IBM StoredIQ

Deployment and Configuration Guide

Note

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

This edition applies to Version 7.6.0.19 of product number 5724M86 and to all subsequent releases and modifications until otherwise indicated in new editions.

Contents

About this publication

IBM StoredIQ Deployment and Configuration Guide provides information about how to plan, deploy, and configure the IBM StoredIQ product.

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/.

IBM StoredIQ components

The IBM StoredIQ solution consists of these components: the application stack, the gateway, the data server, and optionally the Elasticsearch cluster.

Solution components

IBM StoredIQ provides three solution components: the gateway, data servers, and application stack (AppStack).

Gateway
The gateway communicates between the data servers and the application stack. The application stack polls the gateway for information about the data on the data servers. The data servers push the information to the gateway.

Data servers
A data server obtains the data from supported data sources and indexes it. By indexing this data, you gain information about unstructured data such as file size, file data types, file owners.

The data server pushes the information about volumes and indexes to the gateway so it can be communicated to the application stack. Multiple data servers feed into a single gateway.

Data servers can be categorized in two types: DataServer - Classic and DataServer - Distributed. A data server of the type DataServer - Classic uses the embedded PostgreSQL database for storing the index. With a data server of the type DataServer - Distributed, the index is stored in an Elasticsearch cluster. Data servers of this type also provide better performance in search queries. They can manage much larger amounts of data than data servers of the type DataServer - Classic, thus making the IBM StoredIQ deployments more scalable.

You can have both types of data servers in your IBM StoredIQ deployment.

In addition to completing standard administrative tasks, administrators can deploy the IBM StoredIQ Desktop Data Collector and index desktops from the data server.

Application stack
The application stack provides the user interface for the IBM StoredIQ Administrator, IBM StoredIQ Data Workbench, IBM StoredIQ Insights, and the IBM StoredIQ Policy Manager products.

The synchronization feature for integration with a governance catalog is also part of the application stack.

Elasticsearch cluster
The Elasticsearch cluster attached to a data server of the type DataServer - Distributed provides a single data store for all metadata and content of harvested objects. Indexed data is distributed automatically across the nodes in the cluster. Indexing and queries are load-balanced across all nodes. Nodes can be added dynamically without downtime and the indexing process can use these newly added nodes without further setup.

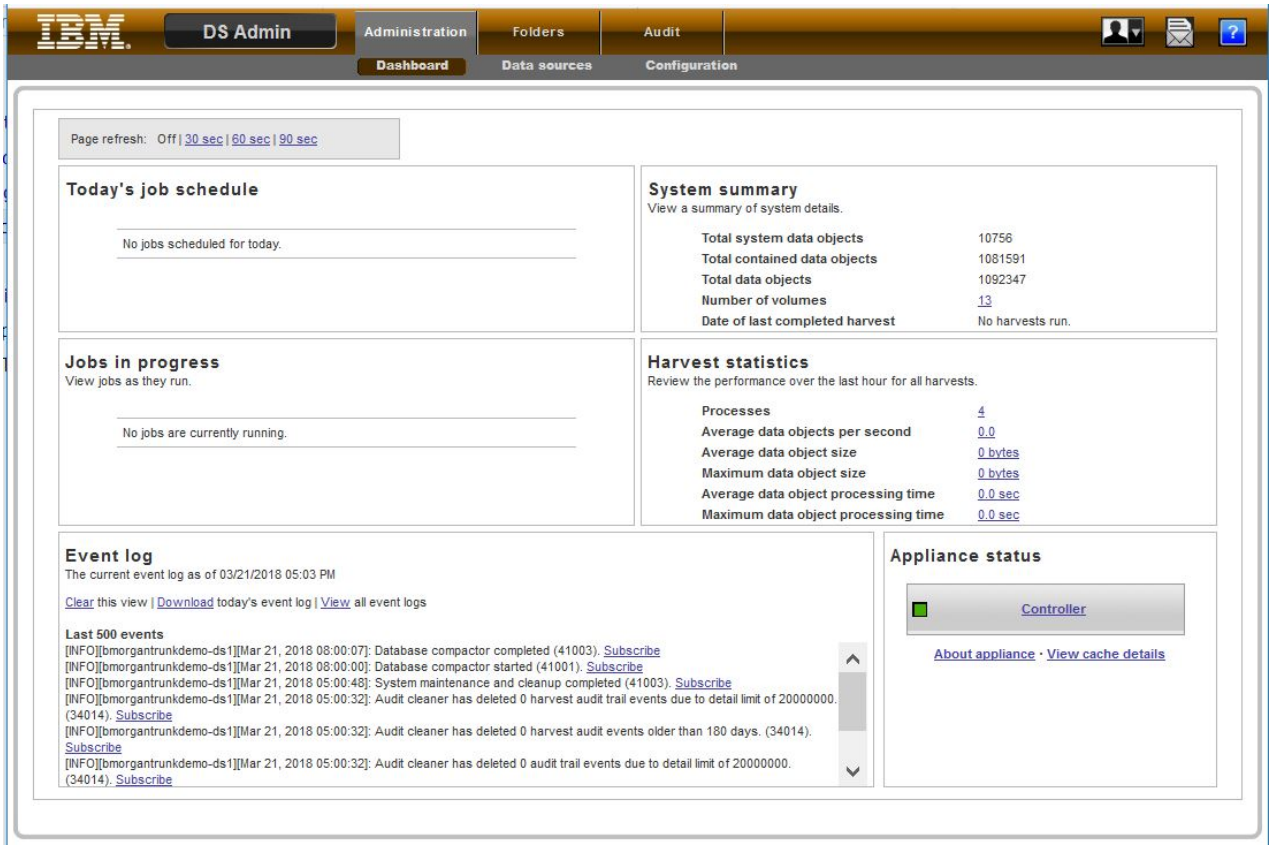
Applications of IBM StoredIQ

IBM StoredIQ provides interface applications that help fulfill its solution goals.

IBM StoredIQ Data Server

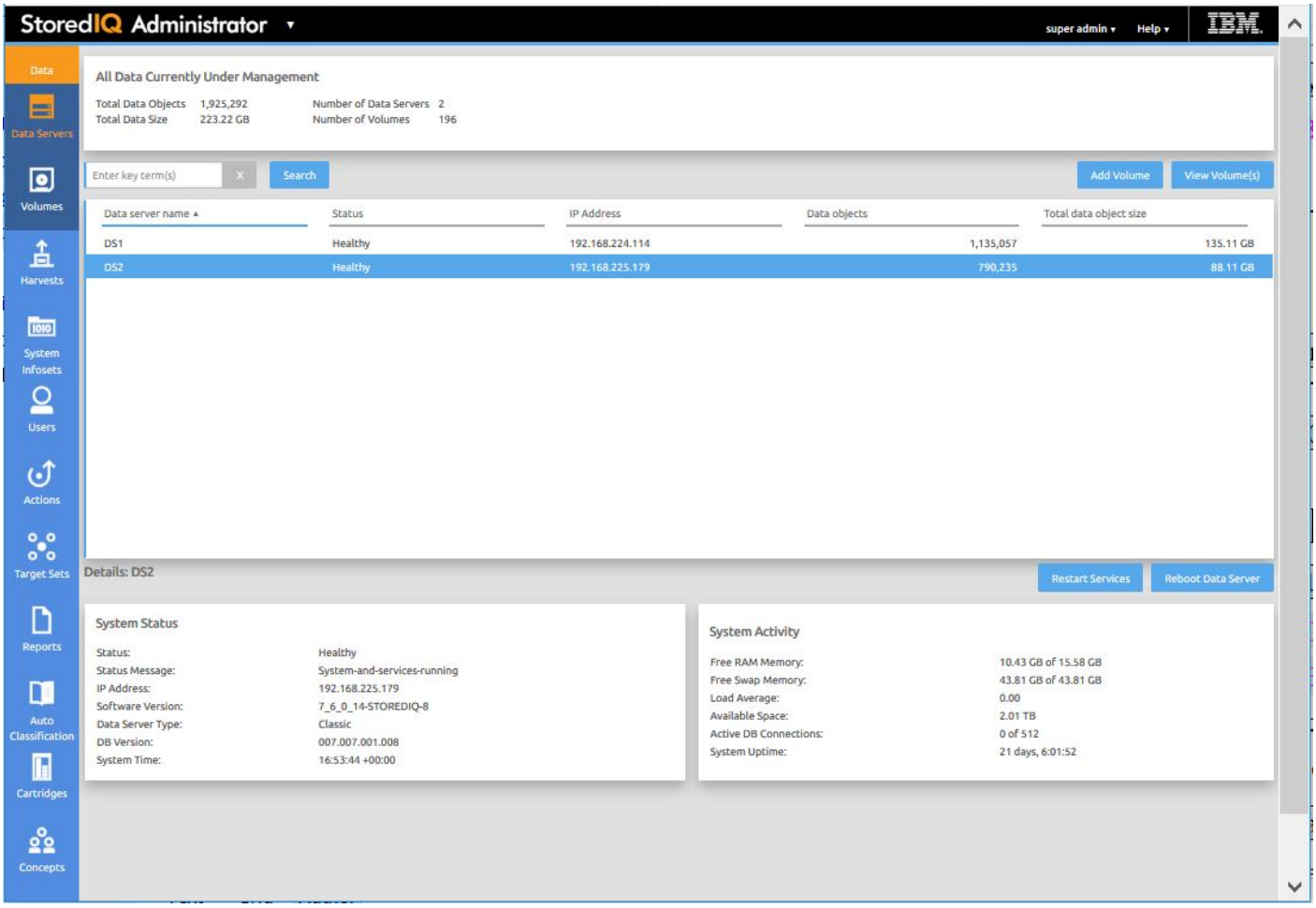
IBM StoredIQ Data Server user interface provides access to data server functionality. It allows administrators to view the dashboard and see the status of the jobs and system details. Administrators

can manage information about servers and conduct various configurations on the system and application settings.



IBM StoredIQ Administrator

IBM StoredIQ Administrator helps you manage global assets common to the distributed infrastructure behind IBM StoredIQ applications.



IBM StoredIQ Administrator provides at-a-glance understanding of the different issues that can crop up in the IBM StoredIQ environment. These views are unique to the IBM StoredIQ Administrator application as they provide an overview of how the system is running. They allow access to various pieces of information that are being shared across applications or allow for the management of resources in a centralized manner.

The administrator is the person responsible for managing the IBM StoredIQ. This individual has strong understanding of data sources, indexes, data servers, jobs, infosets, and actions. This list provides an overview as to how IBM StoredIQ Administrator works:

- **Viewing data servers and volumes**: 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. IBM StoredIQ Administrator also allows the Administrator to see what volumes are currently under management, which data server is responsible for that volume, the state of the volume after indexing, and the amount and size of information that is contained by each volume. Administrators can also add volumes to and delete volumes from data servers through this interface.

  If IBM StoredIQ is configured for integration with Information Governance Catalog, the Administrator can also manage which volumes are published to the governance catalog.
- **Scheduling harvests**: Harvesting, which can also be referred to as indexing, is the process or task by which IBM StoredIQ examines and classifies data in your network. Using IBM StoredIQ Administrator, harvests can be scheduled, edited, and deleted.

- **Creating system infosets**: System infosets that use only specific indexed volumes can be created and managed within IBM StoredIQ Administrator. Although infosets are a core component of IBM StoredIQ Data Workbench, system infosets are created as a shortcut for users in IBM StoredIQ Administrator.
- **Managing users**: The user management area allows administrators to create users and manage users' access to the various IBM StoredIQ applications.
- **Configuring and managing actions**: An action is any process that is taken upon the data that is represented by the indexes. Actions are run by data servers on indexed data objects. Any errors or warnings that are generated as a result of an action are recorded as exceptions in IBM StoredIQ Data Workbench.

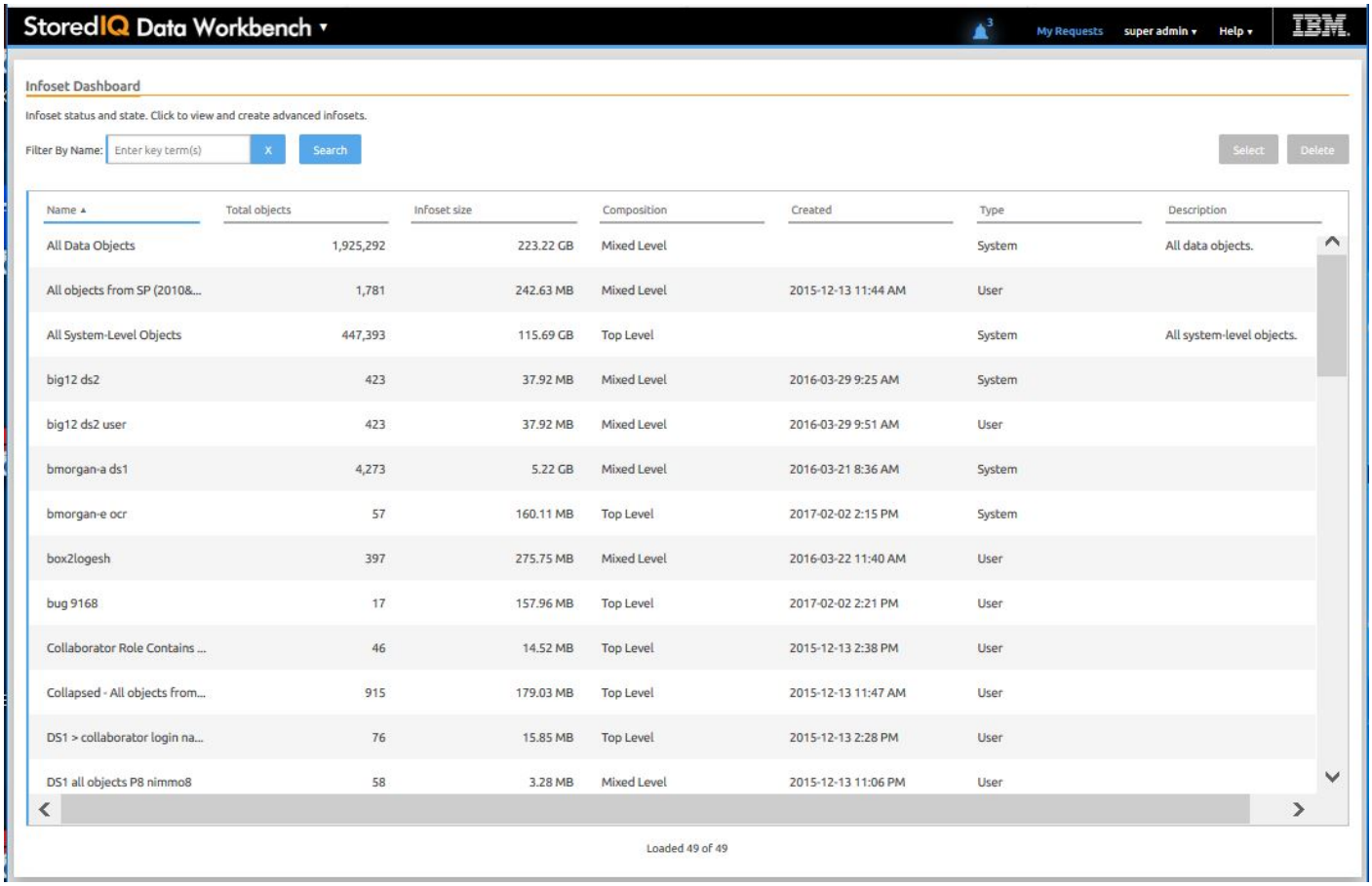
  Note: Actions can be created within IBM StoredIQ Administrator and then made available to other IBM StoredIQ applications such as IBM StoredIQ Data Workbench.
- **Managing target sets**: Provides an interface that allows the user to set the wanted targets for specific actions that require a destination volume for their actions.
- **Reports**: IBM StoredIQ Administrator provides a number of built-in reports, such as summaries of data objects in the system, storage use, and the number of identical documents in the system. You can create custom reports, including Query Analysis Reports for e-discovery purposes, and automatically email report notifications to administrators and other interested parties.
- **Auto-classification**: Automated document categorization, what IBM StoredIQ refers to as auto-classification models, integrates the IBM® Content Classification's classification model into the IBM StoredIQ infoset-generation process. Data Experts can use IBM Content Classification to train a classification model, which is then registered with IBM StoredIQ Administrator. The registered classification model can be applied to an existing infoset in IBM StoredIQ Data Workbench to generate new metadata for the objects in the infoset. Metadata can be used in rule-based filters to create new infosets.
- **Cartridges**: Cartridges are compressed files that contain analysis logic. When you add a cartridge to IBM StoredIQ AppStack, it can detect new data in documents during indexing and make these new insights searchable. For example, a sensitive pattern cartridge can enable IBM StoredIQ to detect passport numbers, phone numbers, and other IDs.

  To apply the analysis logic contained in the cartridge, you must run a Step-up Analytics action that uses the cartridge on an infoset. IBM StoredIQ examines all documents in the infoset, applies the analytics, and then stores the analysis results in the IBM StoredIQ index.
- **Managing concepts**: Provides the ability to relate business concepts to indexed data.
- **Managing Mule scripts**: Helps you to create Mule scripts and upload script packages. These Mule scripts are used by IBM StoredIQ Policy Manager to create policies using the automation workflow.
- **DataServer - Classic**: Data servers can be categorized in two types: DataServer - Classic and DataServer - Distributed. DataServer - Classic refers to the regular data servers. It uses either the current PostgreSQL or Lucene index as an index.
- **DataServer - Distributed**: The distributed data server uses an Elasticsearch cluster instead of an embedded Postgres database. It increases the scalability and flexibility of the IBM StoredIQ deployment in a way that it can manage much larger amounts of data. Without adding more data servers, data that is managed by the IBM StoredIQ deployment can be increased by adding new nodes to the Elasticsearch cluster. Search queries perform better on DataServer - Distributed.
- **Connector API SDK**: A connector is a software component of IBM StoredIQ that is used to connect to a data source such as a network file system and access its data. Using IBM StoredIQ Connector API SDK, developers of other companies can develop connectors to new data sources outside the IBM StoredIQ development environment. These connectors can be integrated with a live IBM StoredIQ application to index, search, manage, and analyze data on the data source.

IBM StoredIQ Data Workbench

Big data is a pervasive problem, not a one-time occurrence. It is easy for most companies to realize that big data is problematic, but it is hard to identify what problems they have. Big data is all about the unknown, but the unknown cannot be off limits. IBM StoredIQ Data Workbench can help you learn about

your data, make educated decisions with your most valuable asset, and turn your company's most dangerous risk into its most valuable asset.



IBM StoredIQ Data Workbench is a data visualization and management tool that helps you to actively manage your company's data. It helps you to determine how much data you have, where it is, who owns it, and when it was last used. When you have a clear understanding of your company's data landscape, IBM StoredIQ Data Workbench helps you take control of data. You can make informed decisions about your data and act on that knowledge by copying, copying to retention, or conducting a discovery export.

Here are just some examples of how you can use IBM StoredIQ Data Workbench.

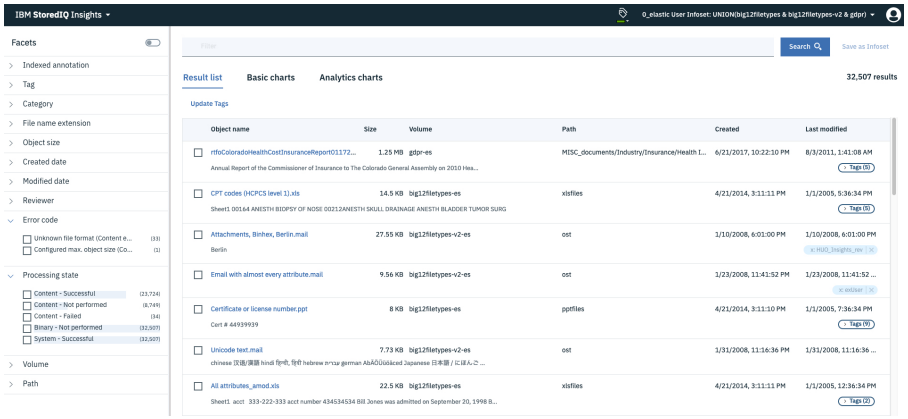
- You need to find all company email that is sent from or received by Eileen Sideways (esideways@thecompany.com). You can use IBM StoredIQ Data Workbench to find all email and then copy that data to a predefined repository. You can also use IBM StoredIQ Data Workbench to find all of the esideways@thecompany.com email that occurred between specific dates and then make that email available for review.
- As an administrator, you want to rid your networks and storage of unused data. You can use IBM StoredIQ Data Workbench to find all files that were not modified in more than five years.
- You want to find all image files that are created in 2007. Not only can IBM StoredIQ Data Workbench find all image files that were created in 2007. It also shows how much space they occupy on your network.
- A user needs to understand how data about Windows is being retained. Using IBM StoredIQ Data Workbench, you can provide that user with a visual overview of the number of objects that are retained and a breakdown of files per data source. Additionally, you can apply overlays to show the user if those files contain forbidden information such as credit-card numbers or Social Security numbers.
- If IBM StoredIQ is configured accordingly, you can select the infosets and filters that are published to the governance catalog for unified governance of structured and unstructured information. When integrating with Information Governance Catalog, you can also analyze and classify the data governed by IBM StoredIQ based on the data classes that are synchronized from the governance catalog.

IBM StoredIQ Insights

IBM StoredIQ Insights provides dynamic and interactive filtering for your data with easy access to all metadata and instant plain-text preview of document content for full-text indexed volumes.

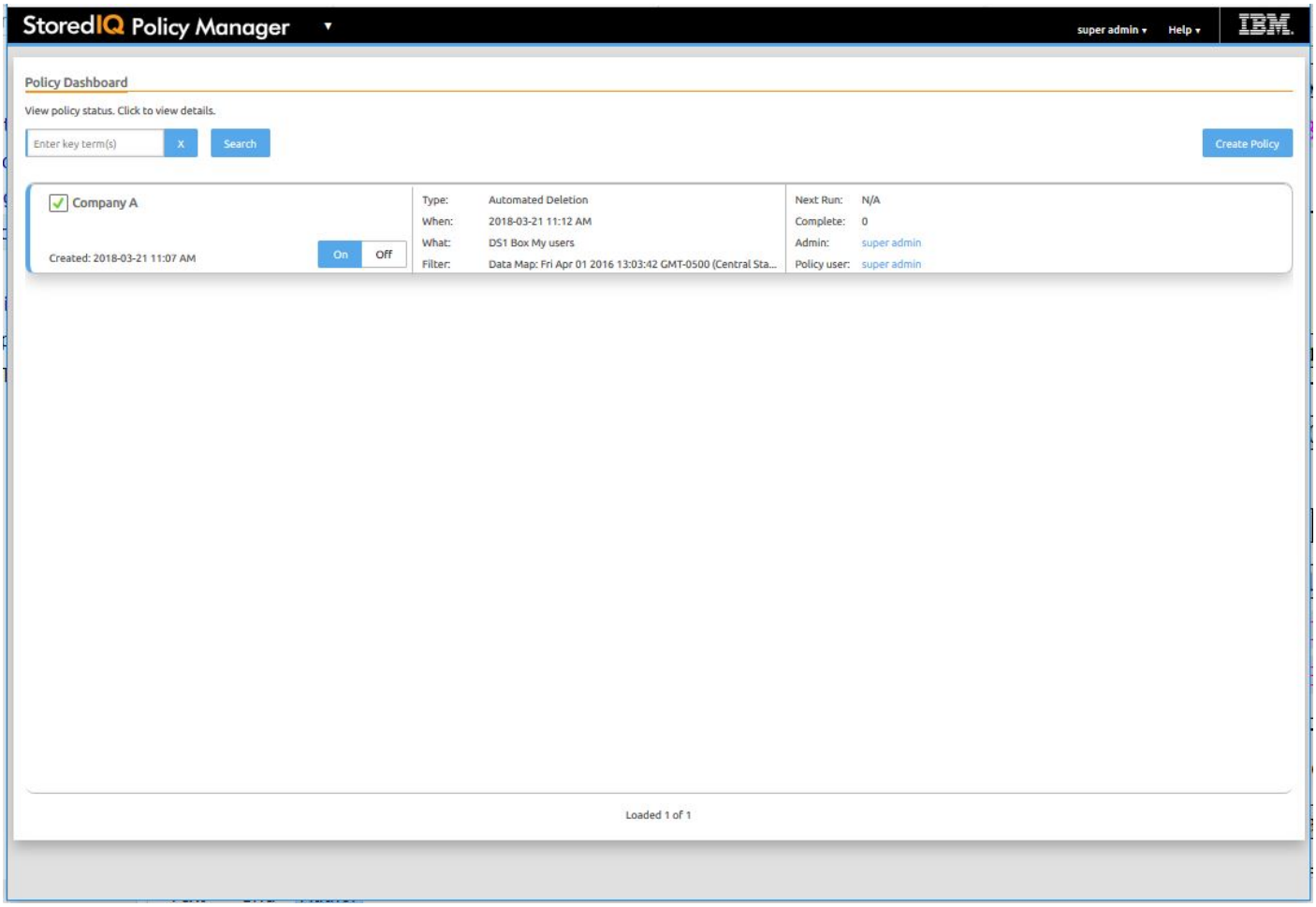
Faceted search lets you drill down to refine your search results as needed. In addition, you can apply any valid IBM StoredIQ filter query. Tags let you categorize the data for easier management. Visual representations of search results help you gain further insights into your data. Several chart types let you look at and explore data from different perspectives, thus helping you identify patterns and relationships very quickly.

With IBM StoredIQ Insights, you can search data that is managed and indexed by a data server of the type DataServer - Distributed. In mixed deployments that have classic and distributed data servers, only the content from distributed data servers will be searchable.



IBM StoredIQ Policy Manager

IBM StoredIQ Policy Manager allows users to run mature policies and processes at scale across a wider range of data.



The users can define and run systemwide policies, focusing on the execution of the process rather than understanding or reviewing affected data objects. Additionally, with reports of IBM StoredIQ Policy Manager, you can record what actions were conducted, when they were conducted, and what data was affected by the policy's execution.

IBM StoredIQ Desktop Data Collector

IBM StoredIQ Desktop Data Collector (also referred to as *desktop client* indexes desktops as volumes. The volumes appear in IBM StoredIQ Data Server and in IBM StoredIQ Administrator, where they can be used like any other data source.

The data server maintains an index using the information sent by the desktop client. After indexing, desktops - even offline or unreachable ones - can be viewed, searched, or targeted for later policy action.

Planning for deployment

When you plan a deployment of IBM StoredIQ, evaluate several infrastructure considerations.

In addition to the information in this section, review the requirements detailed in the IBM Software Product Compatibility Reports (SPCR) tool at: Software Product Compatibility Reports: StoredIQ 7.6

If you plan to use IBM StoredIQ for Legal Identification and Collection to create and manage data boxes and data requests that are to be fulfilled by IBM StoredIQ, also check the system requirements for StoredIQ for Legal at: Software Product Compatibility Reports: StoredIQ for Legal 2.0.3

Generate customized reports with the SPCR tool

Go to the page at Software Product Compatibility Reports to create a high-level report for supported operating systems, related software, hypervisors, and supported translations for any product. You can also create an in-depth report to get detailed system requirements, hardware requirements, and end of service information for each product. You can search for a product in all of the report types and reports are generated based on your query values.

The following report types are the most commonly generated reports from software product compatibility reports:

Detailed system requirements
When you select your product version for the detailed system requirements report, you can set a report filter for **Operating system platforms**, **Product components**, and **Capabilities**, including prerequisites and support software. After you view the report, you can save it as a URL to generate anytime or download it as a PDF.

Hardware requirements
When you select your product version for the hardware requirements report, you can set a report filter by the **Operating system families** option. Set the operating system filter by selecting some or all of the operating systems that are supported by your product. After you view the report, you can save it as a URL to generate anytime or download it as a PDF.

End of service
The end of service report shows the service window of the products that you specify over an eight-year span. For example, you can find out when your product is scheduled to go out of service.

Open Virtual Appliance (OVA) configuration requirements

IBM StoredIQ is deployed as virtual appliances and is supported in VMware ESXi 5.0 (all fix pack levels) or VMware ESXi 6.0 (all fix pack levels) environments. You must have a virtual infrastructure that meets the IBM StoredIQ hardware requirements.

Application stack

- vCPU: 2
- Memory: 4 GB
- Storage:
  - Primary disk (vmdisk1): 21 GB
  - Data disk (vmdisk2): 10 GB

Gateway server

- vCPU: 4
- Memory: 8 GB
- Storage:

- Primary disk (vmdisk1): 100 GB
- Data disk (vmdisk2): 75 GB
- Swap disk (vmdisk3): 40 - 100 GB

DataServer - Classic

- vCPU: 4

  Even though increasing the number of vCPUs increases performance, the actual benefits depend on whether the specific host is oversubscribed or not.

- Memory: 16 GB

  While the minimum value works under light-load condition, as the load increases, the data server quickly starts using swap space. For high load situations, increasing RAM beyond 16 GB can benefit performance. Monitoring swap usage can provide insight.

- Storage:
  - Primary disk (vmdisk1, SCSI 0:0): Default is 150 GB

    This virtual disk has an associated VMDK that contains the IBM StoredIQ operating code. Do not change its size.

    Attention: If you delete the primary disk, you delete the operating system, and the IBM StoredIQ software; the virtual machine might need to be scrapped.

  - Data disk (vmdisk2, SCSI 0:1): Default is 2 TB

    This virtual disk can be resized according to expectations on the amount of harvest data to be stored. For purposes of estimation, the index storage requirement for metadata is about 30 GB per TB of managed source data. Full-text indexing requires an extra 170 GB per TB. Therefore, the default data disk size is targeted for managing 10 TB of source information.

  - Swap disk (vmdisk3, SCSI 0:2): Default is 40 GB

    When under load, the data server can use much RAM; therefore, having ample swap space is prudent. The minimum swap size is equal to the amount of RAM configured for the virtual machine. For best performance under load, place this disk on the highest speed data store available to the host.

The general size limits for a data server are 150 million objects or 500 defined volumes, whichever limit is reached first. Assuming an average object size of 200 KB equals about 30 TB of managed storage across 30 volumes of 5 million objects each, the index storage requirement for metadata on ~30 TB of storage that contains uncompressed general office documents is ~330 GB (11 GB per TB). Add 100 GB per TB of managed storage for full-text or snippet index. For example, to support 30 TB of storage that is indexed for metadata, you need 8 TB indexed for full-text search and extracted text (snippet cache) of 8 TB for auto-classification. A total of 1.9 TB of storage is required (metadata: 330 GB, full-text: 800 GB, snippet cache: 800 GB).

Data-server performance is impacted by the IOPS available from the storage subsystem. For each data server under maximum workload, at least 650 IOPS generally delivers acceptable performance. In the situations where there is a high load on the system, the IOPS that is used can reach up to 7000 with main write operations.

DataServer - Distributed

- vCPU: 4
- Memory: 16 GB
- Storage:
  - Primary disk (vmdisk1, SCSI 0:0): Default is 150 GB
  - Data disk (vmdisk2, SCSI 0:1): Default is 2 TB
  - Swap disk (vmdisk3, SCSI 0:2): Default is 40 GB

If you deploy this type of data server, you must also deploy an Elasticsearch cluster with at least one node. If you deploy a cluster with more nodes, each of the Elasticsearch nodes must meet the listed requirements.

Each Elasticsearch node

- vCPU: 1
- Memory: 32 GB
- Storage:
  - Primary disk (vmdisk1): 100 GB
  - Data disk (vmdisk2): 1 TB

The required memory depends on the index size that is handled by the data node. For a better performance, consider increasing the memory. The total memory available to the node must be distributed between the host operating system, the Elasticsearch java heap space, and the kernel file system caches. For example, if the system has 32 GB memory, 2 GB must be allocated for the host operating system, 15 GB for the java heap space, and the remaining 15 GB for the file system caches.

The data disk (vmdisk2) can be resized according to expectations on the amount of harvest data to be stored.

Network and port requirements

For proper communication, the IBM StoredIQ components must be able to connect to each other.

You must enable network connectivity from the following locations:

- The data server IP address to the gateway IP address on port 11103
- The gateway IP address to and from the application stack IP address on port 8765 and 5432
- Ports 80, 443, and 22 from the administrative workstation to the application stack and data server IP addresses
- Port 22 from the administrative workstation to the gateway IP address.

TCP: port ranges for the firewall

To ensure the network access for desktop volumes, the following port ranges must be open through a firewall.

- 21000-21004
- 21100-21101
- 21110-21130
- 21200-21204

Default open ports

The following ports are open by default on the IBM StoredIQ.

SSH port 22

By default, port 22 is open on all IBM StoredIQ hosts. The port is used for Secure Shell (SSH) communication and allows remote administration access to the VM. In general, traffic is encrypted using password authentication. To add a layer of security, you can establish key-based authentication for passwordless SSH logins to any of the IBM StoredIQ nodes in your environment as described in “Configuring SSH key-based authentication” on page 45.

Default open ports on the AppStack

Port number	Protocol
22	tcp
80	tcp
443	tcp

Default open ports on the IBM StoredIQ data server

Port number	Protocol	Service
22	tcp	PROD-ssh
80	tcp	PROD-web
443	tcp	PROD-https (UI and Web Services APIs)
11103	tcp	PROD-transport (IBM StoredIQ transport services; communication between the gateway and the data server)
11104		

Enable or disable ports or services on the IBM StoredIQ data server

To manage ports, you can use the `/usr/local/storediq/bin/util/port_handler.pyc` script with the appropriate parameter:

```
python /usr/local/storediq/bin/util/port_handler.pyc -parameter
```

-s
To list the current rules in iptables

-l
To list the supported services

-d *port_number*|'*port_range*'
To delete a port or a range of port numbers from iptables, for example:

```
python /usr/local/storediq/bin/util/port_handler.pyc -d '21200:21299'
```

-e '*service_name*'
To enable a specific service, for example, to enable HTTPS services:

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

-d '*service_name*'
To disable a specific service, for example, to disable HTTPS services:

```
python /usr/local/storediq/bin/util/port_handler.pyc -d 'PROD-https'
```

Default open ports on the nodes in the Elasticsearch cluster

Port number	Protocol	Service
21	tcp	ftp

Port number	Protocol	Service
22	tcp	sshd
80	tcp	
443	tcp	
8888	tcp	SimpleHTTPServer (used for copying the `siq-elasticsearch.yml` configuration file from the Elasticsearch node to the data server)
9200	tcp6	docker-proxy (listening for REST requests) You can restrict access to this port by either enabling Search Guard or by setting up a firewall. For more information, see "Securing Elasticsearch cluster communication with Search Guard" on page 51 or "Restricting access to port 9200 on Elasticsearch nodes" on page 52.
9300	tcp6	docker-proxy (internode communication)

Default open ports on the IBM StoredIQ gateway

Port number	Protocol	Service
22	tcp	PROD-ssh
80	tcp	PROD-web
443	tcp	PROD-https (UI and Web Services APIs)
5432	tcp	PROD-postgres
5434	tcp	PROD-transport (IBM StoredIQ transport services; communication between the gateway and the data server)
8765		
7766		
11102		
11103		
11104		

Supported chain and rules on the IBM StoredIQ gateway

In iptables, the following firewall and chain rules are defined:

```
'PROD-transport':['5434','8765','7766','11102','11103','11104'],
            'PROD-https':['443'],
```

```
            'PROD-ssh':['22'],
            'PROD-web':['80'],
            'PROD-postgres':['5432']
```

```
'desktop' service:
            'PROD-broker':['21000'],
            'PROD-collectionsvc':['21300:21399'],
            'PROD-desktopupgrade':['21004'],
            'PROD-objlistmgr':['21100:21199'],
            'PROD-objlistsvc':['21200:21299'],
            'PROD-registration':['21001'],
            'PROD-session':['21002'],
            'PROD-task':['21003'],
```

Open ports for desktop client access to the data server

To open ports for desktop client access to the data server on OVA deployed systems, follow these steps:

1. Log in to the data server as root and run this command:

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

2. Run this command: `iptables -L INPUT`

   In the output of the command, check the list position of the rule that is named `PROD-reject`, for example, the `6th` position on the list.

3. Run this command: `iptables -A INPUT -j PROD-reject`
4. Run this command: `iptables -D INPUT` *`list_position`*

   list_position is the position number of the PROD-reject rule that you determined in step 2.

5. Run the following command:

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

Tip: These steps are required only on an IBM StoredIQ OVA deployed system. The correct ports are open on an upgraded system.

Environment sizing guidelines

To size an environment precisely, you must understand the factors such as harvest frequency, complexity of the source, and use case scenarios that drive application use and action execution.

The general design guidelines for IBM StoredIQ are as follows:

- For data servers of the type DataServer - Classic:
  - One data server for up to 30 TBs of data (which can vary based on the number of volumes, objects per volume, and object types).
  - Up to 500 volumes per data server.

    Tip: When you're sizing an environment that includes Sharepoint data sources, keep in mind that volumes must be defined at Sharepoint site collection level, not the Sharepoint server level.
  - Up to 150 million objects per data server.
- One gateway per 50 data servers.
- One application server.
- NFS is slightly faster than CIFS for metadata only, but assume CIFS/NFS even for this exercise.
- Full-content processing of file (for example, .ZIP, .RAR, .GZ) and email archive (.PST, .NSF, .EMX) processing are slower as items must be extracted from the archives. If there is a significant number of these files in the file system and they are not excluded from content processing, the full-content

processing rate can be too high. Until you have an initial index of the file system, you do not know how to weigh full-content processing of archives.

- An object/time metric is appropriate for metadata only NOT computing a hash, membership in the National Institute of Standards and Technology (NIST) or enumerating objects that are contained in archives. Converting it to a bytes/time metric is a function of the average object size and might vary tremendously. An average object size of 250 KB was used for the metric that is provided earlier.
- A bytes/time metric is appropriate for metadata-only computing a hash and full-content processing. The object per second rate can vary tremendously depending on the object type and sizes encountered. For example, processing an email or file archive is much more expensive than a PDF document.
- Metadata-only not computing a hash, membership in the NIST list, or enumerating objects that are contained in archives is requesting only the file-attribute information from the NAS. Individual files are not opened and read. The processing rate is high, but that does not translate into a large amount of data that traverses a network between the NAS and data server. The bytes/time rate does not translate into bytes served by the NAS and sent over the network.
- Metadata-only computing a hash, membership in the NIST list, or enumerating objects that are contained in archives opens and reads the contents of each file. The content of all requested files traverses the network between the NAS and data server. The maximum load that the data server can place on a NAS is metadata-only processing. It requires all file content to be read to compute a hash or enumerate objects that are contained in archives. The bytes/time rate translates into bytes served up by the NAS and network traffic that must be considered.
- Full-content processing opens and reads the contents of each file to extract all text. The content of all requested files traverses the network between the NAS and data server. The processing time to enumerate archives, extract text, index words, and extract entities on the data server reduces the rate that data is requested from a NAS compared to metadata-only with full hash. The bytes/time rate translates into bytes served up by the NAS and network traffic that must be considered.
- The interrogator process count on the data server for "metadata only not reading all content indexing" is set to eight for optimal performance.
- The interrogator process count for all other processing that involves reading all content default setting is four per data server.
- The interrogator count can be viewed as the number of client connections that are made to a data source actively requesting data. It is important for capacity planning for the data source.
- The data servers are assumed to be "network close" to the NAS data sources. Network latency under 10 ms with at least 1000 Mbps bandwidth is assumed (connected through local area network). The data servers need a low latency high-bandwidth connection to a NAS data source for acceptable indexing performance.
- The gateway and application stack can be located remotely from the data servers. Network connections with latency greater than 10 ms and bandwidth of at least 2+ Mbps are acceptable.

VMware requirements

- VMware vSphere v5.0 and fix packs or v6.0 and fix packs.
- VMware ESXi v5.0 and fix packs, v6.0 and fix packs, or v6.5 and fix packs.
- VMware virtual machine hardware version 8.0 or later. For more information, see the VMware product documentation.
- The appropriate VMware license to enable the required processor cores and memory for the virtual machine.

Stack-provisioning prerequisites

Before a deployment, verify that you meet these prerequisites.

- At least one physical server with sufficient processor, RAM, and hard disk configuration for the planned management project.

- VMware ESX or ESXi on CD/DVD or USB drive.
- IP addresses, cables, and physical switch ports for at least the ESXi/ESX interface, one data server, one gateway server, and one application stack.
- Network connectivity that is enabled from the following locations:
  - The data server IP address to the gateway IP address on port 11103
  - The gateway IP address to and from the application stack IP address on port 8765 and 5432
  - Ports 80, 443, and 22 from the administrative workstation to the application stack and data server IP addresses
  - Port 22 from the administrative workstation to the gateway IP address.
- Network connectivity that is enabled from the data server IP address to any data sources to be harvested and managed.
- A management station computer or notebook from where the load-management work is done.

License usage metrics

Using the IBM License Metric Tool, you can generate license consumption reports that count IBM StoredIQ license usage.

IBM StoredIQ is licensed by Resource Value Unit (RVU). RVU calculation is based on terabytes IBM StoredIQ.

On the IBM StoredIQ application stack, a license program writes usage information to an IBM Software Licence Metric Tag (SLMT) file. This file has the extension .slmtag and can be read periodically by the IBM License Metric Tool (ILMT) after it has been configured to scan for these files. You can generate reports that summarize usage.

By default, the license program retrieves the size of the **All Data Objects** infoset in terabytes once per day and writes this information to the `/var/siq/ilmt/3cd1469042433ee7010fe09f661dc67b.slmtag` file. The .slmtag file can store information up to a maximum file size of 1 MB, after which the file is archived and a new log file is created. A maximum of 10 log files are kept.

The .slmtag file contains usage information in the following format, where new metric records are appended to the end of the file:

```
<SchemaVersion>2.1.1</SchemaVersion>
<SoftwareIdentity>
        <PersistentId>cb98e260a2a14872902578de1b8e2016</PersistentId>
        <Name>IBM StoredIQ Data Assessment</Name>
        <InstanceId>/var/siq/ilmt</InstanceId>
</SoftwareIdentity>
<Metric logTime="2019-01-25T15:58:26+00:00">
        <Type>TERABYTE</Type>
        <SubType>All Data Objects</SubType>
        <Value>0.005</Value>
        <Period>
                <StartTime>2019-01-25T15:58:26+00:00</StartTime>
                <EndTime>2019-01-25T15:58:26+00:00</EndTime>
        </Period>
</Metric>
```

Integration with IBM License Metric Tool

Versions of IBM License Metric Tool (ILMT) that support IBM Software License Metric Tag (SLMT) can generate license consumption reports. An ILMT agent can scan in configurable intervals the file system for .slmtag files, collect information, and send it to the corresponding ILMT server. ILMT reports the number of terabytes managed by IBM StoredIQ. This number is to be used as input for the RVU License Conversion Table specified in the license information (`li_`*`languagecode`* file) that comes with IBM StoredIQ. On the application stack, you can find the license information in the `License` directory.

For more information about using IBM License Metric Tool, see the IBM License Metric Tool documentation.

Security

Plan and implement specific security measures to protect the application and the data it manages, especially when you deploy IBM StoredIQ into sensitive environments.

IBM StoredIQ keeps your data secure through encryption, security hardening, and auditing.

Federal Information Processing Standard (FIPS)

FIPS is a standard recommended by the National Institute of Standards and Technology (NIST) and the US Federal Government. It ensures certain security standards are met for software or hardware components deployed at US government sites. Enabling FIPS ensures that the SSL/TLS engine that is compliant with the US Government recommendation is used. IBM StoredIQ supports FIPS Level 1.

Secure gateway communication can be enabled without FIPS. If FIPS is enabled, IBM StoredIQ uses FIPS compliant versions of OpenSSL.

Secure communication and encryption of data in motion

In a production environment, you should configure or install certificates on the AppStack to enable HTTPS communication and to enable encryption of data in motion between the browser and the AppStack. You can to this during installation and initial configuration or at any time afterward. For details, see the instructions for configuring certificates.

The gateway handles the communication between the data servers and the application stack. By default, the communication between the gateway, any data servers, and the AppStack is in plain text and is not encrypted. If your enterprise security policy mandates encryption of data in motion, enable secure gateway communication. In this case, secure gateway communication must be configured on all three IBM StoredIQ components. You can enable secure gateway communication during installation and initial configuration or at any time afterward. For details, see “Managing the status of secure gateway communication” on page 54.

IBM StoredIQ then uses stunnel to ensure secure communication between the components. If your environment includes data servers of the type DataServer - Distributed, stunnel can also be used to encrypt the communication between the nodes within the Elasticsearch cluster but not for encrypting the communication between the data server and the Elasticsearch cluster.

To secure the communication between the data server and the Elasticsearch cluster and the communication within the Elasticsearch cluster likewise, you can enable Search Guard. For more information, see “Securing Elasticsearch cluster communication with Search Guard” on page 51. If you don't want to do that but still want to restrict client access to port 9200 on the Elasticsearch nodes, you can set up the firewall accordingly. For more information, see “Restricting access to port 9200 on Elasticsearch nodes” on page 52.

If FIPS is not enabled, the following cipher suites and encryption algorithm are used for data at rest:

TLS_ECDHE_ECDSA_WITH_AES_128_GCM_SHA256
TLS_ECDHE_RSA_WITH_AES_128_GCM_SHA256

You can configure these cipher suites in the configuration files listed in the list of key and certificate files. However, if you run the utilities for enabling stunnel, you might need to make the respective configuration changes again.

Encryption of data at rest

Starting with IBM StoredIQ version 7.6.0.15, the disk volume on which the Elasticsearch indexes are stored is encrypted by default. IBM StoredIQ uses Linux Unified Key Setup (LUKS) for disk encryption. For details about key management, see “Key and certificate management” on page 41.

Optionally, you can encrypt the application data on the IBM StoredIQ application stack. For more information, see "Enabling encryption of IBM StoredIQ AppStack application data" on page 50.

Network isolation

If full-text harvesting and Step-up Analytics actions (cartridges) are applied, Elasticsearch indexes can contain potentially sensitive content. Therefore, you should deploy the Elasticsearch nodes in an isolated location on the network (for example, as an enclave or behind a firewall) that is properly secured according to the sensitivity of the data being harvested. Only the IBM StoredIQ application stack and data servers should be allowed to communicate with the Elasticsearch nodes.

Also, any data servers and the gateway should be deployed in an isolated network location to allow for communication with authorized clients only.

Access control

The following administrative accounts are required. The `builder` and `siqadmin` accounts are IBM StoredIQ-specific accounts. For more information about these accounts, see "Default user accounts" on page 17.

`root` and `builder` accounts on the Elasticsearch cluster nodes
Remote login for `root` can be disabled. However, local root login is required, either log in as `root` or use the **su** command to obtain root permissions temporarily.

You set the passwords for the `root` and `builder` accounts during the configuration process when you start the VM for the first time. You can change these passwords anytime.

`siqadmin` account on the AppStack
Administration of the AppStack usually does not require direct root access. For day-to-day administration, the `siqadmin` account can be used.

You set the password for the `siqadmin` account during the configuration process when you start the VM for the first time. You can change this password anytime.

Default user accounts

IBM StoredIQ comes with a set of default user accounts.

For security reasons, change the passwords for these default accounts after the installation is complete so that they are unique and different from the default values. The new password must be 8 to 64 characters long and must contain characters from at least three of these categories:

- Uppercase letters: A - Z
- Lowercase letters: a - z
- Digits: 0 - 9
- Punctuation marks or special characters: . : = * ^ / $ # _ + @ & % -

Account	Default password	Description
`admin`	`admin`	Administrative account for accessing IBM StoredIQ Data Server. Use this account for the initial setup of the data server and to create further administrative accounts for routine administration. Change the password in IBM StoredIQ Data Server under **Administration** > **Configuration** > **Manage users**.

Account	Default password	Description
`audituser`	`Passw0rd!`	Account for accessing the audit database on the AppStack. Change the password by running the **change_audituser_password** script as `siqadmin` user on the AppStack.
`builder`	None. Password is set during initial configuration.	Account for setting up and configuring the Elasticsearch cluster. You can change the password by using the Linux **passwd** command.
`reportuser`	`St0red1q`	Account for accessing the report database on the AppStack. Change the password by running the **change_reportuser_password** script as `siqadmin` user on the AppStack.
`siqadmin`	`Passw0rd!`	Administrative account for managing the AppStack server. For more information, see the administration guide. With new installations, you are prompted for a new password during the initial configuration of the AppStack. In updated deployments, change the password by using the Linux **passwd** command.
`superadmin`	`admin`	Administrative account for accessing IBM StoredIQ Administrator and IBM StoredIQ Data Server. Use this account for the initial setup of IBM StoredIQ and to create further administrative accounts for routine administration so that their actions can be audited. Change the password in IBM StoredIQ Administrator: go to **Users** and edit the `superadmin` account. The `superadmin` account has access to all IBM StoredIQ applications on the application stack. To switch applications, click the triangle arrow icon and select the application that you want to access from the list of the available applications.

Deploying IBM StoredIQ

IBM StoredIQ is a virtual appliance that you deploy and configure in a VMWare virtual host environment.

Deploying the virtual appliances

Use VMware vSphere Client to deploy the virtual appliances to an ESX server. Deploy OVAs for the gateway, the data server, and the application stack, preferably in this order. If an Elasticsearch cluster is also being deployed, those OVAs must be deployed first.

Ensure that all prerequisites for the deployment described in the planning section are met and that the required software packages are available on your local system before you start this task.

Download the latest version of IBM StoredIQ from either IBM Fix Central or IBM Passport Advantage®. For information about the package names and part numbers and the links to the proper download locations, see the download document.

The number of Elasticsearch OVAs deployed depends on the planned size of your Elasticsearch cluster. The default setup is a three-node cluster. Each Elasticsearch node requires a separate OVA deployment.

The number of data server OVAs deployed depends on the number of data servers needed. Each data server requires a separate OVA deployment.

1. Connect to the ESX server or vCenter server.
2. Open VMware vSphere Client.
3. Select **File** > **Deploy OVF Template**.
4. Within the **Deploy OVF Template** wizard, complete these steps.
   a) Within the **Select source** page, click **Local file**, and then browse to and select the appropriate OVF. Click **Next**.
   b) Within the **Review details** page, review the OVF template details.
      These storage requirements are critical and will be used to select a data store during deployment. Click **Next** to proceed.
   c) Within the **Select name and folder** page, enter a name for the deployed template or use the default name. Click **Next**.
   d) Within the **Select a resource** page, select the resource pool where the deployed OVF template runs. Click **Next**.
   e) Within the **Select storage** page, select a data store on which to store the deployed OVF template files. Click **Next**.
   f) Within the **Disk Format** list, select the disk format to be deployed. Note that although **Thin Provision** saves disk space, it can negatively affect performance. If possible, select **Thick Provision Lazy Zeroed**. Click **Next**.
   g) Within the **Network Mapping**, map the network to the appropriate network. Click **Next**.
   h) Within the **Ready to Complete** page, review the deployment settings. Click **Finish** to complete the Deploy OVF Template.
      At this point, you can select the **Power on** check box to turn on the VM after deployment is complete.

      Important: If your IBM StoredIQ environment includes an Elasticsearch cluster, do not select this option when you're deploying the data server OVA. The Elasticsearch cluster setup must be complete before you run the data server first-boot process.
5. Repeat steps 3 and 4 for each OVA.

Configure the components of your IBM StoredIQ environment in this order:

1. Elasticsearch cluster (if applicable)
2. Gateway
3. One or more data servers
4. Application stack

Deploying IBM StoredIQ on Microsoft Hyper-V

As an alternative to installing on an ESX server, the IBM StoredIQ gateway, data server, and application stack can be installed on Microsoft Hyper-V. This option is not supported for the Elasticsearch virtual appliance.

Installing IBM StoredIQ on Microsoft Hyper-V requires the use of third-party software. For the procedure described here, the following software prerequisites must be met:

- Microsoft Hyper-V Manager must be installed on a Windows system.
- 7-zip for Windows or tar for the Linux operating system must be available to extract contents of the OVA.
- 7-zip for Windows or gunzip for the Linux operating system must be installed to decompress the `vmdk.gz` file.
- Microsoft Virtual Machine Converter 3.0 (Windows) or qemu-img (Linux) or must be installed to convert the files from VMWare `.vmdk` files to Hyper-V `.vhdx` files.

Microsoft Hyper-V does not support OVAs and OVFs. The virtual machine needs to be built manually.

Consider the following instructions a sample procedure. The instructions might differ depending on the version of the third-party software.

Complete these steps for the gateway, the data server, and the AppStack:

1. Extract the `vmdk` file from the OVA.
   a) Download the OVA.
   b) Extract the contents of the OVA by using 7-zip for Windows or tar for the Linux operating system.
   c) Decompress the `vmdk.gz` archive by using 7-zip for Windows or gunzip for Linux.
   d) Delete everything except the `vmdk` file when the OVA extraction is complete.
2. Convert the VMWare .vmdk file to a Microsoft Hyper-V .vhdx file.
   For instructions about converting .vmdk files to .vhdx files, see the following links:
   - Use Microsoft Virtual Machine Converter
   - Using qemu-img
3. Build a virtual machine.
   a) Select **New Virtual Machine** > **Next** from Hyper-V Manager.
   b) Enter the name of the virtual machine in the **Specify Name and Location** window and click **Next**.
   c) Select **Generation of the machine** in the **Specify Generation** window.
   d) Enter the startup memory in the **Assign Memory** window.
   e) Select the correct network in the **Configure Network** window and click **Next**.
   f) Select **Use an existing virtual hard disk** in the **Connect Virtual Hard Disk** window. Search the disk location, select one of the primary disks, and then click **Next**.
   g) Click **Finish**.
   h) Repeat these previous steps for all IBM StoredIQ disks.
4. Optional: Add disks.
   a) Select a virtual machine in Microsoft Hyper-V Manager.
   b) Select the settings for the virtual machine under **Actions**.

c) Select the IDE controller and highlight Hard Drive in the **Settings** window.

d) Select Hard Drive and virtual hard disk and click **New**.

e) Click **Next** in the New Virtual Hard Disk wizard.

f) Select **VHDX** in Choose Disk Format and then click **Next**.

g) Select **Dynamically expanding** in Choose Disk Type and then click **Next**.

h) Enter a name for the disk in the **Specify Name and Location** window and the click **Next**.

i) Select **Create a new blank virtual hard disk** and enter a size in the **Configure Disk** window and then click **Finish**.

j) Repeat this procedure for all additional disks.

Complete the installation process by following the instructions in "Configuring IBM StoredIQ" on page 22.

Configuring IBM StoredIQ

After the OVF template deployment, proceed to configure the IBM StoredIQ virtual machines.

Configuring the Elasticsearch cluster

For a deployment of the type distributed, you must deploy and configure an Elasticsearch cluster before you run the data server first-boot process and configure a data server of the type DataServer - Distributed. A classic deployment (data server of the type DataServer - Classic) does not require an Elasticsearch cluster.

The default setup is a three-node Elasticsearch cluster. However, in smaller enterprises or in non-production environments a single node Elasticsearch cluster can be set up.

Configuring a three node Elasticsearch cluster

For a default setup, deploy and configure a three node Elasticsearch cluster.

Ensure that three Elasticsearch OVAs are deployed.

Complete the setup of the Elasticsearch cluster before configuring any of the other IBM StoredIQ components.

1. In VMware vSphere Client, check the VM settings that were created from the OVA and change them if required.

   Primarily, check these settings:

   Memory
   Must be at least 32 GB

   Hard disk 1 (primary disk)
   Must be at least 100 GB

   Hard disk 2 (data disk)
   Must be at least 1 TB
2. Power on the VM.
3. Open the console to launch the configuration wizard (first-boot procedure).

   Depending on your version of vSphere Client, you might either have a **Launch Console** or an **Open Console** link to do so.

   The wizard is text based. To work in the wizard, use the Tab key to navigate, the Space bar to select items, and the Enter key to apply your selections.

   a) Accept the CentOS and IBM Eula license agreements when prompted.

   b) Set up the passwords for the `root` and `builder` accounts.

   With the `builder` account, you can later log in to the virtual machine via SSH.

   c) Configure the network.

   Within the **Network Configuration** window, select either **Static** or **Obtain IP via DHCP**.

   - If you select the **Static IP** option, complete these fields to configure the static address:

Parameter	Value
Hostname	The fully qualified host name of the Elasticsearch node
IP Address	The IPv4 address of the Elasticsearch node
Netmask	The netmask for the assigned IP address

Parameter	Value
Gateway	The IP address of the default gateway for the IP subnet **Note:** This is the network gateway, not the IBM StoredIQ gateway.
Primary DNS Host	The IP address for the domain name server

- If you select the **Obtain IP via DHCP** option, provide this information:

Parameter	Value
Hostname	The fully qualified host name of the Elasticsearch node

Also select **Restart network services** to restart the node after your configuration is complete.

4. Repeat steps "1" on page 22 to "3" on page 22 for each OVA for a minimum of three nodes.
5. Log in to the first node in the Elasticsearch cluster that is used to set up or upgrade the cluster across the other nodes.

   Use the `builder` account and the password that you configured in step "3" on page 22.

   ```
   ssh builder@primary_es_node_ip
   ```

6. Copy the sample `cluster-setup.properties` file to the `builder` home directory.

   At the prompt `[builder@localhost ~]$`, enter the following command:

   ```
   cp /etc/siq/cluster-setup.properties.sample cluster-setup.properties
   ```

7. Edit the file by using this command:

   ```
   vi cluster-setup.properties
   ```

   The file defines the following properties. Only configuration of the ES_HOSTS property is mandatory.

~/cluster-setup.properties	Note
CLUSTER_NAME=Elasticsearch	Cluster name string.
ES_HOSTS=192.0.2.0,192.0.2.10,192.0.2.24	Required. The first entry in the list becomes node1.
STUNNEL_ENABLED=false	Set to `true` to have the inter-node communication within the Elasticsearch cluster encrypted using stunnel. If you want to use Search Guard to secure the Elasticsearch cluster, this property must be set to `false`.
COUNTRY_ABBR=US	Used for x509 certificate generation.
STATE_PROVINCE=Texas	Used for x509 certificate generation.
ES_USER=elasticsearch	Do not modify.
CITY=Austin	Used for x509 certificate generation.
COMPANY_NAME=IBM	Used for x509 certificate generation.
DEPT_NAME=StoredIQ	Used for x509 certificate generation.
CONTACT_EMAIL=storediqsupport@us.ibm.com	Used for x509 certificate generation.
EXPIRATION_DAY=3650	Used for x509 certificate generation.

~/cluster-setup.properties	Note
LUKS=true	Do not modify.
SEARCHGUARD_ENABLED=false	Set to true to secure all communication with and within the Elasticsearch cluster by using Search Guard. Use of Search Guard requires additional configuration steps. For details, see "Securing Elasticsearch cluster communication with Search Guard" on page 51.

8. Run the cluster setup with the properties file that you created.

   At the prompt `[builder@localhost ~]$`, enter:

   ```
   /siq/bin/cluster-setup.sh cluster-setup.properties
   ```

   The setup script generates `properties/configs` for each node and runs the setup against each node by using SSH.
9. Accept the server certificate and enter the password for each node when prompted.

   After the script is executed successfully, the setup of the Elasticsearch cluster is complete.

   Data is stored in the `/siq/var/data/elasticsearch` directory. Log files are written to the `/siq/var/log/elasticsearch` directory.
10. Test the setup by using these commands:

    ```
    curl -X GET 'http://primary_es_node_ip:9200'
    curl -X GET 'http://primary_es_node_ip:9200/_cluster/health?pretty=true'
    curl -X GET 'http://primary_es_node_ip:9200/_cluster/state?pretty'
    ```

Configure the other IBM StoredIQ components: gateway, data server, and application stack.

Deploying a single node Elasticsearch cluster

Deploy a single node Elasticsearch cluster, for example, in a non-production environment.

One Elasticsearch OVA must be deployed.

Complete the setup of the Elasticsearch cluster before configuring any of the other IBM StoredIQ components.

1. In VMware vSphere Client, check the VM settings that were created from the OVA and change them if required.

   Primarily, check these settings:

   Memory
   Must be at least 32 GB

   Hard disk 1 (primary disk)
   Must be at least 100 GB

   Hard disk 2 (data disk)
   Must be at least 1 TB
2. Power on the VM.
3. Open the console to launch the configuration wizard (first-boot procedure).

   Depending on your version of vSphere Client, you might either have a **Launch Console** or an **Open Console** link to do so.

   The wizard is text based. To work in the wizard, use the Tab key to navigate, the Space bar to select items, and the Enter key to apply your selections.

   a) Accept the CentOS and IBM Eula license agreements when prompted.

   b) Set up the passwords for the `root` and `builder` accounts.

With the `builder` account, you can later log in to the virtual machine via SSH.

c) Configure the network.

Within the **Network Configuration** window, select either **Static** or **Obtain IP via DHCP**.

- If you select the **Static IP** option, complete these fields to configure the static address:

Parameter	Value
Hostname	The fully qualified host name of the Elasticsearch node
IP Address	The IPv4 address of the Elasticsearch node
Netmask	The netmask for the assigned IP address
Gateway	The IP address of the default gateway for the IP subnet **Note:** This is the network gateway, not the IBM StoredIQ gateway.
Primary DNS Host	The IP address for the domain name server

- If you select the **Obtain IP via DHCP** option, provide this information:

Parameter	Value
Hostname	The fully qualified host name of the Elasticsearch node

Also select **Restart network services** to restart the node after your configuration is complete.

4. Log in to the Elasticsearch node.

   Use the `builder` account and the password that you configured in step 3.

   ```
   ssh builder@es_node_ip
   ```

5. Copy the sample `cluster-setup.properties` file to the `builder` home directory.

   At the prompt `[builder@localhost ~]$`, enter the following command:

   ```
   cp /etc/siq/cluster-setup.properties.sample cluster-setup.properties
   ```

6. Edit the properties file:

   ```
   vi cluster-setup.properties
   ```

   The file defines the following properties. Only configuration of the ES_HOSTS property is mandatory.

~/cluster-setup.properties	**Note**
`CLUSTER_NAME=Elasticsearch`	Cluster name string.
`ES_HOSTS=192.0.2.0`	Required. Enter a single IP address.
`STUNNEL_ENABLED=false`	If you want to use Search Guard to encrypt the communication between the data server and the Elasticsearch node, this property must be set to `false`.
`COUNTRY_ABBR=US`	Used for x509 certificate generation.
`STATE_PROVINCE=Texas`	Used for x509 certificate generation.
`ES_USER=elasticsearch`	Do not modify.
`CITY=Austin`	Used for x509 certificate generation.

~/cluster-setup.properties	Note
COMPANY_NAME=IBM	Used for x509 certificate generation.
DEPT_NAME=StoredIQ	Used for x509 certificate generation.
CONTACT_EMAIL=storediqsupport@us.ibm.com	Used for x509 certificate generation.
EXPIRATION_DAY=3650	Used for x509 certificate generation.
LUKS=true	Do not modify.
SEARCHGUARD_ENABLED=false	Set to true to secure all communication with the Elasticsearch node by using Search Guard. Use of Search Guard requires additional configuration steps. For details, see "Securing Elasticsearch cluster communication with Search Guard" on page 51.

7. Run the cluster setup with the properties file that you created and the --single-node option:

   ```
   /siq/bin/cluster-setup.sh cluster-setup.properties --single-node
   ```

   The setup script generates properties/configs for the node.
8. Accept the server certificate and enter the password for the node when prompted.

   After the script is executed successfully, the setup of the Elasticsearch cluster is complete.

   Data is stored in the /siq/var/data/elasticsearch directory. Log files are written to the /siq/var/log/elasticsearch directory.
9. Make sure the discovery.zen.minimum_master_nodes property in the docker-compose.yml file is set to 1.

   Use the following command:

   ```
   head -20 /siq/env/docker/docker-compose.yml
   ```

   If the property is set to any other value, complete the following steps:

   a) Stop Elasticsearch by using this command:

   ```
   sudo systemctl stop elasticsearch
   ```

   b) Edit the docker-compose.yml file to set the minimum number of master nodes to 1:

   ```
   vi /siq/env/docker/docker-compose.yml
   ```

   Save the change and exit the file.

   c) Start Elasticsearch again by using this command:

   ```
   sudo systemctl start elasticsearch
   ```
10. Test the setup by using these commands:

    ```
    curl -X GET 'http://es_node_ip:9200'
    curl -X GET 'http://es_node_ip:9200/_cluster/health?pretty=true'
    curl -X GET 'http://es_node_ip:9200/_cluster/state?pretty'
    ```

    Because the cluster contains only one node, replicas are not enabled. Therefore, the node status will be yellow. To update the node status to green, set the number of replicas to 0 for each index:

    ```
    curl -X PUT "http://es_node_ip:9200/volume_*/_settings" -d '{"index":
    {"number_of_replicas" : 0 }}'
    ```

Configure the other IBM StoredIQ components: gateway, data server, and application stack.

Adding a node to an existing Elasticsearch cluster

You can add nodes to any Elasticsearch cluster.

To ensure that the cluster works correctly and all nodes are available, enter

```
curl -X GET 'http://primary_node_ip:9200/_cluster/state?pretty'
curl -X GET 'http://primary_node_ip:9200/_nodes?pretty'
```

Where *primary_node_ip* is the IP address of either the single Elasticsearch node or the primary Elasticsearch node in a cluster of at least three nodes.

Adding a node

You can add a single node to any Elasticsearch cluster with three or more nodes. With a single node deployment, you must add at least two nodes as to create a default three node Elasticsearch cluster.

1. Download the Elasticsearch OVA version that matches the current cluster version from Fix Central and deploy it to vCenter.
2. In VMware vSphere Client, check the VM settings that were created from the OVA and change them if required.

   Primarily, check these settings:

   Memory
   Must be at least 32 GB

   Hard disk 1 (primary disk)
   Must be at least 100 GB

   Hard disk 2 (data disk)
   Must be at least 1 TB
3. Open the console to launch the configuration wizard (first-boot procedure).

   Depending on your version of vSphere Client, you might either have a **Launch Console** or an **Open Console** link to do so.

   The wizard is text based. To work in the wizard, use the Tab key to navigate, the Space bar to select items, and the Enter key to apply your selections.

   a) Accept the CentOS and IBM Eula license agreements when prompted.

   b) Set up the passwords for the `root` and `builder` accounts.

   With the `builder` account, you can later log in to the virtual machine via SSH.

   c) Configure the network.

   Within the **Network Configuration** window, select either **Static** or **Obtain IP via DHCP**.

   - If you select the **Static IP** option, complete these fields to configure the static address:

Parameter	Value
Hostname	The fully qualified host name of the Elasticsearch node
IP Address	The IPv4 address of the Elasticsearch node
Netmask	The netmask for the assigned IP address
Gateway	The IP address of the default gateway for the IP subnet **Note:** This is the network gateway, not the IBM StoredIQ gateway.
Primary DNS Host	The IP address for the domain name server

   - If you select the **Obtain IP via DHCP** option, provide this information:

Parameter	Value
Hostname	The fully qualified host name of the Elasticsearch node

Also select **Restart network services** to restart the node after your configuration is complete.

4. Log in to the Elasticsearch node, where the installation was initially started, by using the `builder` account and password that were configured during the installation of the cluster.
5. Edit the `cluster-setup.properties` file by appending the new node's IP address to the ES_HOSTS line. For example, to add a fourth node to a three node cluster, the ES_HOSTS line looks like:

   ```
   ES_HOSTS=192.0.2.0,192.0.2.10,192.0.2.24,198.51.100.0
   ```

   where `198.51.100.0` is the new node's IP address.
6. Run the cluster setup script. At the prompt `[builder@localhost ~]$`, enter:

   ```
   /siq/bin/cluster-setup.sh cluster-setup.properties
   ```
7. Accept the server certificate and enter the password for the new node when prompted.

   After the script is executed successfully, the new node is part of the Elasticsearch cluster.
8. Add the new node's IP address and port information to `/usr/local/storediq/etc/siq-elasticsearch.yml` on the data servers that use this Elasticsearch cluster.

   Your .yml file should look similar to the example where `198.51.100.0` is the new node's IP address:

   ```
   cluster:
    name: Elasticsearch
   nodes:
   - host: 192.0.2.0
     port: 9200
   - host: 192.0.2.10
     port: 9200
   - host: 192.0.2.24
     port: 9200
   - host: 198.51.100.0
     port: 9200
   ```
9. Enter the command to restart services on the data servers that are using this Elasticsearch cluster:

   ```
   service deepfiler restart
   ```
10. To ensure that the cluster works correctly and all nodes are available, enter:

    ```
    curl -X GET 'http://primary_node_ip:9200'
    curl -X GET 'http://primary_node_ip:9200/_cluster/health?pretty=true'
    curl -X GET 'http://primary_node_ip:9200/_cluster/state?pretty'
    ```

Configuring the gateway

Configure the gateway settings.

- Each VM requires a dedicated IP address. You cannot use one IP address for multiple installations on the same system.
- It is helpful to know your application stack's IP address before you install the gateway. If you do not have this information at this point, you must edit the gateway configuration after the installation is complete to add this information. For more information, see "Updating the gateway configuration" on page 57.

1. In vSphere Client, power on the virtual machine.
2. Open the console to launch the configuration wizard.

   Depending on your version of vSphere Client, you might either have a **Launch Console** or an **Open Console** link to do so.

The wizard is text based. To work in the wizard, use the Tab key to navigate, the Space bar to select items, and the Enter key to apply your selections.

3. Accept the CentOS and IBM license agreements when prompted.
4. Within the **Corporate Network** window, select either **Static** or **Obtain IP via DHCP**.
   - If you select the **Static IP** option, complete these fields to configure the static address:

Parameter	Value
Hostname	The fully qualified host name
Corporate IP	The IPv4 address of the gateway
Netmask	The netmask for the assigned IP address
Gateway	The IP address of the default gateway for the IP subnet **Note:** This is the network gateway, not the IBM StoredIQ gateway.
Primary DNS Host	The IP address for the domain name server

   - If you select the **Obtain IP via DHCP** option, complete these fields:

Parameter	Value
Hostname	The fully qualified host name

5. Click **Next**.
6. In the **Enter root password** window, enter the new password twice, and then press **Enter** when finished.

   The installation of the IBM StoredIQ application starts.

   Note: Depending on the performance of your virtual infrastructure, this process can be lengthy, taking from just minutes to multiple hours.
7. In the **Select FIPS mode** window, you can select to run your system in FIPS-compliant security mode.

   By default, FIPS mode is not enabled.
8. In the **Gateway communication security** window, optionally select **Enable secure gateway communication** to encrypt the communication between the gateway, the data server, and the application stack.

   By default, the communication is in plain text and is not encrypted.

   Secure communication via stunnel can impact performance. Therefore, enable this setting only if your enterprise security policy mandates encryption of data in motion. If you do so, the IBM StoredIQ application stack and the data server must be configured accordingly. Also, check the gateway settings as described in the section about configuring DA Gateway settings in the data server administration documentation.

   You can change the enablement status at any time after the installation.

   For more information, see “Security” on page 16. For additional guidance, contact IBM Support.
9. Select **Done**.
10. At the **Appstack IP** prompt, enter the application stack's IP address.

    If you do not have this information at hand at this point, you can update the gateway configuration later. For more information, see “Updating the gateway configuration” on page 57.

    The installation is finalized and the server restarts.

Configuring the data server

Configure the data server system settings.

Each VM requires a dedicated IP address. You cannot use one IP address for multiple installations on the same system.

1. In vSphere Client, power on the virtual machine.
2. Open the console to launch the configuration wizard.

   Depending on your version of vSphere Client, you might either have a **Launch Console** or an **Open Console** link to do so.

   The wizard is text based. To work in the wizard, use the Tab key to navigate, the Space bar to select items, and the Enter key to apply your selections.
3. Accept the CentOS and IBM license agreements when prompted.
4. Within the **Corporate Network** window, select either **Static** or **Obtain IP via DHCP**.
   - If you select the **Static IP** option, complete these fields to configure the static address:

Parameter	Value
Hostname	The fully qualified host name
Corporate IP	The IPv4 address of the data server
Netmask	The netmask for the assigned IP address
Gateway	The IP address of the default gateway for the IP subnet **Note:** This is the network gateway, not the IBM StoredIQ gateway.
Primary DNS Host	The IP address for the domain name server

   - If you select the **Obtain IP via DHCP** option, complete these fields:

Parameter	Value
Hostname	The fully qualified host name

5. Click **Next**.
6. In the **Enter root password** window, enter the new password twice, and then press **Enter** when finished.
7. In the **Select FIPS mode** window, you can select to run your system in FIPS-compliant security mode.

   By default, FIPS mode is not enabled.
8. If the data server type that is deployed is DataServer - Distributed, select **Enable elasticsearch storage**.

   For a default Elasticsearch deployment, enter the IP address of the first node of the Elasticsearch cluster where the cluster was set up. For a single node Elasticsearch deployment, enter the IP address of the Elasticsearch node.
9. Select **Next**.
10. Optional: Select **Enable secure gateway communication** to encrypt the communication between the data server and the gateway. Also, enter the IP address of the IBM StoredIQ gateway.

    By default, the communication is in plain text and is not encrypted.

    Secure communication via stunnel can impact performance. Therefore, enable this setting only if your enterprise security policy mandates encryption of data in motion. If you do so, the IBM StoredIQ gateway and the application stack must be configured accordingly.

You can change the enablement status at any time after the installation.

For more information, see “Security” on page 16. For additional guidance, contact IBM Support.

11. Select **Done** and press **Enter**.

    The installation of the IBM StoredIQ application starts.

    Note: Depending on the performance of your virtual infrastructure, this process can be lengthy, taking from just minutes to multiple hours.

12. Optional: For a data server of the type DataServer - Distributed, verify that the `/usr/local/storediq/etc/siq-elasticsearch.yml` exists and contains the proper information.

    For a default Elasticsearch deployment, the `-host:` entry in the file must show the IP addresses of all nodes in the Elasticsearch cluster. With a single node Elasticsearch cluster setup, the entry must contain the IP address of this Elasticsearch node.

    If the file does not exist on the data server, copy the `/etc/siq/siq-elasticsearch.yml` from the Elasticsearch node to the data server. In a three node cluster, you can find the file on the primary node.

After the data server virtual machine and the deepfiler services are up, complete the data server configuration in the IBM StoredIQ Data Server user interface:

- Check the gateway settings and modify them if required. For more information, see the topic about gateway settings in the administration guide.
- Verify the network settings. For more information, see the topic about network settings in the administration guide.
- Configure mail server settings as described in the topic about mail server settings in the administration guide.
- Verify system date and time settings. For more information, see the topic about system date and time settings in the administration guide.

Configuring the application stack

Configure the application stack settings.

If you do not want to obtain the application stack IP dynamically, you should have the following information at hand:

- The static IP address for the application stack
- The netmask information for this IP address
- The IP address of the default network gateway for the IP subnet

Remember that IP addresses must be unique within your IBM StoredIQ deployment.

If you want to configure email notification with an authenticated user, make sure to complete the instructions in “Configuring authenticated users for SMTP notifications” on page 38 before configuring the SMTP settings.

If you plan to enable the synchronization with a governance catalog, a working deployment of one of these products must be available:

- IBM Information Server. The minimum required version is Version 11.7.0.
- IBM Cloud Private for Data. The minimum required version for full support is Version 1.2.

For details about this feature, see the information about integrating with IBM Information Governance Catalog.

1. In vSphere Client, power on the virtual machine.
2. Open the console to launch the configuration wizard.

   Depending on your version of vSphere Client, you might either have a **Launch Console** or an **Open Console** link to do so.

The wizard is text based. To work in the wizard, use the Tab key to navigate, the Space bar to select items, and the Enter key to apply your selections.

3. Accept the CentOS and IBM license agreements when prompted.
4. In the **Password Utility** window, enter the new passwords for the `root` user and the `siqadmin` user twice and press **Enter**.
5. Within the **Corporate Network** window, select either **Static IP** or **Obtain IP via DHCP**.
   - If you select the **Static IP** option, complete these fields to configure the static address:

Parameter	Value
Hostname	The fully qualified host name
IP Address	The IPv4 address of the application stack
Netmask	The netmask for the assigned IP address
Gateway	The IP address of the default gateway for the IP subnet **Note:** This is the network gateway, not the IBM StoredIQ gateway.
Primary DNS Host	The IP address for the domain name server

   - If you select the **Obtain IP via DHCP** option, complete these fields:

Parameter	Value
Hostname	The fully qualified host name
Primary DNS Host	The IP address for the domain name server

6. Select the **Restart services** option to commit the IP and restart services.
7. Click **Next**.

Within the **Appstack configuration** window, set these options.

8. In the **Domain name** field, enter the fully qualified domain name (FQDN) or IP address of the application stack.

   This information is used in generated URLs, such as links to reports.

   If you enable the synchronization with the governance catalog, the domain name is used for building the base URLs for REST access to the application stack and for links to IBM StoredIQ artifacts that make these artifacts accessible from the governance catalog.

   Important: If you ever need to change the host name or IP address of the application stack (using the **appstackcfg** utility), you must restart all application stack services afterward by running the command `service appstack restart` from the command line. Selecting the **Restart appstack services** option is not sufficient because this option triggers the restart of only the `uwsgi` and `tomcat` services.
9. In the **StoredIQ Gateway** field, enter the IP address of the StoredIQ Gateway server.
10. Set the following SMTP options to enable the application stack's capability to send and receive notification email.

    Tip: All SMTP settings are optional and can be configured during or after your deployment. If you choose to set or change the SMTP settings at a later time, see “Configuring the application stack to send and receive reports and notifications” on page 58.

    a) Set these options:

Parameter	Value
Server	The mail server's fully qualified domain name or IP address.
Port	The SMTP port. The default port is 25.
Username	The login user name. For the default configuration, leave this field empty. Otherwise, provide the user name of the user with which to authenticate to the Exchange server. • If you authorized any Authenticated User, you can use any user name and password as long as that individual is valid member of the domain. • If you used a specific user, you must use the user name of the single user for which you granted permissions. This must be a fully qualified user name. In this case, you must have completed the instructions in "Configuring authenticated users for SMTP notifications" on page 38 before configuring the SMTP settings.
Password	The login password for the specified user. For the default configuration, leave this field empty.

b) Select **Enable TLS** to enable TLS encryption, if it is supported by the mail server.

For email notification with an authenticated user, enable this option.

For additional information about SMTP notification, see "Configuring authenticated users for SMTP notifications" on page 38.

11. Optional: Set these options to enable the synchronization of specific objects between IBM StoredIQ and a governance catalog.

Tip: These settings can be configured during or after the deployment. If you choose to set or change the synchronization settings at a later time, see "Configuring the application stack to synchronize data with the governance catalog" on page 58.

If the synchronization is not enabled, the values entered here are not validated. However, as soon as you enable synchronization, all entries must be valid. Otherwise, a warning is displayed and synchronization is implicitly disabled.

a) Select **Enable synchronization with the governance catalog** and provide the following settings.

If the data catalog to which you want to publish the IBM StoredIQ object resides in an IBM Cloud Private for Data environment, select the **Server runs in IBM Cloud Private for Data** check box.

Provide or accept the values for these fields:

Parameter	Value
Host	The host name or IP address of the Information Server or IBM Cloud Private for Data installation. The specified host is part of the base URLs for REST access to the Information Governance Catalog or IBM Cloud Private for Data instance

Parameter	Value
	and for links to catalog artifacts that make these artifacts accessible from IBM StoredIQ. Therefore, you should provide the fully qualified domain name of the Information Server or IBM Cloud Private for Data host. If you specify a server port, this port also becomes part of such base URLs. In addition, the host name is also used to address the Information Server Apache Kafka server, which provides all Information Server events as Kafka messages. Specific Kafka messages are consumed by IBM StoredIQ and trigger the synchronization of objects from the governance catalog to IBM StoredIQ.
Port	The port of the governance catalog server. This setting is optional. For connections to an Information Server 11.7 environment: To ensure proper communication, you should set the port to the HTTPS port that is defined in Information Server. The default port is 9443. For connections to an Information Server 11.7 FP1 (or later) environment without Information Server Enterprise Search installed: To ensure proper communication, you should set the port to the HTTPS port that is defined in Information Server. The default port is 9446.. For connections to an Information Server 11.7 FP1 (or later) environment with Information Server Enterprise Search installed: Do not specify a port. For connections to IBM Cloud Private for Data, you can specify the IBM Cloud Private for Data port.
Kafka port	The port of the Information Server Kafka server. The port setting can be overridden. For connections to an Information Server 11.7 environment: the Kafka port defined in Information Server. The default port is 59092. For connections to an Information Server 11.7 FP1 (or later) environment without Information Server Enterprise Search installed: the Kafka port defined in Information Server. The default port is 59092. For connections to an Information Server 11.7 FP1 (or later) environment with Information Server Enterprise Search installed: the Kafka port defined in Information Server. The default port is 9092.

Parameter	Value
	For connections to IBM Cloud Private for Data: the Kafka port defined in IBM Cloud Private for Data. For more information, see the topic Enabling synchronization with IBM StoredIQ in the IBM Cloud Private for Data product documentation.
Username	The user name for authenticating to Information Server or IBM Cloud Private for Data when publishing IBM StoredIQ objects to the governance catalog. This user must be defined in Information Server with the following security roles: Suite User Common Metadata Administrator Information Governance Catalog Information Asset Administrator In IBM Cloud Private for Data, this user must be defined with the Data Stewart role.
Password	The password of the user set with **Username**.
Sync frequency (minutes)	Data is periodically propagated to the governance catalog at the specified interval. The value must be a positive number of minutes. The default value is 15 minutes.
StoredIQ instance name	The name identifying the IBM StoredIQ instance for which data is synchronized. This name can be freely chosen, but must be unique within the governance catalog instance.

12. Optional: Select **Enable FIPS mode at boot time** to enable running your system in FIPS-compliant security mode. By default, FIPS mode is not enabled.
13. Optional: Select **Enable secure gateway communication** to encrypt the communication between the application stack and the gateway.

    By default, the communication is in plain text and is not encrypted.

    Secure communication via stunnel can impact performance. Therefore, enable this setting only if your enterprise security policy mandates encryption of data in motion. If you do so, the IBM StoredIQ gateway and the data server must be configured accordingly. For more information, see "Security" on page 16. For additional guidance, contact IBM Support.

    You can change the enablement status at any time after the installation.
14. Click **Next**.

    Important: For synchronization with the governance catalog to work, HTTPS must be enabled on the AppStack. Therefore, generate and install at least a self-signed certificate.
15. Optional: Within the **Certificate configuration** window, perform the procedure in its entirety to generate a self-signed SSL or TLS certificate.

    SSL or TLS certificates are used to establish secure communications. You can generate self-signed certificates, which should be used in test and development environments only, or certificates that are signed by an internal or a third-party certificate authority (CA). To avoid certificate trust issues, you should obtain and install a certificate signed by a third-party CA.

To skip certificate configuration, tab to **Exit** and click **Enter**. The certificate can be configured at a later time by logging in as `siqadmin` user and running this command: `certcfg`

Important: Synchronization with the governance catalog requires a certificate to be installed. Therefore, do not skip certificate generation now if you enabled the synchronization.

If you choose to generate a certificate, complete the steps of this procedure in the described sequence. In the wizard, use the Up and Down Arrow keys to navigate between options and the space bar to select an option.

a) Generate a self-signed root certificate. Make sure that option 1 is selected and press Enter.

The resulting certificate can be used as a certificate authority (CA).

If you want to use a root certificate from a third party CA to sign your certificates, you can skip this step.

The following table lists the configuration settings for a self-signed root certificate. Required settings are denoted by an asterisk. Edit the settings as required.

Table 1. Creating self-signed root certificate

Setting	**Value**
Common Name *	The name of the certificate. You can use the prefilled value or choose a different name. However, make sure not to use the AppStack host name.
Email Address *	The email address that is used in the certificate.
Country (two-letter) *	An acceptable entry is an ISO-3166-1 alpha-2 code. A listing is available here.
State/Province *	
City *	
Organization *	
Department *	
Key length	The length of the key to be created. The default value is 2048.
Days to expiry *	The number of days before the certificate expires. The default value is 3650.
Root key location *	The fully qualified file name of the root key file. This name can be freely chosen. If the file does not exist, it is created when the certificate is created. However, it is recommended to use the prefilled default file name.
Root certificate location	The fully qualified file name of the root certificate file. This name can be freely chosen. If the file does not exist, it is created when the certificate is created. However, it is recommended to use the prefilled default file name.

Click **Next** to proceed to creating a certificate-signing request.

b) Create a certificate-signing request. Make sure that option 2 is selected and press Enter.

Generate a certificate signing request (CSR) to be signed by a certificate authority. The process creates a key or uses a provided key and generates the CSR from it.

The following table lists the settings for a certificate signing request. Required settings are denoted by an asterisk. Edit the settings as appropriate.

Table 2. Creating certificate signing request

Setting	**Value**
Common Name *	The host name of the AppStack. It must match the domain of the URL that you use. **Important:** This value is prefilled. If you create a self-signed root certificate, make sure to change this value so that it is different from the common name of Step “15.a” on page 36.
Email Address *	
Country (two-letter) *	An acceptable entry is an ISO-3166-1 alpha-2 code. A listing is available here.
State/Province *	
City *	
Organization *	
Department *	
Key length	The length of the key to be created. The default value is 2048.
Key location *	The fully qualified file name of the key file. This name can be freely chosen. If the file does not exist, it is created when the certificate is created. However, it is recommended to use the prefilled default file name.
Certificate request location *	The fully qualified file name of the CSR file. This name can be freely chosen. If the file does not exist, it is created when the request is created. However, it is recommended to use the prefilled default file name.

Click **Next** to proceed to signing the certificate.

c) Generate the signed certificate. Make sure that option 3 is selected and press Enter.

Sign a certificate with a certificate-authority-eligible root certificate based on the certificate-signing request.

The following table lists the settings for signing the certificate. All settings are required. Edit them as appropriate.

Table 3. Signing request with root certificate

Setting	**Value**
Days to expire	The number of days before the certificate expires. The default value is 3650.
Certificate request location	The fully qualified file name of the CSR to sign as specified in step “15.b” on page 36.

Table 3. Signing request with root certificate (continued)	
Setting	**Value**
Certificate location	The fully qualified file name of the signed certificate. This name can be freely chosen. If the file does not exist, it is created when the certificate is created. However, it is recommended to use the prefilled default file name.
Root key location	The fully qualified file name of the root key file as specified in step “15.a” on page 36. Or, if you chose to use a root certificate signed by a third-party CA, the fully qualified file name of the respective root key file. However, it is recommended to create a copy of the third-party root key file with the default name assigned by IBM StoredIQ.
Root certificate location	The fully qualified file name of the root certificate as specified in step “15.a” on page 36. Or, if you chose to use a root certificate signed by a third-party CA, the fully qualified file name of the respective root certificate. However, it is recommended to create a copy of the third-party root certificate with the default name assigned by IBM StoredIQ.

Click **Next** to proceed to updating the AppStack HTTPS certificate.

d) Update the AppStack HTTPS certificate. Make sure that option 4 is selected and press Enter.

Update the application stack to use the provided certificate and key for HTTPS access.

The following table lists the settings for updating the certificate. These settings are required and are prefilled with the information from the previous steps. Do not change these settings.

Table 4. Updating the AppStack HTTPS certificate	
Setting	**Value**
Key location	The fully qualified file name of the key file.
Certificate location	The fully qualified file name of the certificate.

Click **Finish** to complete the certificate configuration.

The application stack installation begins. When the installation is done, the virtual machine restarts to open a console login prompt. The installation and configuration of the application stack is complete.

You can open a browser to log in to the IBM StoredIQ applications. In the address bar, enter the IP address or the host name that you configured in step “5” on page 32. Remember to specify the address in the format `https://`*`IP_address`* or `https://`*`hostname`* if you enabled HTTPS in step “15” on page 35. Use the credentials of the default administrative account for IBM StoredIQ Administrator if you log in for the first time: user ID `superadmin`, password `admin`

Configuring authenticated users for SMTP notifications

The IBM StoredIQ application-stack configuration requires that the username and password be blank for the default configuration in order to send email notifications from the application stack. In order to use an authenticated user with the IBM StoredIQ application-stack configuration, you must provide authorization in both Microsoft Exchange and Active Directory. While it is possible to use the default SMTP Receive Connector, it is recommended that a custom SMTP Receive Connector be created and configured to work

with the IBM StoredIQ application stack. If a custom SMTP Receive Connector is created and defined with the specific IP address of the IBM StoredIQ application stack, network traffic from the application stack will utilize that connector.

Creating an SMTP Receive Connector
Using an SMTP Receive Connector, email messages are transmitted to the Exchange Server for processing.

1. In the Exchange Management Console navigate, to the **Server Configuration** > **Hub Transport** section. Select **New Receive Connector** from the right-side menu.

   The New Receive Connector dialog box appears.
2. In the **Name** text box, enter a name for the new SMTP Receive Connector, and then click **Next**.
3. Configure your local network settings. By default, you can utilize port 25, which is listed by default port. For better security, you can utilize a different port, but you will need to verify that port is open and available.
4. Specify the **Fully Qualified Domain Name (FQDN)**. For example, you might enter `Exchange2010.ibmlab.local`, and then click **Next**.
5. On the Remote Network settings page, click **Add**, and enter the valid IP address for the IBM StoredIQ application stack. Click **OK**.

   The IBM StoredIQ application stack appears within the IP addresses.
6. Select the IP address range of **0.0.0.0-255.255.255.255**, and then click the red **X** to delete it. Click **Next**.
7. Review the configuration details, and then click **New**.

   If the new receive connector is valid, a green check mark and successful completion message will appear.
8. Click **Finish**. The new connector appears within the list of available receive connectors.
9. Restart the Microsoft Exchange Transport service. Go to **Start > Administrative Tools > Services**, and select **Microsoft Exchange Transport** service. Right-click, and select **Restart**.

Authenticating users
There are two methods for authenticating users for SMTP notification: authentication against existing domain users and a single, specific domain user.

1. Log into your Exchange Server.
2. Launch the Active Directory Services Interfaces (ADSI) editor. Click **Start > Administrative Tools > ADSI Edit**.
3. Within ADSI, navigate to **CN=Configuration > CN=Services > CN=First Organization > CN=Administrative Groups > CN=Exchange Administrative Group > CN=Servers > CN=Exchange Server > CN=Protocols > CN=SMTP Receive Connectors**, and select **CN=SMTP Receive Connectors**.

   The newly created SMTP Receive Connector is displayed as one of the available options.
4. Right-click the newly created SMTP Receive Connector and select **Properties**.

   The **Properties** window opens.
5. Click **Security**, click **Add**, and the **Select Users, Computers, Service Accounts, or Groups** dialog box appears.
6. In the **Enter the object names to select** text box, enter either of the following:
   - Enter `Authenticated Users`, and then click **Check Names**. Because `Authenticated Users` is a valid entry, a line appears underneath it.
   - Enter a valid, existing user ID and then click **Check Names**. In order to enter a specific user, that individual must already exist within Active Directory. Because that user exists within Active Directory, it enters the fully qualified user name and domain. For example, if a user ID of `storediqsupport` was created in Active Directory, you would enter the user ID `storediqsupport`.
7. Authenticate the user, either existing domain users or a single, specific domain user.

- If you are authenticating existing domain users, in the **Group or user names** area, select **Authenticated Users**. In the **Permissions for Authenticated Users**, select the **Allow** check box for **Accept Any Sender** and **Accept Authoritative Domain Sender**.
- If you are authenticating a single, specific domain user, in the **Group or user names** area, select the specific user ID. In the **Permissions for Authenticated Users**, select the **Allow** check box for **Accept Any Sender** and **Accept Authoritative Domain Sender**.

8. Click **OK**.

Optional post-installation configuration

After the initial configuration of the system is complete, you can proceed with some mostly security-related optional configuration steps. At any time, update configurations as required.

Key and certificate management

IBM StoredIQ uses keys and certificates to ensure secure communication and encrypt data.

The following table shows the key files, certificates, and configuration files used by the AppStack, the data servers, and the gateway for secure communication. You can replace all keys and self-signed certificates that IBM StoredIQ uses.

Table 5. List of keys and certificates

Component	Key or certificate location	Configuration file	Comment
AppStack Application	`/etc/siq/ssl/client.crt`	`/siq/svc/nginx/ssl.conf`	Run the **certcfg** utility as `siqadmin` user to manage certificates.
Data Server Web application	`/usr/local/apache/conf/ssl.crt/server.crt`	`usr/local/apache/conf/httpd.conf`	Hashing algorithm: PKCS #1 SHA-256 with RSA
AppStack stunnel	`/etc/siq/ssl/stunnel.pem`	`/etc/siq/stunnel.conf`	If you want to use your own certificates, see "Installing your own certificates" on page 43 for details. For details about stunnel key management, see "Managing disk encryption keys for Elasticsearch index volumes" on page 41.
Gateway stunnel	`/etc/deepfile/gateway/stunnel.pem`	`/etc/deepfile/gateway/stunnel.conf`	
Data Server stunnel	`/etc/deepfile/dataserver/stunnel.pem`	`/etc/deepfile/dataserver/stunnel.conf`	
ElasticSearch stunnel	`/siq/env/docker/node1-stunnel.pem`	`/siq/env/docker/node1-stunnel.conf`	

Managing disk encryption keys for Elasticsearch index volumes

IBM StoredIQ uses LUKS to encrypt the disk volume on which the Elasticsearch indexes are stored. You can add your own private keys or passphrases to the LUKS encryption system, or even remove the encryption key that is generated when the cluster is set up.

To protect the private key from unauthorized access, you must be an administrator with root access to be able to complete this procedure.

LUKS provides eight slots to specify the encryption keys. You can manage keys and passphrases by using the **cryptsetup** command.

1. Check which device is encrypted by running the following command:

```
[builder@hostname ~]$ su
Password:
[root@hostname builder]$ blkid -t TYPE=crypto_LUKS
```

The output looks similar to what is shown in this example:

```
/dev/sdb1: UUID="3755df51-cf96-46d9-b5b1-b301a0284bc8" TYPE="crypto_LUKS"
PARTLABEL="primary" PARTUUID="66715c0d-0e22-4e44-a5a3-cd5cd9bbdb5f"
```

By default, the encrypted device should be /dev/sdb1.

2. Check the LUKS key slots to detect which slots are enabled and which slots are disabled.

```
[root@hostname ~]# cryptsetup luksDump /dev/sdb1 | grep "Key Slot"
```

The output looks similar to what is shown in the example:

```
Key Slot 0: ENABLED
Key Slot 1: DISABLED
Key Slot 2: DISABLED
Key Slot 3: ...
```

3. Add a new passphrase.
   - Add a new passphrase to the next available key slot.

     ```
     [root@hostname ~]# cryptsetup luksAddKey /dev/sdb1 -d /root/siq-elasticsearch-luks.key
     Enter new passphrase for key slot:
     Verify passphrase:
     ```

     Check again which key slots are enabled and which ones are disabled:

     ```
     [root@hostname]# cryptsetup luksDump /dev/sdb1 | grep "Key Slot"
     Key Slot 0: ENABLED
     Key Slot 1: ENABLED
     Key Slot 2: DISABLED
     Key Slot 3: DISABLED
     Key Slot 4: ...
     ```

     To remove the passphrase from this slot, you can use this command:

     ```
     [root@hostname ~]# cryptsetup luksKillSlot /dev/sdb1 1 -d /root/siq-elasticsearch-luks.key
     ```

     Check again which key slots are enabled and which ones are disabled:

     ```
     [root@hostname]# cryptsetup luksDump /dev/sdb1 | grep "Key Slot"
     Key Slot 0: ENABLED
     Key Slot 1: DISABLED
     Key Slot 2: DISABLED
     Key Slot 3: DISABLED
     Key Slot 4: ...
     ```

   - You can also add the passphrase to a specific slot by using the -S option.

     ```
     [root@hostname ~]# cryptsetup luksAddKey /dev/sdb1 -S 3 -d /root/siq-elasticsearch-luks.key
     Enter new passphrase for key slot:
     Verify passphrase:
     ```

     Check again which key slots are enabled and which ones are disabled:

     ```
     [root@hostname]# cryptsetup luksDump /dev/sdb1 | grep "Key Slot"
     Key Slot 0: ENABLED
     Key Slot 1: ENABLED
     Key Slot 2: DISABLED
     Key Slot 3: ENABLED
     Key Slot 4: ...
     ```

4. Instead of using a passphrase, you can use a secret key file.
   a) Copy your secret key file to the /root folder.

You can create a random key file by running this command, where `siq-elasticsearch-luks-new.key` is the name of the new key file:

```
dd if=/dev/random of=/root/siq-elasticsearch-luks-new.key bs=2048 count=1
```

b) Change the file permission for this file so that only the root user has read access to it:

```
[root@hostname ~]# chmod 0400 /root/siq-elasticsearch-luks-new.key
```

c) Add the key to the next available key slot.

```
[root@hostname ~]# cryptsetup luksAddKey /dev/sdb1 /root/siq-elasticsearch-luks-new.key -d /root/siq-elasticsearch-luks.key
```

d) Check the key slots.

```
[root@hostname ~]# cryptsetup luksDump /dev/sdb1 | grep "Key Slot"
Key Slot 0: ENABLED
Key Slot 1: ENABLED
Key Slot 2: DISABLED
Key Slot 3: DISABLED
Key Slot 4: ...
```

e) Add the new key file to the `/etc/crypttab` file.

The updated file should look similar to the example:

```
siq-elasticsearch-luks UUID=4f2a067d-6604-46a5-8d0b-8a387a40198b /root/siq-elasticsearch-luks.key luks
siq-elasticsearch-luks UUID=4f2a067d-6604-46a5-8d0b-8a387a40198b /root/siq-elasticsearch-luks-new.key luks
```

5. Optional: Remove the key that is shipped with the product for immediate use.

a) Delete the key from slot 0:

```
[root@hostname ~]# cryptsetup luksKillSlot /dev/sdb1 0 -d /root/siq-elasticsearch-luks-new.key
```

b) Make sure key slot 0 is disabled:

```
[root@hostname ~]# cryptsetup luksDump /dev/sdb1 | grep "Key Slot"
Key Slot 0: DISABLED
Key Slot 1: ENABLED
Key Slot 2: DISABLED
Key Slot 3: DISABLED
Key Slot 4: ...
```

c) Delete the `/root/siq-elasticsearch-luks.key` entry from the `/etc/crypttab` file so that is just contains the `/root/siq-elasticsearch-luks-new.key` entry:

```
siq-elasticsearch-luks UUID=4f2a067d-6604-46a5-8d0b-8a387a40198b /root/siq-elasticsearch-luks-new.key luks
```

If you set a passphrase instead of using a key file,and remove the default key file, you will have to provide this passphrase on each start of the VM.

Important: If you remove a key without adding a new key or a passphrase, you will lock yourself out of the encrypted device. In this case, you will no longer be able to access the device and the data that is stored on this encrypted device is permanently lost.

Installing your own certificates

You can install and use your own certificates when secure gateway communication using stunnel is enabled.

Back up your virtual machines before replacing the certificates. The following procedure requires root authority, even on the AppStack. To perform these steps, `siqadmin` authority is not sufficient.

If you enable secure gateway communication using stunnel, IBM StoredIQ generates its own self-signed root certificate and uses that one. If you want to use your own root certificates along with the IBM StoredIQ intermediate certificate, follow these instructions.

The following certificates and keys must be available:

- Root certificate

- Intermediate certificate (gateway certificate and key)
- AppStack certificate and key
- Data server certificate and key (one for each data server)
- Certificates for the Elasticsearch nodes (one for each Elasticsearch node)

For some steps of this procedure, you must copy files between servers. Use a tool such as **scp** or **sftp** to do so.

For the gateway, the AppStack, and each data server, complete these steps:

1. Copy the root certificate to the gateway, the AppStack, and each data server.

   The `ca.pem` root certificate must be copied to the following locations:

   - On the gateway: `/etc/deepfile/gateway/certs`
   - On the data server: `/etc/deepfile/dataserver/certs`
   - On the AppStack: `/etc/siq/ssl/certs/`
2. Create key pairs on the gateway, the AppStack, and each data server.

   A key pair is a combination of a key and a certificate. For example, you can use the **cat** command to create a combined PEM file as shown here:

   - On the gateway:

     ```
     cat gateway.key gateway.pem > gateway.keypair.pem
     ```
   - On each data server:

     ```
     cat dataserver.key dataserver.pem > dataserver.keypair.pem
     ```
   - On the AppStack:

     ```
     cat appstack.key appstack.pem > appstack.keypair.pem
     ```
3. Copy the AppStack and data server certificates to the gateway.

   Store all certificates (PEM files) in the certificate directory. Copy the `gateway.pem` file, the `dataserver.pem` file, and the `appstack.pem` file to the `/etc/deepfile/gateway/certs` directory.
4. Copy the gateway certificate to the AppStack and to each data server.

   Store the PEM file in the certificate directory.

   - On the data server, copy the `gateway.pem` file to the `etc/deepfile/dataserver/certs` directory.
   - On the AppStack, copy the `gateway.pem` file to the `/etc/siq/ssl/certs` directory.
5. Rehash the certificate directories.

   Use the following commands:

   - On the gateway:

     ```
     /usr/sbin/cacertdir_rehash /etc/deepfile/gateway/certs
     ```
   - On each data server:

     ```
     /usr/sbin/cacertdir_rehash /etc/deepfile/dataserver/certs
     ```
   - On the AppStack:

     ```
     /usr/sbin/cacertdir_rehash /etc/siq/ssl/certs
     ```
6. Copy the key pairs as `stunnel.pem` file to the appropriate location.

   Use the following commands:

   - On the gateway:

```
cp gateway.keypair.pem /etc/deepfile/gateway/stunnel.pem
```

- On each data server:

```
cp dataserver.keypair.pem /etc/deepfile/dataserver/stunnel.pem
```

- On the AppStack:

```
cp appstack.keypair.pem /etc/siq/ssl/stunnel.pem
```

7. Restart service as follows:
   a) On the gateway, run this command:

```
service deepfiler restart
```

   b) On each data server, run this command:

```
service deepfiler restart
```

   c) On the AppStack, run these commands:

```
/siq/bin/monit restart uwsgi
/siq/svc/stunnel/stunnel-ctl.sh stop
/siq/svc/stunnel/stunnel-ctl.sh start
```

8. Test the connectivity as follows:
   - Connectivity between the gateway and the data server

     On the gateway, issue the following API call:

```
curl http://localhost:7766/backchannel/1.0/server
```

   - Connectivity between the gateway and the AppStack

     On the AppStack, issue the following API call:

```
curl -i http://localhost:8765/administrative/1.0/server
```

9. To install certificates on the Elasticsearch nodes, complete the following steps on each node:
   a) Create a key pair, combining key and certificate:

```
cat node.key node.pem > node.keypair.pem
```

   b) Copy the key pair to the appropriate location.

```
cp  node.keypair.pem /siq/env/docker/node1-stunnel.pem
```

   c) Restart the Docker service by using the following commands:

```
cd /siq/env/docker
docker-compose restart
```

   d) To verify that the certificates were accepted, check the `/var/log/stunnel_service.log` log file.

   The communication between the nodes within the Elasticsearch cluster as well as the communication between the data server and the Elasticsearch nodes can also be encrypted by using other methods than stunnel. For more information about those other methods, contact IBM Support.

Configuring SSH key-based authentication

Generate and use an SSH key pair to use RSA-based private key authentication for passwordless SSH logins to any of the IBM StoredIQ nodes in your environment.

Complete this preparatory work:

- Make a list of the home directories of **all** users that you want to authenticate with SSH keys.
- If you want to add an extra layer of security, you can set a passphrase for the SSH key. Determine this passphrase beforehand.

In the following instructions, the sample user `myibmuser` is used. The respective home directory is `/home/myibmuser`.

Complete the procedure on each IBM StoredIQ server where you want to set up SSH key authentication.

1. Log in to the IBM StoredIQ server, for example, the IBM StoredIQ AppStack.
2. Create SSH key pairs for each user that you want to have SSH access to this server.

   An SSH key pair consists of a *private* and a *public* key. To generate the key pair, issue the following command:

   ```
   sudo -u myibmuser ssh-keygen -t rsa -b 4096 -C "myibmuser@example.local" -f /home/
   myibmuser/.ssh/id_rsa
   ```

   Messages similar to the following ones are written to the console:

   ```
   Generating public/private rsa key pair.
   Created directory '/home/myibmuser/.ssh'.
   Enter passphrase (empty for no passphrase):
   Enter same passphrase again:
   Your identification has been saved in /home/myibmuser/.ssh/id_rsa.
   Your public key has been saved in /home/myibmuser/.ssh/id_rsa.pub.
   The key fingerprint is:
   c3:6b:19:da:75:a0:4a:ca:e0:0e:ba:c3:79:2c:6c:1a myibmuser@example.local
   The key's randomart image is:
   +--[ RSA 4096]----+
   |                 |
   |                 |
   |      .          |
   |    . . .        |
   | . . S . .       |
   | . o o + * .     |
   |E + o o =        |
   |oX o .           |
   |*o+              |
   +-----------------+
   ```

   The first time that you create keys, the following files are created in the `/home/myibmuser/.ssh/` directory:

   The `id_rsa` file, which is the private key. That is the key file that you distribute to those users that will be allowed to log in to the IBM StoredIQ server via SSH.
   The private key allows access to the system. Therefore, keep this key file protected from unauthorized access.
   The `id_rsa.pub` file, which is the public key. This key file remains on the IBM StoredIQ server.

   Consider to save a copy of the key files because in each subsequent run of the **ssh-keygen** tool you are asked whether you want to overwrite the files. In this case, any previously generated key file is replaced so that the previous key can no longer be used for authentication.
3. Add the public key to the IBM StoredIQ server.

   The authorized keys for each account are stored in their respective home directories. With this configuration, anyone with the private key can connect to this IBM StoredIQ host as `myibmuser`:

   ```
   cat /home/myibmuser/.ssh/id_rsa.pub >> /home/myibmuser/.ssh/authorized_keys
   ```
4. Transfer the private key (the `id_rsa` file) to a host that needs SSH access to the IBM StoredIQ server by using a secure copy tool such as **scp**.
5. On the remote host, verify that key-based authentication works.

   For example, on OS X, you can use the -i flag to specify the key:

   ```
   ssh -i id_rsa myibmuser@storediqmachine.example.local
   ```

   As a result, the prompt should change to something like this one:

```
[myibmuser@storediq ~]
```

6. Disable password-based authentication in the SSH configuration.
   a) Display the default configuration for IBM StoredIQ by entering the following command:

   ```
   grep -E '^PasswordAuthentication' /etc/ssh/sshd_config
   ```

   This command should return `PasswordAuthentication yes`.
   b) Change the setting to no.
   Enter the following command:

   ```
   sed -i '/^PasswordAuthentication/s/yes/no/g' /etc/ssh/sshd_config
   ```

   c) To verify the new setting, display the configuration once again:

   ```
   grep -E '^PasswordAuthentication' /etc/ssh/sshd_config
   ```

   This command should now return `PasswordAuthentication no`.
   d) Restart the OpenSSH daemon:

   ```
   service sshd restart
   ```

   Now the IBM StoredIQ server does no longer not accept passwords for SSH logins.

Enabling encryption of IBM StoredIQ gateway and data server application data

IBM StoredIQ uses Linux Unified Key Setup (LUKS) for encrypting data at rest at the file system level. You can enable disk encryption on the gateway or a data server for security reasons.

Complete this procedure for new installations or for update installations where LUKS encryption was not yet enabled. After you complete the procedure, all application data on the gateway or the respective data server that is stored in the `/deepfs` directory structure is encrypted.

To enable disk encryption on the gateway and on a data server, complete the following steps:

1. Add two new disks to the VM by using your VMWare client.

   The IBM StoredIQ application data on the gateway or a data server is stored in the `/deepfs` directory structure. One of the disks is mounted at `/deepfs` and the other device is mounted at `/deepfs/nas`. The device mounted at `/deepfs/nas` contains all the Postgres database data.

   If you are encrypting the IBM StoredIQ application data on the gateway, both disks should have at least 75 GB of space. These disks will be used to encrypt the application data in the `/deepfs` and `/deepfs/nas` directories.

   If you are encrypting the IBM StoredIQ application data on a data server, one disk should have at least 124 GB and the other disk should have hat least 2 TB of space. The disk with at least 124 GB of space is used for `/deepfs` and the larger disk with 2 TB of space is used for `/deepfs/nas`.

   If your gateway or data server have disks whose current sizes (under `/deepfs` and `/deepfs/nas`) is larger than the stated default sizes, then use the current sizes when creating the new disks. You can also create the new disks with larger sizes if you feel the current sizes are not enough.
2. Restart the VM.
3. Log in to the VM as root and verify that the new disks were added by using the **fdisk** command.
   For example, if the two new disk were added as the devices `/dev/sdc` and `/dev/sdd`, run the following commands:

   - For the device `/dev/sdc`:

   ```
   fdisk -l /dev/sdc
   ```

The output should be similar to what is shown in the example:

```
Disk /dev/sdc: 79.1 GB, 79076261888 bytes
255 heads, 63 sectors/track, 9613 cylinders
Units = cylinders of 16065 * 512 = 8225280 bytes
Sector size (logical/physical): 512 bytes / 512 bytes
I/O size (minimum/optimal): 512 bytes / 512 bytes
Disk identifier: 0x00000000
```

- For the device /dev/sdd:

```
fdisk -l /dev/sdd
```

The output should be similar to what is shown in the example:

```
Disk /dev/sdb: 53.7 GB, 53687091200 bytes
255 heads, 63 sectors/track, 6527 cylinders
Units = cylinders of 16065 * 512 = 8225280 bytes
Sector size (logical/physical): 512 bytes / 512 bytes
I/O size (minimum/optimal): 512 bytes / 512 bytes
Disk identifier: 0x00000000
```

4. To transfer the IBM StoredIQ application data to your new disk devices and encrypt them, run the /usr/local/storediq/bin/encrypt.sh script.

   If you are encrypting data on a data server, specify the disk with at least 124 GB of space directly after the passphrase argument so that this disk is used to encrypt the data under /deepfs. Specify the disk with at least 2 TB of space as the last argument so that it is used to encrypt the data under /deepfs/nas.

   You must provide a passphrase argument that is later used to access the new devices after they are encrypted:

   ```
   /usr/local/storediq/bin/encrypt.sh mypassphrase deepfs_device deepfs_nas_device
   ```

   The output should be similar to what is shown in the example:

   ```
   # ./encrypt.sh mypassphrase /dev/sdc /dev/sdd
   INFO: /mnt/deepfs is already unmounted
   INFO: Successfully formatted LUKS device /dev/sdc.
   INFO: Successfully created LUKS mapped device /dev/mapper/deepfs
   INFO: Successfully created ext3 filesystem on device /dev/sdc.
   INFO: Successfully mounted /dev/mapper/deepfs to /mnt/deepfs
   INFO: /mnt/nas is already unmounted
   INFO: Successfully formatted LUKS device /dev/sdd.
   INFO: Successfully created LUKS mapped device /dev/mapper/nas
   INFO: Successfully created ext3 filesystem on device /dev/sdd.
   INFO: Successfully mounted /dev/mapper/nas to /mnt/nas
   INFO: Stopping deepfiler services.
   standalone system using auto-generated /etc/deepfile/monitrc
   Stopping storediq processes: monit daemon with pid [2144] killed
   killing findex loaders
   ******** Stop Sat Jul 13 02:47:20 UTC 2019 ********
   /etc/init.d/deepfiler: line 219:  4513 Killed                  /usr/local/bin/siq-loader stop
   killing findex writers

   INFO: Successfully stopped deepfiler services.
   INFO: Stopping postgres.
   Stopping postgresql service: [  OK  ]
   INFO: Successfully stopped postgres
   INFO: Copying content of /deepfs/nas/. to /mnt/nas/.
   INFO: Successfully copied /deepfs/nas/. to /mnt/nas/.
   INFO: Successfully unmounted /mnt/nas
   INFO: Successfully unmounted /deepfs/nas
   INFO: Copying content of /deepfs/. to /mnt/deepfs/.
   INFO: Successfully copied /deepfs/. to /mnt/deepfs/.
   INFO: Successfully unmounted /mnt/deepfs
   INFO: Successfully unmounted /tmp
   INFO: Successfully unmounted /deepfs
   INFO: Successfully mounted /dev/mapper/deepfs to /deepfs
   INFO: Successfully mounted /dev/mapper/nas to /deepfs/nas
   INFO: Successfully mounted /deepfs/sys/tmp to /tmp
   INFO: Saved passphrase to /usr/local/storediq/luks-keys/slot0-key
   INFO: Added LUKS key to /etc/crypttab file for device /dev/sdc and name deepfs.
   INFO: Added LUKS key to /etc/crypttab file for device /dev/sdd and name deepfs_nas.
   ```

```
INFO: Successfully made a backup of /etc/fstab to /etc/fstab.orig
grep: fstab.orig: No such file or directory
INFO: Added LUKS /dev/mapper/deepfs entry to /etc/fstab file.
INFO: Added LUKS /dev/mapper/deepfs_nas entry to /etc/fstab file.
INFO: Starting postgres.
Starting postgresql service:
INFO: Successfully started postgres
INFO: Starting deepfiler services.
standalone system using auto-generated /etc/deepfile/monitrc
Verifying/remounting FINDEX exports ..
getAllowedHosts: ['198.51.100.18', '0.0.0.0', 'siq-gw6']
Verifying NAS symlinks
Stopping reactor
Verifying database ..
Starting storediq processes..
Removing leftovers from /deepfs/data/tmp..
mkdir: cannot create directory `/deepfs/config/apache': File exists
Stopping siqsnmpagent process
Starting siqsnmpagent process

Xvfb does not appear to be running
rpc.mountd appears to be running, pid: 3942
starting Web Server
AH00558: httpd: Could not reliably determine the server's fully qualified domain name, using
0.0.0.0. Set the 'ServerName' directive globally to suppress this message
/etc/deepfile/gateway/apache_gateway_ctl start: httpd started
starting Apache Tomcat Servlet Engine
starting NFS Mount Daemon
Starting/checking rpc.mountd
rpc.mountd already running, pid: 3942
starting SIQ Log Service
starting Findex Writer 1
starting Findex Writer 2
starting DA siqtransport
starting Query Replication Service
starting Query Status Service
starting Distributed Search Service
starting Distributed Object Viewer Service
starting Distributed Job Execution Service
starting Distributed Cube Replication Service
starting Distributed Policy Audit Service
starting Gateway API service
getAllowedHosts: ['198.51.100.18', '0.0.0.0', 'siq-gw6']
starting Gateway backchannel service
getAllowedHosts: ['198.51.100.18', '0.0.0.0', 'siq-gw6']
starting Gateway ContentType Controller Service
starting Gateway FileType Controller Service
starting Gateway AgeGroup Controller Service
starting Gateway SizeGroup Controller Service
starting Gateway DataExplorer Controller Service
starting Gateway Volume Controller Service
starting Gateway VolumeType Controller Service
starting Distributed Administration Service
starting Gateway stunnel service
stunnel not enabled for this gateway
starting Gateway service for monitoring statistics
starting Gateway service for subscriber registration
starting Autoclassification Replication Service
starting HashSet Replication Service
starting Cartridge Replication Service
starting Report Service
turning on monitoring for StoredIQ processes
INFO: Successfully started deepfiler services.
```

For more information about the usage of the `encrypt.sh` script, use the -h command option.

The application data on the IBM StoredIQ gateway or data sever is now encrypted.

At any time, you can change the passphrases for the encrypted disks. To change the passphrases for key slot 0, run the following commands replacing *deepfs_device* and *deepfs_nas_device* with the appropriate values:

```
cryptsetup luksDump deepfs_device | deepfs_nas_device
cryptsetup luksAddKey deepfs_device | deepfs_nas_device
```

Enter the passphrases when prompted.

Enabling encryption of IBM StoredIQ AppStack application data

IBM StoredIQ uses Linux Unified Key Setup (LUKS) for encrypting data at rest at the file system level. You can enable disk encryption on the AppStack for security reasons.

Complete this procedure for new installations or for update installations where LUKS encryption was not yet enabled. After you complete the procedure, all application data on the AppStack that is stored in the `/var/siq` directory structure is encrypted.

To enable disk encryption:

1. Add a new disk with at least 25 GB of space to the AppStack VM by using your VMWare client.
2. Restart the AppStack VM.
3. Log in to the AppStack as root and verify that the new disk was added by using the **fdisk** command.

   For example, if the new disk was added as the device `/dev/sdc`, run the following command:

   ```
   fdisk -l /dev/sdc
   ```

   The output should be similar to what is shown in the example:

   ```
   Disk /dev/sdc: 25.2 GB, 17179869184 bytes
   255 heads, 63 sectors/track, 2088 cylinders
   Units = cylinders of 16065 * 512 = 8225280 bytes
   Sector size (logical/physical): 512 bytes / 512 bytes
   I/O size (minimum/optimal): 512 bytes / 512 bytes
   Disk identifier: 0x08040000
   ```

4. To transfer the IBM StoredIQ application data to your new disk device and encrypt it, run the `/siq/bin/encrypt.sh` script.

   You must provide a passphrase argument that is later used to access the new device after it is encrypted:

   ```
   /siq/bin/encrypt.sh mypassphrase /dev/sdc
   ```

   The output should be similar to what is shown in the example:

   ```
   INFO: Successfully formatted LUKS device /dev/sdc.
   INFO: Successfully created LUKS mapped device /dev/mapper/siq.
   INFO: Successfully created ext3 filesystem on device /dev/sdc.
   INFO: Successfully mounted /dev/mapper/siq to /mnt/siq.
   INFO: Stopping appstack services.
   Monit daemon with pid [25318] killed
   Monit did not go away. Doing a hard kill on pid 25318
   Services stopped
   INFO: Successfully stopped appstack services.
   INFO: Successfully copied /var/siq/. to /mnt/siq/.
   INFO: Successfully unmounted /dev/mapper/siq from /mnt/siq.
   INFO: Successfully renamed /var/siq to /var/siq.orig.1537897246
   INFO: Successfully mounted /dev/mapper/siq to /var/siq.
   INFO: Starting appstack services.
   Starting Monit 5.12.2 daemon with http interface at [localhost]:2812
   Services and processes started
   INFO: Successfully started appstack services.
   ```

The application data on the IBM StoredIQ application stack is now encrypted.

At any time, you can change the passphrase for the encrypted disk. To change the passphrase for key slot 0, run the following commands:

```
cryptsetup luksDump /dev/sdc
cryptsetup luksAddKey /dev/sdc
```

Enter the passphrases when prompted.

Enabling or disabling FIPS

At any time, you can enable or disable FIPS as required on the IBM StoredIQ components.

On the gateway or any data server, you must be logged in as root to check or modify the FIPS enablement status. On the AppStack, you must be logged in as `siqadmin` user.

- To check the status of FIPS on the gateway, any data server, or the AppStack, check the `/proc/sys/crypto/fips_enabled` file on each component.

  On the gateway or any data server, use edit the file directly. On the AppStack, run the **edit_proc_sys_crypto_fips_enabled** command. The entry 0 indicates that FIPS is disabled and 1 means it is enabled.
- To enable FIPS:
  - On the gateway or any data server, run this command: `python /usr/local/storediq/bin/fipscfg.pyc --enable`
  - On the AppStack, run the `appstackcfg` utility and select **Enable FIPS mode at boot time**.
- To disable FIPS:
  - On the gateway or any data server, run this command: `python /usr/local/storediq/bin/fipscfg.pyc --disable`
  - On the AppStack, run the `appstackcfg` utility and deselect **Enable FIPS mode at boot time**.

Securing Elasticsearch cluster communication with Search Guard

To better secure the IBM StoredIQ Elasticsearch cluster, you can enable Search Guard on each node in the cluster.

If you want to use Search Guard to secure the Elasticsearch cluster, the primary Elasticsearch node must be configured with specific property settings in the `cluster-setup.properties` file. To edit the file, you must be logged in with the `builder` account. Update the properties file as follows:

- The `STUNNEL_ENABLED` property must be set to `false`.
- For deployments upgraded to IBM StoredIQ 7.6.0.17, add the following entry to the end of the properties file:

  ```
  SEARCHGUARD_ENABLED=true
  ```

  For new deployments (starting with IBM StoredIQ 7.6.0.17), the `SEARCHGUARD_ENABLED` property must be set to `true`.

After editing the properties file, run the cluster setup with the updated properties file. At the prompt `[builder@localhost ~]$`, enter:

```
/siq/bin/cluster-setup.sh cluster-setup.properties
```

Search Guard enables HTTPS connections between the data server and the Elasticsearch cluster and between the nodes of the Elasticsearch cluster. It also ensures that all connections to the Elasticsearch other than those with authorized clients are blocked.

1. Initialize the Search Guard configuration.
   a) Log in to the primary Elasticsearch node as `builder`. Then, switch to the root user:

   ```
   su root
   ```

   b) Log in to the Docker shell by using this command:

   ```
   docker exec -it node-name bash
   ```

where *node-name* is the name of the primary Elasticsearch node.

c) Navigate to the Search Guard configuration directory:

```
cd /usr/share/elasticsearch/config
```

d) Run the Search Guard **sgadmin** tool with the following options:

```
/usr/share/elasticsearch/plugins/search-guard-5/tools/sgadmin.sh -cd /usr/share/
elasticsearch/plugins/search-guard-5/sgconfig -cn cluster_name -cert sgcerts/es-
admin.crt.pem -key sgcerts/es-admin.key.pem -cacert sgcerts/chain-ca.pem -nhnv
```

where *cluster_name* is the name of the Elasticsearch cluster.

2. Test the setup by using this command:

```
curl -vk https://localhost:9200/_cluster/health?pretty -E ./sgcerts/es-admin.crt.pem --key ./
sgcerts/es-admin.key.pem --cacert ./sgcerts/chain-ca.pem
```

3. To enable secure communication between the data server and the Elasticsearch nodes, copy the data server certificates that are available on the Elasticsearch nodes to each data server.

a) Log in to an Elasticsearch node as `root`.

b) Navigate to the directory where the certificates are stored:

```
cd /siq/env/docker/sgcerts
```

c) Copy the following files to the `/etc/deepfile/dataserver` directory on the data server.

Use a secure copy tool such as scp to do so.

- `es-admin.crt.pem`
- `es-admin.key.pem`
- `chain-ca.pem`

d) Log in to the data server as `root` and rename the certificates as follows:

`es-admin.crt.pem` to `client.crt.pem`
`es-admin.key.pem` to `client.key.pem`
`chain-ca.pem` to `cert-chain.pem`

e) Restart all services on the data server.

4. Test the connection from the data server to the Elasticsearch cluster:

```
curl -vk https://primary_node_ip:9200/_cluster/health?pretty -E ./client.crt.pem --key ./
client.key.pem --cacert ./cert-chain.pem
```

where *primary_node_ip* is the IP address of the primary Elasticsearch node.

Restricting access to port 9200 on Elasticsearch nodes

By default, port 9200 is open on the nodes in the Elasticsearch cluster. You can restrict client access to this port by changing the firewall setup.

Create a **firewalld** service unit for adding or removing the firewall rules that allow the Elasticsearch nodes and the data server to use port 9200, and reject all communication coming from other sources. Then, enable the firewall to activate the rules. To lift the restriction, disable the firewall.

Important: You must complete this setup on each node in the Elasticsearch cluster.

1. Log in to an Elasticsearch node as `builder`. Then, switch to the root user:

```
su root
```

Alternatively, you can run the commands from the `builder` account via **sudo**.

2. Create the **firewalld** service unit.

The rules in the service unit add the IP addresses that are allowed to access port 9200 and block all other traffic. At a minimum, the IP addresses of the other nodes in the Elasticsearch cluster and the data server to which the Elasticsearch cluster is attached must have access to port 9200.

At the command prompt, enter the following information. Replace *IP_addresses* with either a Classless Inter-Domain Routing (CDIR) network address such as `192.0.2.0/24` or a comma-separated list of IP addresses such as `192.0.2.10,192.0.2.12,192.0.2.15,192.0.2.20`.

```
cat > /etc/systemd/system/esfw.service
[Unit]
Description=Dynamic firewall rules for elasticsearch service
Requires=elasticsearch.service
After=elasticsearch.service

[Service]
Type=oneshot
WorkingDirectory=/siq/env/docker
ExecStart=/usr/bin/firewall-cmd --zone=public --direct --add-rule ipv4 filter DOCKER-USER 0 -
p tcp -i eth0 -s IP_addresses --dport 9200 -j ACCEPT
ExecStart=/usr/bin/firewall-cmd --zone=public --direct --add-rule ipv4 filter DOCKER-USER 1 -
p tcp -i eth0 --dport 9200 -j DROP
ExecStop=/usr/bin/firewall-cmd --zone=public --direct --remove-rule ipv4 filter DOCKER-USER
1 -p tcp -i eth0 --dport 9200 -j DROP
ExecStop=/usr/bin/firewall-cmd --zone=public --direct --remove-rule ipv4 filter DOCKER-USER
0 -p tcp -i eth0 -s IP_addresses --dport 9200 -j ACCEPT
RemainAfterExit=yes
TimeoutStartSec=0
User=root

[Install]
WantedBy=multi-user.target
```

3. To exit the input mode and save the file, press Crtl+D.
4. To pick up the `esfw.service` unit, reload `systemd` and enable the service unit during startup.

   Enter the following commands:

   ```
   systemctl daemon-reload
   systemctl enable esfw
   ```

5. Enable the firewall by entering the following command:

   ```
   systemctl enable firewalld
   ```

6. Restart Elasticsearch by using this command:

   ```
   systemctl restart elasticsearch
   ```

7. Check the status of the Elasticsearch cluster by submitting the following request from one of the allowed hosts:

   ```
   curl -XGET 'http://primary_node_ip:9200/_cluster/health?pretty'
   ```

   The response should look similar to this example:

   ```
   {
    "cluster_name" : "es-cluster",
    "status" : "green",
    "timed_out" : false,
    "number_of_nodes" : 3,
    "number_of_data_nodes" : 3,
    "active_primary_shards" : 241,
    "active_shards" : 242,
    "relocating_shards" : 0,
    "initializing_shards" : 0,
    "unassigned_shards" : 0,
    "delayed_unassigned_shards" : 0,
    "number_of_pending_tasks" : 0,
    "number_of_in_flight_fetch" : 0,
    "task_max_waiting_in_queue_millis" : 0,
    "active_shards_percent_as_number" : 100.0
   }
   ```

The firewall is enabled so that client access other than from the other Elasticsearc nodes or the data server is blocked. If you want to lift the restriction, you can disable this configuration at any time. To do so, complete these steps on each Elasticsearch node as root user. Alternatively, you can run the commands from the `builder` account via **sudo**.

1. Disable the firewall by using this command:

   ```
   systemctl disable firewalld
   ```

2. Disable the `esfw.service` unit by using the following command:

   ```
   systemctl disable esfw
   ```

3. Restart Elasticsearch by using this command:

   ```
   systemctl restart elasticsearch
   ```

Managing the status of secure gateway communication

The gateway handles the communication between the data servers and the application stack. You can check whether secure gateway communication is used to encrypt the communication and change the enablement status as required.

You must be logged in to the component that you want to manage: on the gateway or any data server as `root` user, on the application stack as `siqadmin` user.

Secure gateway communication must be enabled or disabled on all components. Follow these steps to see the enablement status of these components, and to disable or enable secure gateway communication as required.

Secure communication via stunnel can impact performance. Therefore, enable this setting only if your enterprise security policy mandates encryption of data in motion. If you do so, the IBM StoredIQ gateway and the data server must be configured accordingly. For more information, see “Security” on page 16. For additional guidance, contact IBM Support.

You can change the enablement status at any time after the installation.

- To check the status:
  - On the application stack, enter `service appstack status` on the command prompt.

    If **Process appstackStunnel** is listed and running, secure gateway communication is enabled on the application stack. Otherwise, it is disabled.
  - On the gateway, enter `service deepfiler status` on the command prompt.

    If **Process GatewayStunnel** is in the **running** state, secure gateway communication is enabled. If it is in the **not monitored** state, it is disabled.
  - On a data server, enter `service deepfiler status` on the command prompt.

    If **Process DataserverStunnel** is listed and running, secure gateway communication is enabled. Otherwise, it is disabled.
- To enable secure gateway communication:
  a) On the gateway, enter `/usr/local/storediq/bin/util/stunnelcfg` on the command prompt. Then, enter `service deepfiler restart` to restart the services.
  b) On a data server, enter `/usr/local/storediq/bin/util/stunnelcfg` *`gateway_ip`* on the command prompt.

     Then, you have to log in to the IBM StoredIQ Data Server user interface to complete this configuration:

     a. Go to **Administration** > **Configuration** > **DA Gateway settings**.
     b. Enter `127.0.0.1` in the **Host** field.

c. Enter the name in the **Node name** field.

d. Click **OK**.

e. Log off.

Back on the data server, enter `service deepfiler restart` on the command line to restart the services.

c) On the application stack, complete these steps:

a. Run the `appstackcfg` utility.

b. Select **Enable secure gateway communication**.

c. Select **Restart appstack services** to restart the services on exiting the configuration utility and then select **Save and exit**.

Alternatively, you can just select **Save and exit** and then restart the services from the command line by using this command: `service appstack restart`

- To disable secure gateway communication:

a) On the gateway and on the data server, enter `/usr/local/storediq/bin/util/stunnelcfg disable` on the command prompt. Then, enter `service deepfiler restart` to restart the services.

b) On the application stack, complete these steps:

a. Run the `appstackcfg` utility.

b. Deselect **Enable secure gateway communication**.

c. Select **Restart appstack services** to restart the services on exiting the configuration utility and then select **Save and exit**.

Alternatively, you can just select **Save and exit** and then restart the services from the command line by using this command: `service appstack restart`

Securing the data server against host header injection vulnerabilities

To prevent host header injection attacks, the data server allows sessions to be established only with clients (usually by using the data server admin interface) that use specific host names or IP addresses in the URLs used to connect to the data server. These allowed hosts can be determined automatically by the system and provided in a configuration file.

By default, the data server attempts to determine both the primary host name and the primary IP address. If the default configuration is not modified, these are the only references to the data server that work in the URLs that a client uses to establish a session with the data server. Check the `AppServer.out` file in the `/deepfs/config` directory to see which hosts are allowed. The contents of the file might look like this example:

```
*** RESTART ***
Wed Jun 19 21:15:57 UTC 2019
getAllowedHosts: ['198.51.100.0', 'storediq-ds1']
```

The data server lists the allowed hosts, in this case 198.51.100.0 and storediq-ds1, that can be used by clients. In general, two default hosts are automatically determined by the data server on startup: the primary IP address (198.51.100.0) and the primary host name (storediq-ds1). In addition, one or more hosts that represent the local data server system might also be automatically determined, for example, the `localhost` IP address of 127.0.0.1.

Often the default hosts are not sufficient, and host names need to be added to the system to allow clients to use those host names in their URLs. To do so, modify the `settings.py` file in the `/usr/lib/python2.6/site-packages/deepfile/ui/djangoweb` with a list of allowed hosts to add.

1. Using an SSH tool, log in to the data server as root or as a user with sudo access.

2. Go to the /usr/lib/python2.6/site-packages/deepfile/ui/djangoweb directory.
3. Back up the settings.py file located in this directory.
4. Edit the settings.py file.
   a) Locate the line that starts with ALLOWED HOSTS.
      This line is usually within the first 15 lines:

      ```
      ALLOWED_HOSTS = getAllowedHosts()
      ```

   b) To provide an extra allowed host, insert it into the ALLOWED_HOSTS line as follows:

      ```
      ALLOWED_HOSTS = getAllowedHosts('dataserver')
      ```

      Thus, the dataserver host is added to the allowed hosts when the AppServer service is restarted.
   c) To provide more than one allowed host, add them in the same way, separating each with a comma:

      ```
      ALLOWED_HOSTS = getAllowedHosts('dataserver', 'myds')
      ```

5. Save the settings.py file.
6. Restart the AppServer service to pick up the new configuration by running the following command:

   ```
   monit restart AppServer -c /etc/deepfile/monitrc
   ```

   The AppServer restart takes a little while. You can monitor the progress by using the following command:

   ```
   monit summary AppServer -c /etc/deepfile/monitrc
   ```

   The AppServer service must be restarted for any changes in settings.py to take effect.

After a restart with the modified ALLOWED_HOSTS line (as in the example), the output in AppServer.out might look like this:

```
*** RESTART ***
Wed Jun 19 21:25:37 UTC 2019
getAllowedHosts: getAllowedHosts: ['198.51.100.0', 'dataserver', 'myds', 'storediq-ds1']
```

This security approach means that URLs employed in browsers to access the data server user interface must use one of the allowed IP addresses or host names listed in the AppServer.out file when the AppServer service initializes. If a client attempts to establish a session using a host name or IP address that was not automatically determined or specified in the ALLOW_HOSTS line of settings.py, the session request is rejected.

A user tries to access the data server admin user interface used in these examples with this URL: https://storediq-dataserver/login

Provided the host name storediq-dataserver is resolved to the data server system (either through DNS resolution or an entry in the client system's hosts file), the user is presented with the login page as usual. However, the login attempt results in the following error message:

```
Unhandled Exception
An unhandled exception was thrown by the application.
```

This is because storediq-dataserver is not an allowed host. If they instead use the URL https://storediq-ds1/login, the login is processed in the normal way and access should be granted to the data server admin user interface provided suitable credentials are supplied.

Updating initial configuration settings

Update configuration settings by completing configuration steps that you skipped during installation or by changing existing settings.

You can add or change the following settings:

Gateway configuration
The AppStack IP address

Data server configuration
DA gateway settings

AppStack configuration
Email notification settings

Settings for the synchronization with a governance catalog

HTTPS enablement for the AppStack by using the **certcfg** utility

Updating the gateway configuration

If you did not set the AppStack IP address during the initial configuration of the gateway, you must update the configuration with this information after the installation. Otherwise, the gateway cannot communicate with the AppStack.

You can also use this procedure to update the AppStack IP address in the gateway configuration if required.

1. Using an SSH tool, log into the gateway server as `root`.
2. Run the **/usr/bin/set-appstack-ip.sh** script and enter the AppStack IP address when prompted.

   This script adds an entry into the `pg_hba.conf` file, allowing trusted access to the gateway database from the AppStack IP address. Without this access, reports cannot run from the AppStack. This script must be run whenever the AppStack IP address happens to change.

Configuring the data server gateway settings

Check and, if required, modify the gateway settings on the IBM StoredIQ data server.

You must complete this task for every data server. If you change any settings, you must restart services.

1. Log in to the IBM StoredIQ Data Server user interface.

   In a web browser, enter the IP address that you obtained during the installation of the data server. Alternatively, you can specify the host name.
2. In the login window, enter the credentials of the default administrative account for IBM StoredIQ Data Server: user ID `admin`, password `admin`
3. Go to **Administration** > **Configuration** > **DA Gateway settings**.
4. 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`.
5. The **Node name** field shows the name of the data server that you assigned during installation.

   Change as required.
6. If you changed any of the settings, restart services in either of the following ways:
   - Go to the data server dashboard, click **About Appliance** and then click **Restart Services**.
   - Using an SSH tool, log in to the data server VM as root and then run this command: `service deepfiler restart`.

Configuring the application stack to send and receive reports and notifications

At any time after deployment, you can set up or change the application stack's configuration for sending and receiving reports and notifications through email.

If the email server was not configured during deployment or if you want to change any settings later, complete these steps to configure the application stack accordingly.

1. Using an SSH tool, log into the application stack as `siqadmin`.
2. Launch the Appstack Configuration utility using this command: `appstackcfg`
3. In the **Domain name** field, enter the fully qualified domain name (FQDN) or IP address of the application stack.
4. In the **StoredIQ Gateway** field, enter the IP address for the IBM StoredIQ gateway server.
5. Set the following SMTP options:

Parameter	Value
Server	The mail server's fully qualified domain name or IP address.
Port	The SMTP port. The default port is 25.
Username	The login user name. For the default configuration, leave this field empty. Otherwise, provide the user name of the user with which to authenticate to the Exchange server. • If you authorized any Authenticated User, you can use any user name and password as long as that individual is valid member of the domain. • If you used a specific user, you must use the user name of the single user for which you granted permissions. This must be a fully qualified user name. In this case, you must have completed the instructions in "Configuring authenticated users for SMTP notifications" on page 38 before configuring the SMTP settings.
Password	The login password for the specified user. For the default configuration, leave this field empty.

6. Select **Enable TLS** to enable TLS encryption, if it is supported by the mail server.
7. Select **Restart appstack services**.

   Alternatively, you can restart the application stack services from the command line after you save the configuration by running the following command: **monit restart uwsgi**
8. Select **Save and exit** and wait for all IBM StoredIQ services to restart.

   You can now send email notifications from the IBM StoredIQ application stack or when generating reports.

Configuring the application stack to synchronize data with the governance catalog

At any time after deployment, you can set up or change the application stack's configuration for making data from IBM StoredIQ data sources discoverable in IBM Information Server or IBM Cloud Private for Data, to make governance catalog data classes available for use in IBM StoredIQ, and to keep the information in sync.

Before you can configure and use the synchronization feature, a working deployment of one of these products must be available:

- IBM Information Server. The minimum required version is Version 11.7.0.
- IBM Cloud Private for Data. The minimum required version for full support is Version 1.2.

If the synchronization was not enabled during deployment or if you want to change any settings later, complete these steps to configure the application stack accordingly.

Important: Synchronization requires HTTPS to be enabled on the AppStack. Therefore, it is mandatory that at least a self-signed certificate is installed.

To enable the synchronization with Information Governance Catalog or the IBM Cloud Private for Data catalog at any time after deployment of the AppStack:

1. Using an SSH tool, log in to the application stack as `siqadmin`.
2. Launch the Appstack Configuration utility by using this command: `appstackcfg`
3. Check the value in the **Domain name** field.

   The information that you provide here is used to build the base URLs for REST access to the application stack and for links to IBM StoredIQ artifacts that make these artifacts accessible from the governance catalog. Therefore, you should provide the fully qualified domain name or the IP address of the application stack.
4. If you want your data experts to receive email notifications for any changes to the synchronized governance catalog objects, ensure that the application stack configuration includes the appropriate SMTP settings.

   For details, see “Configuring the application stack to send and receive reports and notifications” on page 58.
5. Select **Enable synchronization with the governance catalog** and provide the following settings.

   If the data catalog to which you want to publish the IBM StoredIQ object resides in an IBM Cloud Private for Data environment, select the **Server runs in IBM Cloud Private for Data** check box.

   Provide or accept the values for these fields:

Parameter	Value
Host	The host name or IP address of the Information Server or IBM Cloud Private for Data installation. The specified host is part of the base URLs for REST access to the Information Governance Catalog or IBM Cloud Private for Data instance and for links to catalog artifacts that make these artifacts accessible from IBM StoredIQ. Therefore, you should provide the fully qualified domain name of the Information Server or IBM Cloud Private for Data host. If you specify a server port, this port also becomes part of such base URLs. In addition, the host name is also used to address the Information Server Apache Kafka server, which provides all Information Server events as Kafka messages. Specific Kafka messages are consumed by IBM StoredIQ and trigger the synchronization of objects from the governance catalog to IBM StoredIQ.
Port	The port of the governance catalog server. This setting is optional. For connections to an Information Server 11.7 environment: To ensure proper communication,

Parameter	Value
	you should set the port to the HTTPS port that is defined in Information Server. The default port is 9443. For connections to an Information Server 11.7 FP1 (or later) environment without Information Server Enterprise Search installed: To ensure proper communication, you should set the port to the HTTPS port that is defined in Information Server. The default port is 9446.. For connections to an Information Server 11.7 FP1 (or later) environment with Information Server Enterprise Search installed: Do not specify a port. For connections to IBM Cloud Private for Data, you can specify the IBM Cloud Private for Data port.
Kafka port	The port of the Information Server Kafka server. The port setting can be overridden. For connections to an Information Server 11.7 environment: the Kafka port defined in Information Server. The default port is 59092. For connections to an Information Server 11.7 FP1 (or later) environment without Information Server Enterprise Search installed: the Kafka port defined in Information Server. The default port is 59092. For connections to an Information Server 11.7 FP1 (or later) environment with Information Server Enterprise Search installed: the Kafka port defined in Information Server. The default port is 9092. For connections to IBM Cloud Private for Data: the Kafka port defined in IBM Cloud Private for Data. For more information, see the topic Enabling synchronization with IBM StoredIQ in the IBM Cloud Private for Data product documentation.
Username	The user name for authenticating to Information Server or IBM Cloud Private for Data when publishing IBM StoredIQ objects to the governance catalog. This user must be defined in Information Server with the following security roles: Suite User Common Metadata Administrator Information Governance Catalog Information Asset Administrator

Parameter	Value
	In IBM Cloud Private for Data, this user must be defined with the Data Stewart role.
Password	The password of the user set with **Username**.
Sync frequency (minutes)	Data is periodically propagated to the governance catalog at the specified interval. The value must be a positive number of minutes. The default value is 15 minutes.
StoredIQ instance name	The name identifying the IBM StoredIQ instance for which data is synchronized. This name can be freely chosen, but must be unique within the governance catalog instance.

6. Select **Restart appstack services**.

   Restarting the services is required for any configuration changes to take effect. Selecting this option is an alternative to restarting the application-stack services `uwsgi` and `tomcat` from the command line.

7. Select **Save and exit** and wait for all IBM StoredIQ services to restart.

   On the first synchronization run, the governance catalog is initially populated with the IBM StoredIQ artifacts that you selected for publishing to the catalog.

Backing up the IBM StoredIQ image

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

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. This type of backup is supported only for data servers.

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

1. Stop services on all data servers and the gateway:
   a) Log in to each data server and to the gateway as `root`.
   b) To stop all IBM StoredIQ services, enter the following command:

      ```
      service deepfiler stop
      ```

   c) To stop the `postgresql` database service, enter the following command:

      ```
      service postgresql stop
      ```

   d) Log out.

   Important: Wait 10 minutes after a harvest before you use this command to stop services.
2. Stop the IBM StoredIQ services on the application stack:
   a) Log in to the application stack as `siqadmin` user.
      Alternatively, you can log in as `root` user.
   b) Enter the following command:

      ```
      service appstack stop
      ```

   c) Log out.
3. 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.
4. Restart services on all data servers and the gateway:
   a) Log in to each data server and to the gateway as `root`.
   b) To restart all IBM StoredIQ services, enter the following command:

      ```
      service deepfiler restart
      ```

   c) To restart the `postgresql` database service, enter the following command:

      ```
      service postgresql restart
      ```

5. Start the IBM StoredIQ services on the application stack:
   a) Log in to the application stack as `siqadmin` user.
      Alternatively, you can log in as `root` user.
   b) Enter the following command:

      ```
      service appstack start
      ```

Upgrading IBM StoredIQ

You can obtain the upgrade ISO from IBM Fix Central. Upgrades consist of upgrading the gateway and data server first and then the application stack. These instructions apply to upgrades from the preceding release version to the current version. Direct upgrades from earlier versions are not supported.

Important:

- Back up all of your IBM StoredIQ VMware images (gateway, data servers, application stack, and all nodes of the Elasticsearch cluster, if deployed) before your start the upgrade in case the upgrade fails. You can use one of the following methods:
  - Create a snapshot of the current state of the image.
  - Create a clone of the current state of the image.
- In deployments with a data server of the type DataServer - Distributed, you must upgrade the Elasticsearch cluster before you upgrade any data server. For more information, see "Upgrading the Elasticsearch cluster" on page 67.

Go to IBM Fix Central. In the **Product selector** field, enter `StoredIQ` and then select the latest fix pack as **Installed Version**. The upgrade ISO is part of the fix pack.

Considerations when not upgrading from the preceding version

IBM StoredIQ update packages are not cumulative. Therefore, you must upgrade to the preceding release version before you can upgrade to the current version.

Direct upgrades are supported from IBM StoredIQ 7.5.1.0 to 7.6.0.1. The upgrade path for any release version starting with version 7.6.0.1 requires you to install all 7.6.0.*x* upgrade packages one by one in ascending order.

For new deployments, directly install the most current fix pack without installing the base version 7.6.0. Each fix pack does not only contain an upgrade package but also an OVA for a fresh installation.

Some of the upgrades from earlier versions require additional steps:

- "Upgrading from 7.6.0.2 to 7.6.0.4" on page 63
- "Upgrading the gateway and data server from 7.6.0.3 to 7.6.0.4" on page 64
- "Upgrading to 7.6.0.15" on page 64
- "Upgrading to 7.6.0.16" on page 64
- "Upgrading to 7.6.0.17" on page 65
- "Upgrading to 7.6.0.18" on page 66

Upgrading from 7.6.0.2 to 7.6.0.4

Gateway

When updating the gateway from 7.6.0.2 to 7.6.0.4, run this command: `sed -i '/^icc/d' /etc/deepfile/package_excludes` and then proceed with the normal upgrade process.

Application stack

For application stack upgrades from 7.6.0.2 to 7.6.0.4, complete the following steps before the actual upgrade:

1. Update the gateway and open port 80.
2. Log in as root using PuTTY and run the following commands:

a. `wget http://<gatewayIP>/products/appstack/centos65/7.6.0.4+32/siq-appstack-1.1.15+81-1.x86_64.rpm`
b. `rpm -Uvh siq-appstack-1.1.15+81-1.x86_64.rpm`

3. Proceed with the normal application stack upgrade process.

Upgrading the gateway and data server from 7.6.0.3 to 7.6.0.4

For DVD and ISO upgrades from 7.6.0.3 to 7.6.0.4, the following steps must be done before the upgrade of the gateway and data servers.

1. Copy the 7.6.0.4 upgrade ISO into the `tmp` directory on the gateway and data servers.
2. Log in as root using PuTTY.
3. Create mount point by running this command:

```
mkdir /mnt/cdrom
```

4. On the data server and gateway that are being updated, mount the update ISO:

```
mount -o loop /tmp/7_6_0_4-STOREDIQ-38-upgrade.iso /mnt/cdrom
```

5. Upgrade the updater package:

```
rpm -Uvh /mnt/cdrom/storediq-appliance-updater*.rpm
```

6. Unmount the ISO:

```
umount /mnt/cdrom
```

7. When upgrading the gateway, run this command:

```
sed -i '/^icc/d' /etc/deepfile/package_excludes
```

Do **not** run this command on the data servers

8. Proceed with the gateway and data server normal upgrade process.

Upgrading to 7.6.0.15

The 7.6.0.14-IBMStoredIQ_IF001 was uploaded to Fix Central in January, 2018. Before you upgrade to IBM StoredIQ Version 7.6.0.15, deploy the Spectre (CVE-2017-5715, CVE-2017-5753) and Meltdown (CVE-2017-5754) fix pack to the 7.6.0.14 system and then upgrade.

Also check the information about enabling disk encryption on an Elasticsearch cluster.

Upgrading to 7.6.0.16

The following migration steps are required:

Change existing Linux account passwords

Starting with IBM StoredIQ Version 7.6.0.16, hashing algorithm for Linux account passwords was changed to SHA512 with 999,999 rounds, which is now the default for new installations. For upgraded deployments, the passwords for all Linux user accounts that can log in to the system remotely or that can execute tasks with elevated privileges must be updated manually on each OVA after the upgrade. Do this to have upgraded deployments match the security standards for new installations where the passwords are hashed using SHA512 with 999,999 rounds by default.

For each user ID, change the current password and specify the hashing algorithm that you want to use as shown in the examples:

```
/usr/sbin/authconfig --passalgo=sha512 --update
```

This command changes the password for the given user ID using SHA512 as the hashing algorithm.

```
echo userID:mynewpassword | chpasswd --crypt-method SHA512 --sha-rounds 999999
```

This command changes the password for the given user ID using SHA512 as the hashing algorithm with 999,999 rounds of hashing.

Set the password for the synchronization anew

If synchronization with the governance catalog is enabled, you must update the configuration of the application stack after the product upgrade is complete by setting the password for the governance catalog user once again. If you do not update the password, synchronization with the governance catalog will no longer be possible. If the synchronization is not enabled, you can skip this update.

1. Launch the Appstack Configuration utility by using the **/siq/bin/appstackcfg** command.
2. In the section where the settings for the synchronization are configured, navigate to the **Password** field and specify the password for the user whose credentials are used for authenticating to Information Server.
3. Select **Restart appstack services**.
4. Select **Save and exit** and wait for all IBM StoredIQ services to restart.

Upgrading to 7.6.0.17

Important: Upgrading the AppStack to version 7.6.0.17 entails an upgrade of PostgreSQL. Therefore, it is strongly recommended that you back up your AppStack data before starting the upgrade. Otherwise, you will not be able to recover your data if the upgrade fails. For more information, see the topic about backing up the IBM StoredIQ image in the administration guide.

1. After you issue the `bootstrap product.upgrade appstack HEAD` command, monitor the upgrade process by tailing the `appstack.log` file:

   ```
   tail -f /var/bootstrap/appstack.log
   ```

   The process stops if the available space on the Postgres data volume is not sufficient for the upgrade. In this case, fix the issue as indicated and restart the upgrade.

   The process stops in any case at some point to give you the opportunity to back up your AppStack data before upgrading in case you haven't already done so.

   Tip: Copy the command shown in the message.

   Exit the log file. Back up your data if required. To proceed with the upgrade, paste the command that you copied earlier to the command line or enter the following command manually:

   ```
   nohup /etc/bootstrap/upgrade-appstack.sh /var/bootstrap/products/appstack/repository HEAD
   production Continue &>/dev/null &
   ```

   Tail the log file again to monitor the remaining upgrade process.
2. Examine the `/var/siq/log/tomcat.log` log file for any errors written by the audit persistence service.

   During the upgrade, the schema of the audit database is updated. Errors that occur during this update do not prevent the audit service from starting. However, new audit records might not match the database schema causing auditing to fail. In this case, auditing is no longer possible. Therefore, contact IBM Support if you come across such issues.
3. Apply additional security updates by running the following command:

   ```
   /siq/bin/postupgrade.sh
   ```

When the script completes, the message `Update complete` is displayed.

Upgrading to 7.6.0.18

Gateway and data server

After the update package is installed and the system is restarted, complete the upgrade by applying some additional security update:

1. Log out from the system (as `util` user) and log back in as `root` user.
2. Run the following command:

   ```
   /usr/local/storediq/bin/tools/postupgrade.sh
   ```

   When the script completes, the message `Update complete` is displayed.
3. To verify the security updates, complete these steps:
   a. Restart the OpenSSH daemon by running this command:

      ```
      service sshd restart
      ```

   b. Check the new configuration by running the following command:

      ```
      sshd -T|grep -E "^(ciphers|macs|kexalgo)"
      ```

      This command should return the following output:

      ```
      ciphers aes128-ctr,aes192-ctr,aes256-ctr
      macs hmac-sha2-256,hmac-sha2-512
      kexalgorithms diffie-hellman-group-exchange-sha256
      ```

Application stack

After the production mode is activated, apply additional security updates:

1. Run the following command:

   ```
   /siq/bin/postupgrade.sh
   ```

   When the script completes, the message `Update complete` is displayed.
2. To verify the security updates, complete these steps:
   a. Restart the OpenSSH daemon by running this command:

      ```
      service sshd restart
      ```

   b. Check the new configuration by running the following command:

      ```
      sshd -T|grep -E "^(ciphers|macs|kexalgo)"
      ```

      This command should return the following output:

      ```
      ciphers aes128-ctr,aes192-ctr,aes256-ctr
      macs hmac-sha2-256,hmac-sha2-512
      kexalgorithms diffie-hellman-group-exchange-sha256
      ```

Then, you can check the vault status and close port 80 on the gateway.

Enabling disk encryption on an Elasticsearch cluster upgraded to 7.6.0.15

For an Elasticsearch cluster that you upgraded to IBM StoredIQ 7.6.0.15, follow these instructions to encrypt the volume on which the Elasticsearch indexes are stored. IBM StoredIQ uses Linux Unified Key Setup (LUKS) for disk encryption.

Make sure that the cluster works correctly and all nodes are available by running these commands:

```
curl -X GET 'http://your first node IP:9200/_cluster/state?pretty'
curl -X GET 'http://your first node IP:9200/_nodes?pretty'
```

New installations of IBM StoredIQ Version 7.6.0.15 and later by default encrypt the disk volume on which the Elasticsearch indexes are stored. However, if you installed a previous version and upgraded to Version 7.6.0.15, the existing volumes were not encrypted automatically during the upgrade process. For security reasons, you should encrypt the volume. However, you cannot decrypt the volume again later.

1. For each node, determine whether the Elasticsearch data on the disk is less or greater than the available storage on the root partition.

   The root partition can be used for backing up the existing data.

   - If the data is less than the available storage, proceed with step "2" on page 67.
   - If the data on the disk is greater than the available storage, add a disk of the required size to the VM and mount it at `/mnt/backup`.
2. Restart the VM on each node to ensure that no outside processes or connections are using the data partitions. If you added a disk, ensure that it is still mounted at `/mnt/backup`.
3. Log in to the primary master node from PuTTY by using the **builder** account and password that were configured during the installation of the cluster. Edit the `cluster-setup.properties` file by appending LUKS=`true` on a new line at the end of the file.
4. On the primary master node, run the cluster setup script. At the prompt `[builder@localhost ~]$`, enter:

   ```
   /siq/bin/cluster-setup.sh cluster-setup.properties
   ```

   Note: If these steps were not followed correctly, the process might fail.

   Enter the password to continue for each node.
5. Indexes are now encrypted but the key is accessible on non-encrypted disks.

   If you want to secure the key with a LUKS password, complete these steps for each node in the Elasticsearch cluster:

   a) Log in from the vCenter Console by using the root account and password that were configured during the installation of the cluster.

   b) Add a key by using the **cryptsetup** tool. Issue the following command:

   ```
   [root@localhost ~]$ sudo cryptsetup -d /root/siq-elasticsearch-luks.key luksAddKey /dev/
   sdb1
   ```

   c) Enter the new passphrase for the key slot and verify it.

   Important: You must remember that passphrase or keep it in a safe place. If you store the passphrase, make sure to keep it in a place away from the information it protects.

   d) Remove the LUKS key after you verify the password by using shred utility:

   ```
   [root@localhost ~]$ sudo shred -zvu -n 5 /root/siq-elasticsearch-luks.key
   ```

Upgrading the Elasticsearch cluster

If you have a DataServer - Distributed, you must upgrade the Elasticsearch cluster first before you upgrade any data server. If you have only the DataServer - Classic, skip this procedure and proceed to the next topic.

Follow these steps to upgrade the Elasticsearch cluster.

1. Download the release upgrade ISO to the primary master node of the Elasticsearch cluster where the installation was initially started.

   Download the latest version from IBM Fix Central. For information about the package names and part numbers and the links to the proper download locations, see the download document.

2. Log on to the primary master node from PuTTY by using the `builder` account and password that were configured during the installation of the cluster. Run the following commands to upgrade the cluster.
3. To prepare the upgrade, run the following commands replacing *STOREDIQ-Upgrade.iso* with the image name stated in the download document:

```
[builder@hostname ~]# su
[root@hostname builder]# cd /home/builder
[root@hostname builder]# mount STOREDIQ-Upgrade.iso /mnt
[root@hostname builder]# cd /mnt
[root@hostname mnt]# /usr/bin/python -m SimpleHTTPServer 80 >/dev/null 2>&1 &
[root@hostname mnt]# export killpid=$! && cd
```

   Hint: When you change to the `root` account, you will be prompted for the password. Enter the password for `root` that was configured during the installation of the cluster.
4. In a second PuTTY session, log in to the primary master node by using the `builder` account and password that were configured during the installation of the cluster.

   Then, run the following commands to upgrade the cluster.

```
[builder@hostname ~]$ bootstrap product.set elasticsearch repository http://IP of primary
master node/products/elasticsearch
```

   Press Enter to continue for each node.

```
[builder@hostname ~]$ bootstrap product.upgrade elasticsearch 7.6.0.19+3
```

   Press Enter to continue for each node.
5. Go back to the PuTTY session from step 3 where you're logged in as `root` and run these commands to finish the upgrade:

```
[root@hostname ~]$ kill $killpid
[root@hostname ~]$ umount /mnt
```

Upgrading the gateway and data servers

Download the upgrade ISO and then upgrade the gateway and data servers.

Important: The upgrade to version 7.6.0.19 entails an upgrade of PostgreSQL. Therefore, it is strongly recommended that you back up your data before starting the upgrade. Otherwise, you will not be able to recover your data if the upgrade fails. For more information, see the topic about backing up the IBM StoredIQ image.

To avoid space issues during the PostgreSQL upgrade, make sure that at least 15% of the space on the Postgres data volume are free before you start the upgrade.

To avoid space issues with the `/boot` partition during the upgrade, complete the steps in “Cleaning up the /boot partition” on page 70 on the gateway and each data server before you start the upgrade.

You must upgrade the gateway before upgrading any data servers.

1. Download the release upgrade ISO from IBM Fix Central and store it in the `/tmp` folders on the gateway and on each data server that you want to upgrade.

   Download the latest version to your local machine by using one of the options provided on Fix Central. Then, copy it to the gateway and data servers by using a copy tool, for example, PuTTY Secure Copy client on Windows.

   For information about the package names and part numbers and the links to the proper download locations, see the download document.

Complete the following steps for the gateway first and then for each data server that you want to upgrade.

2. Log in to the virtual machine as `util util`.
   a) Use the arrow and Tab keys to select **Manage Software Repositories**, and then select **OK**.

b) Use the arrow and Tab keys to select **DVD Update Repository** and then select **Edit**.

c) Use the arrow keys, the space bar, and the Tab key within **Edit Update Repository** to select **ISO** within **Type**, and then select **Next**.

d) Enter the location of the upgrade ISO and the name of the ISO such as `/tmp/7.6.0.19-STOREDIQ-Upgrade-NN.iso`, and then select **Finish**.

e) In the **Manage Update Repositories** screen, select **Done**.

3. When prompted, save the repository configuration changes by selecting **Yes**.
4. In the **Appliance Manager Screen**, select **Update This Node** and select **OK**.
5. In **Choose Update Repositories**, select **DVD Update Repository** and then select **OK**.
6. After the new package updates are listed, select **OK**.

   The **Back up database?** message appears. If you took a VM snapshot, select **No**. However, it is essential that you have a current backup or VM snapshot. Otherwise, you will not be able to recover your data if the upgrade fails. When you're done, you are notified that a new update package was installed.

   If the PostgreSQL upgrade fails, an error message similar to the following one is shown:

   ```
   The database upgrade script failed on this node: '198.51.100.10'
   ```

   The error might be due to insufficient space on the Postgres data volume. Therefore, complete the following steps as root user before contacting IBM Support:

   a. Check the update log file on the node by running this command:

   ```
   /deepfs/config/update/update.log
   ```

   The output might look similar to this example:

   ```
   ********************************************************************************
   ***
   More than   85% of the Postgres data volume capacity is currently used.
   'The remaining capacity might not be sufficient for data migration.

    Back up your data and enlarge the volume mounted on /dev/sdb before restarting the
   upgrade.

   ********************************************************************************
   ***
   Processing stopped
   ```

   b. Back up the data and increase the size of the volume as directed.

   c. To complete the package update, run the following command:

   ```
   /usr/local/storediq/bin/continue_update.py
   ```

   For easier reference, redirect the output of the `continue_update.py` script to a log file:

   ```
   /usr/local/storediq/bin/continue_update.py > continue_update.log
   ```

   When you run the script, skip steps “7” on page 69 to “10” on page 69 of the regular update procedure.

7. Select **Restart**. At the command prompt, enter `su util`.
8. Select **Update This Node** again and then select **OK**.

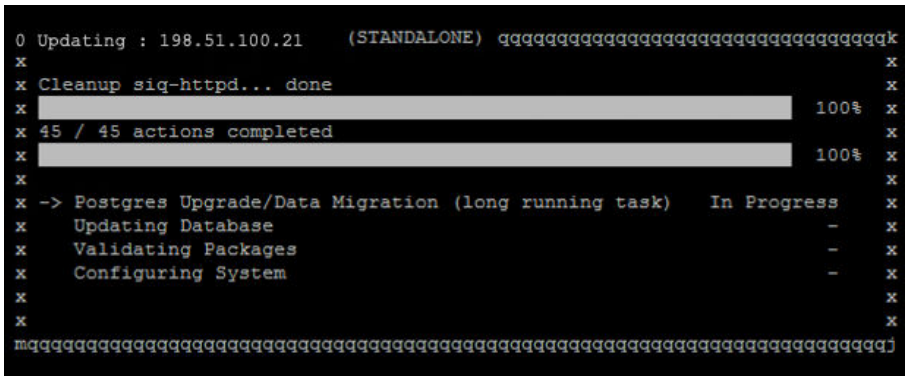
   The **Choose Update Repositories** window opens.
9. Select **DVD** and then select **OK**.

   The available packages are listed again.
10. After the new package updates are listed, select **OK**.

The **Back up database?** message appears. If you took a VM snapshot, select **No**. You are notified that a new update package was installed.

Several informational messages appear, indicating that the repository is being created, what new package updates are available, that update transactions are being created (including the PostgreSQL upgrade), and that the system is restarted.

Important: Depending on the amount of data that needs to be migrated, the PostgreSQL upgrade might take very long. Do not try to pause or stop the upgrade.



11. To complete the upgrade, you must apply some additional security update:
    a) Log out from the system (as `util` user) and log back in as `root` user.
    b) Run the following command:

    ```
    /usr/local/storediq/bin/tools/postupgrade.sh
    ```

    When the script completes, the message `Update complete` is displayed.

You might want to extend the range of allowed host names on the data server to allow clients to use other host names in their URLs than the default ones. For more information, see "Securing the data server against host header injection vulnerabilities" on page 55

Cleaning up the /boot partition

To avoid space issues with the `/boot` partition during the upgrade of the gateway or a data server, clean up the partition before you start the upgrade.

Complete these steps on the gateway and on each data server.

1. Log in to the VM as root user or a user with sudo access.
2. On the `/boot` partition, check the available space by running the following command:

   ```
   df -h
   ```

3. Check which kernel is in use by running this command:

   ```
   uname -r
   ```

4. Check which other versions are installed by running this command:

   ```
   rpm -qa |grep kernel
   ```

5. To free space, remove all unused kernels. Run the following command replacing *kernel_version_123* with the appropriate version:

   ```
   rpm -e kernel_version_123
   ```

Upgrading the application stack

After you upgrade the gateway and data servers, you can proceed to upgrade the application stack.

Verify that the gateway and data servers are already upgraded.

Important: It is strongly recommended that you back up your AppStack data before starting the upgrade. Otherwise, you will not be able to recover your data if the upgrade fails. For more information, see the topic about backing up the IBM StoredIQ image in the administration guide.

- Keep the services running in the AppStack before you start an upgrade.
- Check the `/var/bootstrap/appstack.log` file whenever an upgrade is running.
- Check the `/var/siq/log/tomcat.log` after upgrading.

1. Use an SSH tool to log in to the gateway and complete these steps.
   a) Open port 80 by running this command:

   ```
   iptables -A PROD-web -p tcp -m state --state NEW -m tcp --dport 80 -j ACCEPT
   ```

   b) Restart the services on the gateway using the **service deepfiler restart** command.
   c) In a browser, you can go to `http://`*`gatewayIP`*`/products` to check whether the product package is available.
2. Use an SSH tool to log in to the application stack as `root` user and run these commands.
   a) Run this command:

   ```
   bootstrap product.set appstack repository "pkgsource=\"http://gatewayIP/products\""
   ```

   b) Run this command:

   ```
   bootstrap product.upgrade appstack HEAD
   ```

   c) Monitor the upgrade process by tailing the `appstack.log` file:

   ```
   tail -f /var/bootstrap/appstack.log
   ```

   When the upgrade is complete, you are notified that the production mode is activated.
   d) Examine the `/var/bootstrap/appstack.log` log file for any traceback or any keywords such as ProgrammingError, OperationalError. Contact IBM Support immediately if you come across such issues.
3. Apply additional security updates by running the following command:

   ```
   /siq/bin/postupgrade.sh
   ```

   When the script completes, the message `Update complete` is displayed.
4. Check whether the vault service is running by using this command:

   ```
   service vault status
   ```

   If the vault service is not running after the upgrade (which is indicated by the message `Vault server is not running`), start it manually by using this command:

   ```
   service vault start
   ```

   Tip: You can run these commands as `root` or as `siqadmin` user.
5. After the upgrade of the application stack is complete, log in to the gateway as `root` by using an SSH tool and complete these steps.
   a) Close port 80 by running this command:

   ```
   iptables -D PROD-web -p tcp -m state --state NEW -m tcp --dport 80 -j ACCEPT
   ```

b) Restart services on the gateway by running the **service deepfiler restart** command.

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

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

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

Any references in this information to non-IBM 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 Director of Licensing
IBM Corporation
North Castle Drive, MD-NC119
Armonk, NY 10504-1785
US

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.

The performance data discussed herein is presented as derived under specific operating conditions. Actual results may vary.

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

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

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

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.

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.

IBM Online Privacy Statement

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.

This Software Offering does not use cookies or other technologies to collect personally identifiable information.

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.

Index

L

M

N

O

P

R

S

T

U

V