table 66: ERROR event log messages. This table lists all ERROR event-log messages. *(continued)*

Event number	Reason	Sample message	Required customer action
17012	An error occurred when a volume was created during Centera Discovery.	Unable to create Centera Volume Company_jpool_ 2009_ FEB_1 in pool jpool. Error:<database error description> (17012)	Contact Customer Support. This message occurs due to a database error.
17501	Generic retention discovery failed in a catastrophic manner.	Generic retention discovery fatal failure: <17501>	Contact Customer Support.
17503	Generic retention discovery creates volume sets associated with primary volumes. When that fails, IBM StoredIQ sends this message. This failure likely occurred due to database errors.	Error creating/loading volumeset for <server>:<share>	Contact Customer Support.
17505	Unable to query object count for a discovered volume due to a database error.	Unable to determine object count for <server>:<share>	Contact Customer Support.
17506	Generic retention discovery could not create discovered volume.	Error creating volume <server>:<share:>	Contact Customer Support.
18001	SMB connection fails.	Windows Share Protocol Exception when connecting to the server <server-name> : <reason>. (18001)	Make sure IBM StoredIQ still has appropriate permissions to a volume. Verify that there is network connectivity between the data server and your volume. Contact Customer Support.
18002	The SMB volume mount failed. Check the share name.	Windows Share Protocol Exception when connecting to the share <share-name> on <server-name> : <reason>. (18002)	Verify the name of the server and volume to make sure that they are correct. If this message persists, then contact Customer Support.
18003	There is no volume manager.	Windows Share Protocol Exception while initializing the data object manager: <reason>. (18003)	Contact Customer Support.
18006	Grazer volume crawl threw an exception.	Grazer._run : Unknown error during walk. (18006)	Verify the user that mounted the specified volume has permissions equivalent to your current backup solution. If this message continues, contact Customer Support.

Table 66: ERROR event log messages. This table lists all ERROR event-log messages. *(continued)*			
Event number	**Reason**	**Sample message**	**Required customer action**
18021	An unexpected error from the server prevented the harvest to reach the end of the activity stream on the NewsGator data source that is harvested. The next incremental harvest attempts to pick up from where the current harvest was interrupted.	Unable to fetch trailing activity stream from NewsGator volume. Will retry in next harvest. (18021)	Check to ensure the NewsGator server has sufficient resources (disk space, memory). It is likely that this error is transient. If the error persists across multiple harvests, contact Customer Support.
18018	Start directory has escape characters, and the data server is configured to skip them.	Cannot graze the volume, root directory Nunez has escape characters (18018)	Consider turning off escape character checking.
19001	An exception occurred during interrogator initialization.	Interrogator._init __exception: <reason>. (19001)	Contact Customer Support.
19002	An unknown exception occurred during interrogator initialization.	Interrogator.__ init __exception: unknown. (19002)	Contact Customer Support.
19003	An exception occurred during interrogator processing.	Interrogator.process exception (<volumeid>, <epoch>): <reason>. (19003)	Contact Customer Support.
19004	An unknown exception occurred during interrogator processing.	Interrogator.process exception (<volumeid>, <epoch>). (19004)	Contact Customer Support.
19005	An exception occurred during viewer initialization.	Viewer.__init__: Exception - <reason>. (19005)	Contact Customer Support.
19006	An unknown exception occurred during viewer initialization.	Viewer.__init__: Unknown exception. (19006)	Contact Customer Support.

Table 66: ERROR event log messages. This table lists all ERROR event-log messages. *(continued)*			
Event number	**Reason**	**Sample message**	**Required customer action**
33003	Could not mount the volume. Check permissions and network settings.	Unable to mount the volume: <error reason> (33003)	Verify whether user name and password that are used for mounting the volume are accurate. Check the user data object for appropriate permissions to the volume. Make sure that the volume is accessible from one of the built-in protocols (NFS, Windows Share, or Exchange). Verify that the network is properly configured for the appliance to reach the volume. Verify that the appliance has appropriate DNS settings to resolve the server name.
33004	Volume could not be unmounted.	Unmounting volume failed from mount point : <mount point>. (33004)	Restart the data server. If the problem persists, then contact Customer Support.
33005	Data server was unable to create a local mounting point for the volume.	Unable to create mount_point using primitive.thread-Safe-Makedirs(). (33005)	Restart the data server. If the problem persists, then contact Customer Support.
33010	Failed to make SMB connection to Windows Share server.	Mounting Windows Share volume failed with the error : <system error message>. (33010)	Verify user name and password that is used for mounting the volume are accurate. Check the user data object for appropriate permissions to the volume. Make sure that the volume is accessible from one of the built-in protocols (Windows Share). Verify that the network is properly configured for the data server to reach the volume. Verify that the data server has appropriate DNS settings to resolve the server name.
33011	Internal error. Problem accessing local /proc/mounts	Unable to open /proc/mounts. Cannot test if volume was already mounted. (33011)	Restart the data server. If the problem persists, then contact Customer Support.
33012	Database problems when a volume was deleted.	An exception occurred while working with HARVESTS_TABLE in Volume._delete(). (33012)	Contact Customer Support.

Table 66: ERROR event log messages. This table lists all ERROR event-log messages. *(continued)*			
Event number	**Reason**	**Sample message**	**Required customer action**
33013	No volume set was found for the volumes set name.	Unable to load volume set by its name. (33013)	Contact Customer Support.
33014	System could not determine when this volume was last harvested.	An error occurred while performing the last_harvest operation. (33014)	Contact Customer Support.
33018	An error occurred mounting the Exchange share.	Mounting Exchange Server failed : <reason>. (33018)	Verify user name and password that is used for mounting the share are accurate. Check for appropriate permissions to the share. Make sure that the share is accessible. Verify that the network is properly configured for the data server to reach the share. Verify that the data server has appropriate DNS settings to resolve the server name.
33019	Failed to connect and authenticate to the Hitachi Archivas Content Archive server.	Mounting HCAP volume failed : Cannot connect to HCAP share. (33019)	Ensure the connectivity, credentials, and permissions to the Hitachi volume and try again.
33022	Failed to connect to the Discovery Accelerator with the information for the Volume.	Mounting Discovery Accelerator volume failed with the error: insufficient permissions to review CaseOne (33022)	Verify the information that is used to add the volume and ensure that all details are entered correctly before you try it again. If the error points to network issues with connectivity, address them and try again.
33027	The attempt to connect and authenticate to the IBM FileNet server failed.	Mounting IBM FileNet volume failed : <reason>. (33027)	Ensure the connectivity, credentials, and permissions to the FileNet volume and try again.
33029	Failed to create the volume as the number of active volume partitions exceeds the limit of 500.	Exceeded maximum number of volume partitions (33029).	Contact Customer Support.
34002	Could not complete the copy action because the target disk was full.	Copy Action aborted as the target disk has run out of space (34002)	Verify that there is space available on your policy destination and try again.

Table 66: ERROR event log messages. This table lists all ERROR event-log messages. *(continued)*

Event number	Reason	Sample message	Required customer action
34009	Could not complete the move action due to full target disk.	Move Action aborted as the target disk has run out of space.(34009)	Verify that there is space available on your policy destination, then run another harvest before you run the policy. When the harvest completes, try running the policy again.
34015	The policy audit could not be deleted for some reason.	Error Deleting Policy Audit: <error message> (34016)	Contact Customer Support.
34020	The copy to Centera action could not be run because of insufficient permissions.	Copy to Centera failed as we do not have read/write permissions on the access profile used. (34020)	Check permissions on the access profile that is provided for the Centera pool on which the volume is defined, and check whether the appropriate permissions are provided.
34021	The move to Centera action could not be run because of insufficient permissions.	Move to Centera failed as we do not have read/write permissions on the access profile used. (34020)	Check permissions on the access profile that is provided for the Centera pool on which the volume is defined, and check whether the appropriate permissions are provided.
34030	Discovery export policy is started since it detected the target disk is full.	Production Run action aborted because the target disk has run out of space. (34030)	Create sufficient space on target disk and run discovery export policy again.
34034	The target volume for the policy could not be mounted. The policy is started.	Copy objects failed, unable to mount volume: QA1.COMPANY. COM:SHARE. (34034)	Ensure the connectivity, login credentials, and permissions to the target volume for the policy and try again.
41004	The job is ended abnormally.	<job-name> ended unexpectedly. (41004)	Try to run the job again. If it fails again, contact Customer Support.
41007	Job failed.	[Job name] has failed (41007).	Look at previous messages to see why it failed and refer to that message ID to pinpoint the error. Contact Customer Support.
42001	The copy action could not run because of parameter errors.	Copy data objects did not run. Errors occurred:<error-description>. (42001)	Contact Customer Support.

Table 66: ERROR event log messages. This table lists all ERROR event-log messages. *(continued)*			
Event number	**Reason**	**Sample message**	**Required customer action**
42002	The copy action was unable to create a target directory.	Copy data objects failed, unable to create target dir:<target-directory-name>. (42002)	Check permissions on the target. Make sure the permissions that are configured to mount the target volume have write access to the volume.
42004	An unexpected error occurred.	Copy data objects terminated abnormally. (42004)	Contact Customer Support.
42006	The move action could not run because of parameter errors.	Move data objects did not run. Errors occurred:<error-description>. (42006)	Contact Customer Support.
42007	The move action was unable to create a target directory.	Move data objects failed, unable to create target dir:<target-directory-name>. (42007)	Check permissions on the target. Make sure the permissions that are configured to mount the target volume have write access to the volume.
42009	An unexpected error occurred.	Move data objects terminated abnormally. (42009)	Contact Customer Support.
42017	An unexpected error occurred.	Delete data objects terminated abnormally. (42017)	Contact Customer Support.
42025	The policy action could not run because of parameter errors.	Policy cannot execute. Attribute verification failed. (42025)	Contact Customer Support.
42027	An unexpected error occurred.	Policy terminated abnormally. (42027)	Contact Customer Support.
42050	The data synchronizer could not run because of an unexpected error.	Content Data Synchronizer synchronization of <server-name>: <volume-name> failed fatally.	Contact Customer Support.
42059	Invalid set of parameters that are passed to discovery export policy.	Production Run on objects did not run. Errors occurred: The following parameters are missing: action_limit. (42059)	Contact Customer Support.
42060	Discovery export policy failed to create target directory for the export.	Production Run on objects (Copying native objects) failed, unable to create target dir: production/10. (42060)	Verify that the discovery export volume has write permission and re-execute policy.

Table 66: ERROR event log messages. This table lists all ERROR event-log messages. *(continued)*

Event number	Reason	Sample message	Required customer action
42062	Discovery export policy was ended abnormally.	Production Run on objects (Copying native objects) terminated abnormally. (42062)	Contact Customer Support.
42088	The full-text optimization process failed; however, the index is most likely still usable for queries.	Full-text optimization failed on volume <volume-name> (42088)	Contact Customer Support.
45802	A full-text index is already being modified.	Time allocated to gain exclusive access to in-memory index for volume= 1357 has expired (45802)	Contact Customer Support.
45803	The index for the specified volume does not exist. This message can occur under normal conditions.	Index '/deepfs/full-text/volume_index/volume_1357' not found. (45803)	No user intervention is required.
45804	Programming error. A transaction was never initiated or was closed early.	Transaction of client: node.client.com_ FINDEX_ QUEUE_ 1357_117251522 _3_ 2 is not the writer (45804)	Contact Customer Support.
45805	The query is not started or expired. The former is a programming error. The latter is normal.	Query ID: 123 does not exist (45805)	No user intervention is required.
45806	The query expression is invalid or not supported.	Failed to parse 'dog pre\3 bar' (45806)	Revise your full-text query.
45807	Programming error. A transaction was already started for the client.	Client: node.client.com_ FINDEX_QUEUE _1357_1172515 222_3_2 is already active (45807)	Contact Customer Support.
45808	A transaction was never started or expired.	No transaction for client: node.client.com_ FINDEX_ QUEUE_1357_ 1172515222_3_2 (45808)	No user intervention is required. The system handles this condition internally.
45810	Programming error.	Invalid volumeId. Expected: 1357 Received:2468 (45810)	Contact Customer Support.
45812	A File I/O error occurred while the system was accessing index data.	Failed to write disk (45812).	Try your query again. Contact Customer Support for more assistance if necessary.
45814	The query expression is too long.	Query: 'a* b* c* d* e*' is too complex (45814)	Refine your full-text query.

Table 66: ERROR event log messages. This table lists all ERROR event-log messages. *(continued)*			
Event number	**Reason**	**Sample message**	**Required customer action**
45815	The file that is being indexed is too large or the query expression is too complex. The engine temporarily ran out of memory.	Java heap exhausted while indexing node with ID: '10f4179cd5ff22f 2a6b 79a1bc3aef247 fd94ccff' (45815)	Check the skipped file list in the audit log for files that failed to load due to their sizes. Revise your query expression and try again.
46023	Tar command failed while it persists full-text data to Windows Share or NFS share.	Failed to back up full-text data for server:share. Reason: <reason>. (46023)	Check disk space and permissions.
46024	Unhandled unrecoverable exception while persisting full-text data into a .tgz file.	Exception <exception> while backing up fulltext data for server:share (46024)	Contact Customer Support.
46025	Was not able to delete partial .tgz file after a failed full-text backup.	Failed to unlink incomplete backup image. Reason: <reason>. (46025)	Check permissions.
47002	Synchronization failed on a query.	Synchronization failed for query '<query-name>' on volume '<server-and-volume> (47002)	Contact Customer Support.
47101	An error occurred during the query of a full-text expression.	Cannot process full-text expression (Failed to read from disk (45812) (47101)	Restart services and contact Customer Support.
47203	No more database connections are available.	Database connections exhausted (512/511) (47203)	Contact Customer Support.
47207	User is running out of disk space.	Disk usage exceeds threshold. (%d)	Contact Customer Support. In rare cases, this message can indicate a program error leaking disk space. In most cases, however, disk space is almost full, and more storage is required.
47212	Interrogator failed while the system processed a file. The current file is missing from the volume cluster.	Harvester 1 Does not exist. Action taken : restart. (47212)	If the problem persists (that is, the system fails on the same file or type of files), contact Customer Support.
47214	SNMP notification sender is unable to resolve the trap host name.	Unable to resolve host name nomachine.no where.com (47214)	Check spelling and DNS setup.
50011	The DDL/DML files that are required for the database versioning were not found in the expected location on the data server.	Database version control SQL file not found. (50011)	Contact Customer Support.

Table 66: ERROR event log messages. This table lists all ERROR event-log messages. *(continued)*

Event number	Reason	Sample message	Required customer action
50018	Indicates that the pre-upgrade database restoration failed, which was attempted as a result of a database upgrade failure.	Database restore is unsuccessful. Contact Customer Support. (50018)	Contact Customer Support.
50020	Indicates that the current database requirements do not meet those requirements that are specified for the upgrade and cannot proceed with the upgrade.	Versions do not match! Expected current database version: <dbversion>. (50020)	Contact Customer Support.
50021	Indicates that the full database backup failed when the system attempts a data-object level database backup.	Database backup failed. (50021)	Contact Customer Support.
61003	Discovery export policy failed to mount volume.	Production policy failed to mount volume. Aborting. (61003)	
61005	The discovery export load file generation fails unexpectedly. The load files can be produced correctly, but post-processing actions like updating audit trails and generating report files might not complete.	Production load file generation failed. Load files may be produced, but post-processing may be incomplete. (61005)	Contact Customer Support.
61006	The discovery export load file generation was interrupted because the target disk is full.	Production load file generation interrupted. Target disk full. (61006)	Free up space on the target disk, void the discovery export run and run the policy again.
68001	The gateway and data server must be on the same version to connect.	Gateway connection failed due to unsupported data server version.	Update your data server to the same build number as the gateway and restart services. If your encounter issues, contact Customer Support.
68003	The data server failed to connect to the gateway over an extended period.	The data-server connection to the gateway cannot be established.	Contact Customer Support.
8002	The system failed to open a connection to the database.	Failed to connect to the database (80002)	The "maximum database connections" configuration parameter of the database engine might need to be increased. Contact Customer Support.

INFO event log messages

The following table contains a complete listing of all INFO event-log messages.

Table 67: INFO event log messages

Event number	Reason	Sample message	Required customer action
9001	No conditions were added to a query.	Harvester: Query <query name> cannot be inferred because no condition for it has been defined (9001).	Add conditions to the specified query.
9002	One or more conditions in a query were incorrect.	Harvester: Query < query name> cannot be inferred because of regular expression or other condition error (9002).	Verify that regular expressions are properly formed.
9003	Volume Harvest is complete and explorers are being calculated.	Volume statistics computation started (9003).	No user intervention is required.
9004	Explorer calculations are complete.	Volume statistics computation completed (9004).	No user intervention is required.
9005	Query membership calculations started.	Query inference will be done in <number> steps (9005).	No user intervention is required.
9006	Query membership calculations progress information.	Query inference step <number> done (9006).	No user intervention is required.
9007	Query membership calculations completed.	Query inference completed (9007).	No user intervention is required.
9012	Indicates the end of dumping the content of the volume cache.	Dump of Volume cache(s) completed (9012).	No user intervention is required.
9013	Indicates the beginning of the load process.	Postprocessing for volume 'Company Data Server:/mnt/demo-A' started (9013).	No user intervention is required.
9067	Indicates load progress.	System metadata and tagged values were successfully loaded for volume 'server:volume' (9067).	No user intervention is required.
9069	Indicates load progress.	Volume 'data server: /mnt/demo-A': System metadata, tagged values and full-text index were successfully loaded (9069).	No user intervention is required.
9084	The volume export finished.	Exporting volume 'data server:/mnt/demo-A' (1357) completed (9084)	No user intervention is required.

Table 67: INFO event log messages (continued)

Event number	Reason	Sample message	Required customer action
9087	The volume import finished.	Importing volume 'dataserver:/mnt/demo-A' (1357) completed (9087)	No user intervention is required.
9091	The load process was ended by the user.	Load aborted due to user request (9091).	No user intervention is required.
15008	The volume load step was skipped, per user request.	Post processing skipped for volume <server>:<vo-lume>. (15008)	No user intervention is required.
15009	The volume load step was run but the harvest step was skipped, per user request.	Harvest skipped for volume <server>:<vol-ume>. (15009)	No user intervention is required.
15012	The policy that ran on the volume is complete and the volume load can now proceed.	Volume <volume> on server <server> is free now. Proceeding with load. (15012)	No user intervention is required.
15013	The configured time limit on a harvest was reached.	Harvest time limit reached for server:share. Ending harvest now. (15013)	No user intervention is required.
15014	The configured object count limit on a harvest was reached.	Object count limit reached for server:share. Ending harvest now. (15014)	No user intervention is required.
15017	Check box is selected for nightly load job.	Deferring post processing for volume server:vol (15017)	No user intervention is required.
15018	Harvest size or time limit is reached.	Harvest limit reached on server:volume. Synthetic deletes will not be computed. (15018)	No user intervention is required.
15019	User stops harvest process.	Harvest stopped by user while processing volume dpfsvr:vol1. Rest of volumes will be skipped. (15019)	No user intervention is required.
15020	The harvest vocabulary changed. Full harvest must run instead of incremental.	Vocabulary for dpfsvr:jhaide-A has changed. A full harvest is recommended (15020).	Full harvest must be run instead of an incremental harvest.
15022	The user is trying to run an ACL-only harvest on a volume that is not a Windows Share or SharePoint volume.	Permission-only harvest: permission checks not supported for <server>:<share>	No action is needed as the volume is skipped.
15023	The user is trying to run an ACL-only harvest on a volume that was assigned a user list.	Permission-only harvest: volume <server>:<share> has no associated user list.	No action is needed as the volume is skipped.

Table 67: INFO event log messages (continued)

Event number	Reason	Sample message	Required customer action
17002	Sent when Centera Discovery sends the query to the Centera server	Centera External Iterator : Starting to populate using pool QAPOOL. (17002)	No user intervention is required.
17003	Sent when Centera Discovery autocreates a new volume	<servername>: Created new volume QAPOOL:_QA-POOL_2009_JAN _1 (17003)	No user intervention is required.
17004	A Centera Discovery auto-created volume reached the preconfigured limit, starting a new one.	Object limit reached for _QAPOOL_2009 _JAN_1, starting new volume (17004)	No user intervention is required.
17007	Pending data return from Centera.	Centera Harvester: No items returned from Centera for over 5 minutes. Still waiting. (17007)	Check whether a Centera node is down. A Centera cluster might be overloaded.
17009	Configured time limit is reached for Centera discovery step.	Centera Discovery : time limit for discovery reached. Ending this run. (17009)	No user intervention is required.
17010	Configured object count limit that is reached for Centera discovery step.	Centera Discovery: configured item count limit for discovery reached. Ending this run. (17010)	No user intervention is required.
17507	Limit (time or object count) reached for generic retention discovery.	Retention discovery limit reached for <server>:<share>	Contact Customer Support.
17508	Generic retention discovery found no new items for this master volume.	No new items discovered. Post-processing skipped for volume <server>:<share>	No user intervention is required unless the user is certain that new items must be discovered.
17509	Generic retention discovery created a new volume.	Created new discovered volume <server>:<share> in volume set <autodis-covered volume set name>.	No user intervention is required.
18004	Job was stopped.	Walker._process File: Grazer Stopped. (18004)	No user intervention is required.
18005	Grazer queue was closed.	Walker._processFile: Grazer Closed. (18005)	No user intervention is required.

Table 67: INFO event log messages (continued)

Event number	Reason	Sample message	Required customer action
18016	Displays the list of top-level directories that are selected by matching the start directory regular expression. Displays at the beginning of a harvest.	Choosing top-level directories: <directories> (18016)	No user intervention is required.
34001	Marks current progress of a copy action.	<volume>: <count> data objects processed by copy action. (34001)	No user intervention is required.
34004	Marks current progress of a delete action.	<volume>: <count> data objects processed by delete action. (34004)	No user intervention is required.
34008	Marks current progress of a move action.	<volume>: <count> data objects processed by move action. (34008)	No user intervention is required.
34014	A policy audit was deleted.	Deleting Policy Audit # <audit id> <policy name> <start time> (34014)	No user intervention is required.
34015	A policy audit was deleted.	Deleted Policy Audit # <audit id> <policy name> <start time> (34015)	No user intervention is required.
34031	Progress update of the discovery export policy, every 10000 objects processed.	Winserver:top share : 30000 data objects processed by production action. (34031)	No user intervention is required.
41001	A job was started either manually or was scheduled.	<jobname> started. (41001)	No user intervention is required.
41002	The user stopped a job that was running.	<jobname> stopped at user request (41002)	No user intervention is required.
41003	A job is completed normally with or without success.	<jobname> completed. (41003)	No user intervention is required.
41006	Rebooting or restarting services on the controller or compute node causes all jobs to stop.	Service shutdown. Stopping outstanding jobs. (41006)	Rerun jobs after restart if you want the jobs to complete.
41008	Database compactor (vacuum) job cannot run while there is database activity.	Database compactor was not run because other jobs are active (41008).	Set the database compactor's job schedule so that it does not conflict with long-running jobs.
42005	The action completed or was ended. Shows results of copy action.	Copy complete: <number> data objects copied,<number> collisions found. (42005)	No user intervention is required.

Table 67: INFO event log messages (continued)

Event number	Reason	Sample message	Required customer action
42010	The action completed or was ended. Shows results of move action.	Move complete: <number> data objects moved,<number > collisions found. (42010)	No user intervention is required.
42018	The action completed or was ended. Shows results of deleted action.	Copy data objects complete: <number> data objects copied,<number> collisions found. (42018)	No user intervention is required.
42020	Third-party application deleted clips on discovered Centera volumes.	Centera Deleted Files Synchronizer complete: 10000 objects passed. 2 objects were missing from the Centera cluster. (42020)	No user intervention is required.
42024	The synchronizer was completed normally.	Content Data Synchronizer complete. (42024)	No user intervention is required.
42028	The action completed or was ended. Shows results of policy action.	Policy completed (42028).	No user intervention is required.
42032	The action completed or was ended. Shows results of report action.	<report name> completed (42032).	No user intervention is required.
42033	The synchronizer started automatically or manually with the GUI button.	Content Data Synchronizer started. (42033)	No user intervention is required.
42048	Reports that the synchronizer is skipping a volume if synchronization is determined not to be required.	Content Data Synchronizer skipping <server-name>:<volume-name> as it does not need synchronization. 42048)	No user intervention is required.
42049	Reports that the synchronizer started synchronization of a volume.	Content Data Synchronizer starting synchronization for volume <server-name>:<volume-name>	No user intervention is required.
42053	The policy that was waiting for participant volumes to be loaded before it continues, is now starting.	Proceeding with execution of <policy-name>.	No user intervention is required.
42063	Report on completion of discovery export policy execution phase.	Production Run on objects (Copying native objects) completed: 2003 data objects copied, 25 duplicates found. (42063)	No user intervention is required.

Table 67: INFO event log messages (continued)

Event number	Reason	Sample message	Required customer action
42065	A discovery export policy that was held up for want of resources, is now done waiting, and begins execution.	Proceeding with execution of 'Production case One'. (42065)	No user intervention is required.
42066	A new discovery export run started.	New run number 10 started for production Production Case 23221. (42066)	Note the new run number to tie the current run with the corresponding audit trail.
42067	Discovery export policy is preparing the audit trail in XML format. It might take a few minutes.	Production Run producing Audit Trail XML. (42067)	No user intervention is required.
42074	A query or tag was replicated to a member data server successfully.	Successfully sent query 'Custodian: Joe' to member data server San Jose Office (42074)	No user intervention is required.
46001	The backup process began. Any selected backups in the system configuration screen are run if necessary.	Backup Process Started. (46001)	No user intervention is required.
46002	The backup process did not complete all its tasks successfully. One or more backup types did not occur.	Backup Process Failed: <error-description>. (46002)	Check your backup volume.
46003	The backup process completed attempting all the necessary tasks successfully. Any parts of the overall process add their own log entries.	Backup Process Finished. (46003)	No user intervention is required.
46004	The Application Data backup, as part of the overall backup process, needed to run but did not succeed.	Application Data backup failed. (46004)	Check your backup volume. Look at the setup for the Application Data backup. If backups continue to fail, contact Customer Support.
46005	The Application Data backup, as part of the overall backup process, needed to run and succeeded.	Application Data backup finished. (46005)	No user intervention is required.

Table 67: INFO event log messages (continued)

Event number	Reason	Sample message	Required customer action
46006	The Application Data backup, as part of the overall backup process, was not configured.	Application Data backup not configured, skipped. (46006)	No user intervention is required.
46007	The Harvested Volume Data backup, as part of the overall backup process, needed to run but did not succeed.	Harvested Volume Data backup failed. (46007)	Check your backup volume. Look at the setup for the Harvested Volume Data backup. If backups continue to fail, contact Customer Support.
46008	The Harvested Volume Data backup, as part of the overall backup process, needed to run and succeeded.	Harvested Volume Data backup finished. (46008)	No user intervention is required.
46009	The Harvested Volume Data backup, as part of the overall backup process, was not configured.	Harvested Volume Data backup not configured, skipped. (46009)	No user intervention is required.
46010	The System Configuration backup, as part of the overall backup process, needed to run but did not succeed.	System Configuration backup failed. (46010)	Check your backup volume. Look at the setup for the System Configuration backup. If backups continue to fail, contact Customer Support.
46011	The System Configuration backup, as part of the overall backup process, needed to run and succeeded.	System Configuration backup finished. (46011)	No user intervention is required.
46012	The System Configuration backup, as part of the overall backup process, was not configured.	System Configuration backup not configured, skipped. (46012)	No user intervention is required.
46013	The Audit Trail backup, as part of the overall backup process, needed to run but did not succeed.	Policy Audit Trail backup failed. (46013)	Check your backup volume. Look at the setup for the Audit Trail backup. If back-ups continue to fail, contact IBM support.
46014	The Audit Trail backup, as part of the overall backup process, needed to run and succeeded.	Policy Audit Trail backup finished. (46014)	No user intervention is required.

Table 67: INFO event log messages (continued)

Event number	Reason	Sample message	Required customer action
46015	The Audit Trail backup, as part of the overall backup process was not configured.	Policy Audit Trail backup not configured, skipped. (46015)	No user intervention is required.
46019	Volume cluster backup failed.	Indexed Data backup failed: <specific error> (46019)	Contact Customer Support.
46020	Volume cluster backup finished.	Indexed Data backup finished. (46020)	No user intervention is required.
46021	Volume is not configured for indexed data backups.	Indexed Data backup not configured, skipped. (46021)	No user intervention is required.
46022	Full-text data was successfully backed up.	Successfully backed up full-text data for server:share (46022)	No user intervention is required.
47213	Interrogator was successfully restarted.	Harvester 1 is now running. (47213)	No user intervention is required.
60001	The user updates an object on the system. It includes any object type on the data server, including the updating of volumes.	Query cities was updated by the administrator account (60001).	No user intervention is required.
60002	The user creates an object. It includes any object type on the data server, including the creation of volumes.	Query cities was created by the administrator account (60002).	No user intervention is required.
60003	The user deletes an object. It includes any object type on the data server, including the deletion of volumes.	Query cities was deleted by the administrator account (60003).	No user intervention is required.
60004	The user publishes a full-text query set or a query.	Query cities draft was published by the administrator account (60004).	No user intervention is required.
60005	The user tags an object. It includes a published query, a draft query, or tag.	Query tagging for cities class was started by the administrator account (60005).	No user intervention is required.
60006	A user restarted services on the data server.	Application services restart for all data servers was requested by the administrator account (60006).	No user intervention is required.
61001	Concordance discovery export is now preparing the load files.	Preparing for upload of load file(s). (61001)	No user intervention is required.

Table 67: INFO event log messages (continued)			
Event number	**Reason**	**Sample message**	**Required customer action**
61002	Concordance discovery export is ready to upload the load files.	Load file(s) ready for upload. (61002)	No user intervention is required.
65000	The log file finished downloading.	Log file download complete (65000)	No user intervention is required.

WARN event log messages

The following table contains a complete listing of all WARN event-log messages, reasons for occurrence, sample messages, and any required customer action.

Table 68: WARN event log messages			
Event number	**Reason**	**Sample message**	**Customer action**
1002	An Interrogator process failed because of an unknown error. The data object that was processing is skipped. A new process is created to replace it.	Processing could not be completed on object, interrogator died : <data object name>. (1002)	Classify the document manually and Contact Customer Support.
1003	Interrogator child process did not properly get started. There might be problems to access the volume to be harvested.	Interrogator terminated before accessing data objects. (1003)	Try to readd the volume that is harvested. If that fails, contact Customer Support.
1004	Interrogator child process was ended because it was no longer responding. The data object that was processing is skipped. A new process is created to replace it.	Processing was not completed on object, interrogator killed : <data object name>. (1004)	Contact Customer Support.
6001	A user email might not be sent. The mail server settings are incorrect.	Failed to send an email to user <email address>; check mail server configuration settings (6001).	Verify that your SMTP server is configured correctly. Make sure that the IP address that is configured for the data server can relay on the configured SMTP server.
8001	The database needs to be vacuumed.	The Database is approaching an operational limit. Please run the Database maintenance task using the Console interface (8001)	Run the database maintenance task to vacuum the database.

Table 68: WARN event log messages (continued)

Event number	Reason	Sample message	Customer action
9068	Tagged values were loaded, but full-text index loading failed.	System metadata and tagged values were loaded successfully for volume 'server:volume', but loading the full-text index failed (9068)	Contact Customer Support.
9070	Tagged values and full-text index loading failed.	Loading system metadata, tagged values and the full-text index failed for volume 'server:volume' (9070)	Contact Customer Support.
15003	The volume mount appeared to succeed, but the test for mount failed.	Volume <volume name> on server <server name> is not mounted. Skipping. (15003)	Contact Customer Support.
15004	A component cleanup failure on stop or completion.	[<component>] Cleanup failure on stop. (15004)	Contact Customer Support.
15005	There was a component run failure.	[<component>] Run failure. (15005)	Contact Customer Support.
15006	Cleanup failed for component after a run failure.	[<component>] Cleanup failure on abort. (15006)	Contact Customer Support.
15007	A component that is timed out needs to be stopped.	Component [<component>] unresponsive; autostopping triggered. (15007)	Try your action again. If this error continues, contact Customer Support.
15010	The same volume cannot be harvested in parallel. The harvest is skipped and the next one, if any are in queue, started.	Volume <volume-name> on server <server-name> is already being harvested. Skipping. (15010)	No user intervention is required. You might want to verify that the volume harvest is complete.
15011	A volume cannot be harvested if it is being used by another job. The harvest continues when the job is complete.	Volume <volume-name> on server <server-name> is being used by another job. Waiting before proceeding with load. (15011)	No user intervention is required.
15015	Configured harvest time limit is reached.	Time limit for harvest reached. Skipping Volume v1 on server s1. 1 (15015)	Reconfigure harvest time limit.
15016	Configured harvest object count limit is reached.	Object count limit for harvest reached. Skipping Volume v1 on server s1 (15016)	Reconfigure harvest data object limit.

Table 68: WARN event log messages (continued)

Event number	Reason	Sample message	Customer action
17008	Query that ran to discover Centera items ended unexpectedly.	Centera External Iterator : Centera Query terminated unexpectedly (<error description>). (17008)	Contact Customer Support.
17011	Running discovery on the same pool in parallel is not allowed.	Pool Jpool appears to have another discovery running. Skipping. (17011).	Make sure that two jobs are not running at the same time that discovers the same pool.
17502	Generic retention discovery is already running for this master volume.	Volume <server>:<share> appears to have another discovery running. Skipping.	No user intervention is required as the next step, if any, within the job is run.
17504	Sent when a retention discovery is run on any volume other than a Windows Share retention volume.	Volume <server>:<share> is not supported for discovery. Skipping.	Contact Customer Support.
18007	Directory listing or processing of data object failed in Grazer.	Walker._walktree: OSError - <path><rea-son> (18007)	Make sure that the appliance still has appropriate permissions to a volume. Verify that there is network connectivity between the appliance and your volume. Contact Customer Support.
18008	Unknown error occurred while processing data object or listing directory.	Walker._walktree: Unknown exception - <path>. (18008)	Contact Customer Support.
18009	Grazer timed out processing an object.	Walker._process File: Grazer Timed Out. (18009)	Contact Customer Support.
18010	The skipdirs file is either not present or not readable by root.	Unable to open skipdirs file: <filename>. Cannot skip directories as configured. (18010)	Contact Customer Support.
18011	An error occurred reading the known extensions list from the database.	Grazer._run: couldn't read extensions - <reason>. (18011)	Contact Customer Support.
18012	An unknown error occurred reading the known extensions list from the database.	Grazer._run: couldn't read extensions. (18012)	Contact Customer Support.
18015	NFS initialization warning that NIS is not available.	NIS Mapping not available. (18015)	User name and group names might be inaccurate. Check that your NIS server is available and properly configured in the data server.

Table 68: WARN event log messages (continued)

Event number	Reason	Sample message	Customer action
18017	A folder in an Enterprise Vault case that is harvested was skipped because of insufficient permissions.	Skipping Folder (ID=3) in volume evdiscaccel. company.com: CaseOne due to insufficient permissions. Check Review permissions on the folder. (18017)	If all folders were expected to be harvested in the volume, verify that the user name used to add the Volume has Folder Review permission on all folders.
18019	The checkpoint that is saved from the last harvest of the NewsGator data source failed to load. Instead of conducting an incremental harvest, a full harvest is run.	Unable to load checkpoint for NewsGator volume. A full harvest will be performed instead. (18019)	If the message repeats in subsequent harvests, contact Customer Support.
18020	The checkpoint noted for the current harvest of the NewsGator data source might not be saved. The next incremental harvest of the data source is not able to pick up from this checkpoint.	Unable to save checkpoint for NewsGator harvest of volume. (18020)	If the message repeats in subsequent harvests, contact Customer Support.
33016	System might not unmount this volume.	Windows Share Protocol Session teardown failed. (33016)	Server administrators can see that connections are left hanging for a predefined time. These connections will drop off after they time out. No user intervention required.
33017	System encountered an error while it tries to figure out what query uses this volume.	An error occurred while retrieving the query instances pointing to a volume. (33017)	Contact Customer Support.
33023	A connection to the Discovery Accelerator might not be torn down completely, and a few resources on the server might be tied up for a short while.	HTTP Connection tear-down to Discovery Accelerator failed. (33023)	None
33028	The tear-down operation of the connection to a FileNet volume failed. Some connections can be left open on the FileNet server until they time out.	IBM FileNet tear-down operation failed. (33028)	None
34003	Skipped a copy data object because disk full error.	Copy action error :- Target disk full, skipping copy : <source volume> to <target volume>. (34003)	Verify that there is space available on your policy destination and try again.

Table 68: WARN event log messages (continued)

Event number	Reason	Sample message	Customer action
34010	Skipped a move data object because disk full error.	Move action error :- Target disk full, skipping copy : <source volume> to <target volume>. (34010)	Verify that there is space available on your policy destination. After verifying that space is available, run another harvest before you run your policy. Upon harvest completion, try running the policy again.
34029	Discovery export policy detects the target disk is full and skips production of an object.	Discovery export Run action error: Target disk full, skipping discovery export: `share-1/saved/years.pdf to production/10/documents/1/0x0866e5d6c898d9ffdbea 720b090a6f46d3058605 .txt`. (34029)	Create sufficient space on target disk and run discovery export policy again.
34032	The policy that was run has no volumes in scope that is based on the configured query and scoping. The policy cannot be run.	No volumes in scope for policy. Skipping policy execution. (34032)	Check policy query and scoping configuration, and re-execute policy.
34035	If the global hash setting for the system is set to not compute data object hash, no hash can be computed for the target objects during a policy action.	Copy objects : Target hash will not be computed because Hashing is disabled for system. (34035)	If target hashes need to be computed for the policy audit trail, turn on the global hash setting before you run the policy.
34036	The policy has no source volumes in scope, which means that the policy cannot be run.	The policy has no source volume(s) in scope. Wait for the query to update before executing the policy. (34036)	Confirm that the query used by the policy has one or more volumes in scope.
42003	The job in this action is stopped by the user.	Copy data objects stopped at user request. (42003)	No user intervention is required.
42008	The job in this action is stopped by the user.	Move data objects stopped at user request. (42008)	No user intervention is required.
42016	The job in this action is stopped by the user.	Delete data objects stopped at user request. (42016)	No user intervention is required.
42026	The job in this action is stopped by the user.	Policy stopped at user request. (42026)	No user intervention is required.

Table 68: WARN event log messages (continued)

Event number	Reason	Sample message	Customer action
42035	When the job in this action is stopped by the user.	Set security for data objects stopped at user request. (42035)	No user intervention is required.
42051	Two instances of the same policy cannot run at the same time.	Policy <policy-name> is already running. Skipping. (42051)	No user intervention is required.
42052	Policies cannot be run until after volumes are loaded, if those volumes are participants to the policy by virtue of being in the query.	One or more volume(s) needed by policy <policy-name> are being loaded. Waiting for the bulk load(s) to finish. (42052)	No user intervention is required.
42061	Discovery export policy was stopped by user.	Discovery export run on objects (Copying native objects) stopped at user request. (42061)	No user intervention is required.
42064	Discovery export policy execution is delayed because a conflicting discovery export run is in progress.	A Discovery export run related to policy 'Discovery export case One' is in progress. Waiting for it to finish. (42064)	The discovery export policy execution is held up for required resources. Execution must begin as soon as resource becomes available.
42068	Policy failed to set appropriate permissions on the target directory. Objects that are created from the policy might not have appropriate permissions set.	Copy objects warning, unable to set permissions on target directory: share-1/saved. (42068)	Policy might not be able to set appropriate permissions on the objects it creates. If it is not acceptable, verify that target volume has appropriate write permissions and re-execute.
42069	If the “Copy data objects modified since last harvest” option is selected for a discovery export policy, it is valid only if the discovery export itself is defined to act on the original file/email archive, as opposed to their members. If it is not true, the warning tells the user that modified objects are skipped.	Discovery export DAT_Export is configured to act on members of containers, and cannot act on objects modified after the last harvest. Discovery export run X will skip modified objects. (42069)	If the modified objects need to be acted upon, either use a discovery export action only on the original file/email archive, or conduct an incremental harvest on the source volumes.
46026	Volume is being harvested or policies are running against it. If there are other full-text indexes to be backed up, the system works on those actions. Try this volume again.	Volume volume:share is in use. Unable to back up full-text index. Will retry later. (46026)	Rerun backup when volume is not in use.

Table 68: WARN event log messages (continued)

Event number	Reason	Sample message	Customer action
47201	Database connections are down to a normal level.	Database connections at normal level again (512/100) (47201)	No user intervention is required.
47202	The system is starting to run low on database connections. This situation is abnormal. An indication of process restarts and connections are not being cleared.	Database connections usage seems excessive (512/415) (47202)	Contact Customer Support.
47215	Someone internally or externally is trying (and failing) to SSH into the data server.	SSHD: Failed password for root from 172.17.18.185 port 57982. (47125)	Contact your local IT manager. It might be either a mistyped password by a legitimate user or in the worst case scenario, a genuine break-in attempt.
61003	One of the load files cannot be uploaded because the compute node might not be accessed to obtain.	Failed to mount transaction cache dump '/deepfs/ postgres/pro-duction_cache'. (61003)	Some of the load files will be missing after the discovery export completes. These load files are reproduced on a new run. If problem persists across runs, Contact Customer Support.
61004	Warns the user that one of the transaction cache memory dump processes encountered an error. If a discovery export runs, it means that the discovery export fails to produce one of the load files. Note: If multiple memory dumps fail, there is one warning per failed memory dump.	Transaction Cache Dump failed with error - Validation failed during creation of load file. (61004)	Run the discovery export policy that saw the error again. If the error persists, and you cannot find any cluster/data server configuration issues, contact Customer Support.

Notices

This information was developed for products and services offered in the U.S.A. This material may 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
Armonk, NY 10504-1785
U.S.A.

For license inquiries regarding double-byte (DBCS) information, contact the IBM Intellectual Property Department in your country or send inquiries, in writing, to:

Intellectual Property Licensing
Legal and Intellectual Property Law
IBM Japan Ltd.
19-21, Nihonbashi-Hakozakicho, Chuo-ku
Tokyo 103-8510, Japan

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 Web sites are provided for convenience only and do not in any manner serve as an endorsement of those Web sites. The materials at those Web sites are not part of the materials for this IBM product and use of those Web sites 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.

Licensees of this program who wish to have information about it for the purpose of enabling: (i) the exchange of information between independently created programs and other programs (including this one) and (ii) the mutual use of the information which has been exchanged, should contact:

IBM Corporation
J46A/G4
555 Bailey Avenue

San Jose, CA 95141-1003
U.S.A.

Such information may be available, subject to appropriate terms and conditions, including in some cases, payment of a fee.

The licensed program described in this document and all licensed material available for it are provided by IBM under terms of the IBM Customer Agreement, IBM International Program License Agreement or any equivalent agreement between us.

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.

All 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 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. 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.

Each copy or any portion of these sample programs or any derivative work, must include a copyright notice as follows: © (your company name) (year). Portions of this code are derived from IBM Corp. Sample Programs. © Copyright IBM Corp. 2004, 2010. All rights reserved.

If you are viewing this information softcopy, the photographs and color illustrations may not appear.

Privacy policy considerations

IBM Software products, including software as a service solutions, (“Software Offerings”) may use cookies or other technologies to collect product usage information, to help improve the end user experience, to tailor interactions with the end user or for other purposes. In many cases no personally identifiable information is collected by the Software Offerings. Some of our Software Offerings can help enable you to collect personally identifiable information. If this Software Offering uses cookies to collect personally identifiable information, specific information about this offering’s use of cookies is set forth below.

[Depending upon the configurations deployed, this Software Offering may use [session] [and] [persistent] cookies that collect each user’s [name,] [user name,] [password,] [profile name,] [position,] or [other personally identifiable information] for purposes of [session management,] [authentication,] [enhanced user usability,] [single sign-on configuration] [or other usage tracking or functional purposes.] These

cookies [cannot be disabled] [can be disabled, but disabling them will also [likely] eliminate the functionality they enable].

If the configurations deployed for this Software Offering provide you as customer the ability to collect personally identifiable information from end users via cookies and other technologies, you should seek your own legal advice about any laws applicable to such data collection, including any requirements for notice and consent.

For more information about the use of various technologies, including cookies, for these purposes, See IBM's Privacy Policy at http://www.ibm.com/privacy and IBM's Online Privacy Statement at http://www.ibm.com/privacy/details the section entitled "Cookies, Web Beacons and Other Technologies" and the "IBM Software Products and Software-as-a-Service Privacy Statement" at http://www.ibm.com/software/info/product-privacy.

Terms and conditions for product documentation

Permissions for the use of these publications are granted subject to the following terms and conditions.

Applicability

These terms and conditions are in addition to any terms of use for the IBM website.

Personal use

You may reproduce these publications for your personal, noncommercial use provided that all proprietary notices are preserved. You may not distribute, display or make derivative work of these publications, or any portion thereof, without the express consent of IBM.

Commercial use

You may reproduce, distribute and display these publications solely within your enterprise provided that all proprietary notices are preserved. You may not make derivative works of these publications, or reproduce, distribute or display these publications or any portion thereof outside your enterprise, without the express consent of IBM.

Rights

Except as expressly granted in this permission, no other permissions, licenses or rights are granted, either express or implied, to the publications or any information, data, software or other intellectual property contained therein.

IBM reserves the right to withdraw the permissions granted herein whenever, in its discretion, the use of the publications is detrimental to its interest or, as determined by IBM, the above instructions are not being properly followed.

You may not download, export or re-export this information except in full compliance with all applicable laws and regulations, including all United States export laws and regulations.

IBM MAKES NO GUARANTEE ABOUT THE CONTENT OF THESE PUBLICATIONS. THE PUBLICATIONS ARE PROVIDED "AS-IS" AND WITHOUT WARRANTY OF ANY KIND, EITHER EXPRESSED OR IMPLIED, INCLUDING BUT NOT LIMITED TO IMPLIED WARRANTIES OF MERCHANTABILITY, NON-INFRINGEMENT, AND FITNESS FOR A PARTICULAR PURPOSE.

Trademarks

IBM, the IBM logo, and ibm.com are trademarks or registered trademarks of International Business Machines Corp., 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" http://www.ibm.com/legal/copytrade.shtml.

Adobe and PostScript are either registered trademarks or trademarks of Adobe Systems Incorporated in the United States, and/or other countries.

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

Microsoft and Windows 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.

VMware, VMware vCenter Server, and VMware vSphere are registered trademarks or trademarks of VMware, Inc. or its subsidiaries in the United States and/or other jurisdictions.

Index

A

administration knobs
  configuring 118
alias 26
attributes
  SharePoint 138
audit
  saving results from 115
audit messages
  policy 115
audit results
  saving 115
audit settings
  configuring 22
Audit tab
  Configuration subtab 3
  Dashboard subtab 3
  Data Sources subtab 3
audit trails
  searching 114
audits
  harvest 106
  policy 111

B

back up 15
backup 14

C

Centera
  advanced retention 34
Centera pool
  creating 34
Chatter
  configuring messages 74
CIFS
  configuring server platform 27
CIFS (Windows platforms) 76
CIFS/SMB or SMB2 36, 140
clients
  upgrading for registered workstations 24
CMIS 36, 140
configuring SMB properties 70
Connections
  setting up the administrator access on Connections 33
content based hash 22
content typing 22
customized web services 117

D

DA Gateway settings
  configuring 9
data object types 22
data object typing 22
data server
  LDAP connections 15
data sources
  adding 26
DCOM traffic
  allowing through Windows firewall 36
delete policy 103
Desktop collection 101
desktop service
  disabling 24
desktop services
  enable 24
desktop settings
  configuring 24
desktop volume, deleting 87
Discovery Accelerator
  configuring 32
  configuring customer settings 32
  required permissions 32
discovery export volume
  creating 86
Documentum
  as a data source 74
  configuring as server platform 30
  installing 30
Domino
  adding as primary volume 72

E

Encrypted File System
  configuring 102
encrypted file-system user
  adding 25
Enterprise Vault
  configuring site 30
  retention volumes 35
ERROR event log messages 147
event
  subscribing to 110
event log
  clearing current 110
  downloading 110
  viewing 110
event log messages
  ERROR 147
  INFO 147, 157
  WARN 147, 165
event logs 110
Exchange
  servers 27
Exchange 2003
  improving performance 28
Exchange 2007 Client Access Server support 72
Exchange Admin Role 27

Exchange servers
configuring 27

F

file types
supported 120, 128
FileNet
configuring 74
folder
copying items to different 105
creating 104
deleting 104
filtering items within view 105
Library 104
moving 104
renaming 104
Workspace 104
full hash 22
full-text index settings
configure 20

G

global copy ignore target vc 118

H

harvest
full 92
incremental 92
lightweight 93
post-processing 92
troubleshooting 99
harvest audits
viewing 108
harvest list
downloading 109
harvester
lightweight harvest
harvester settings 93
harvester settings 18
harvesting
about 92
harvests
incremental 92
hash
content based 22
full 22
metadata based 22
partial 22
settings 22
hash computation
SharePoint 138
HDFS 36
Hitachi 76
Hitachi HCAP
configuration requirements 36

I

IBM Connections 33
IBM Connections data source 33
IBM Content Manager 36, 75, 140
IBM StoredIQ Desktop Data Collector
installation 101
installer 24
IBM StoredIQ Desktop Data Collector files 103
IBM StoredIQ image 15
IIS 6.0
improving performance 28
import audits 109
incremental harvests 92
INFO event log messages 157
installation methods 101
installation of IBM StoredIQ Desktop Data Collector 101
installer
downloading from the application 24
interface
buttons 5
icons 5
items
saving into different folders 105

J

Jive 36, 140
job
creating 95, 96
deleting 98
discover retention volumes 96
editing 96
predefined 97
saving as 97
starting 97
jobs
Centera deleted files synchronizer 95
CIFS/NFS retention volume deleted files synchronizer 95
Database Compactor 95
Harvest every volume 95
Hitachi deleted files synchronizer 95
System maintenance and cleanup 95
Update Age Explorers 95

L

LDAP connections
data server 15
legal
notices 172
trademarks 174
lightweight harvest
full-text settings 94
hash settings 94
performing 93
volume configuration 93
Livelink 36
logs 106

M

mail settings
configuring 11
messages
event log 147

metadata based hash 22
Microsoft Exchange 140
migrations 91

N

network settings 10
NewsGator
    required privileges 33
NFS 27, 140
NFS v2, v3 36
NFS v3 76
notices
    legal 172
NSF files
    importing from Lotus Notes 17

O

OCR processing 19
Optical Character Recognition 19

P

partial hash 22
policy audit
    by name 111
    by volume 111
    viewing by discovery export 114
    viewing by name 113
    viewing by time 113
    viewing by volume 113
    viewing details 112
policy audit failure messages 115
policy audit messages 115
policy audit success messages 115
policy audit warning messages 115
policy audits 111
predefined job
    running 97
primary volume
    Domino 72
primary volumes
    creating 36
    Exchange 2007 Client Access Server support 72
processing
    monitoring 98
protocols
    supported 140

R

recovery agent 25
remote DCOM
    enabling 35
restore 14
retention servers 34
retention volume
    adding 76
    configuring 81
    Enterprise Vault 81
retention volumes
    Enterprise Vault 35

S

search depth
    volume indexing 26
server alias 26
server platforms
    configuring 26
    supported 140
services
    restarting 7
SharePoint
    alternate-access mappings 29
    privileges for social data 29
    servers 29
SharePoint 2013 web services 117
sharepoint column mapping file 118
sharepoint custom webservice location 118
sharepoint harvest docs only 118
SharePoint objects
    supported types 138
SharePoint targets 91
SharePoint volume
    performance considerations with versioning 73
SharePoint volumes
    performance considerations 73
    permissions for mounting 92
    versioning 73
SiqDocument 75
Skipped directory 36
SMB 70
SMB properties 70
SMB1 70
SMB2 70
SMBv2 70
SNMP settings
    configuring 11
supported file types
    archive 120, 128
    CAD 120, 128
    database 120, 128
    email 120, 128
    graphic 120, 128
    multimedia 120, 128
    presentation 120, 128
    spreadsheet 120, 128
    system 120, 128
    text and markup 120, 128
    word processing 120, 128
system
    restart 7
system backup configurations 14
system restore configurations 14
system status
    checking 6
system volume
    adding 82

T

text extraction 22
trademarks 174

U

user account
- administering 16
- creating 16
- deleting 16
- locking 16
- resetting passwords 16
- unlocking 16

user interface
- Administration tab 3
- Audit tab 3
- Folders tab 3

V

verifying the location
- hosted web service 117

volume
- adding primary 36
- discovery export 86
- system 82

volume cache
- deleting 99

volume data
- exporting 83
- exporting to a system volume 83
- importing 83
- importing to a system volume 84

volume definitions
- editing 7

volume import audit
- viewing details 110

volume indexing 26
volume-import audits 109
volumes
- action limitations by type 87
- creating 26
- deleting 87
- export 140
- primary 140
- retention 76, 140
- system 140

W

WARN event log messages 165
web service
- uninstalling existing instance 117

web services
- deploy 117

Windows authentication
- enabling integration on Exchange servers 28

Windows Share
- server 27