About This Document

This document is a guide written to provide users with information about configuring and using the Data Dynamics StorageX Management Portal (StorageX Management Portal) and StorageX Retrieval Portal (Retrieval Portal) products, including the system and software, steps for configuration, and how to use the product in your environment. This preface contains the following sections:

- Documentation conventions ........................................... i
- Documentation feedback .............................................. ii
- Contacting Support ..................................................... ii

Documentation conventions

This section describes text formatting conventions and important notices formats.

Text formatting

The narrative-text formatting conventions that are used in this document are as follows:

**bold text** Identifies command names

Identifies the names of user-manipulated GUI elements

Identifies keywords and operands

Identifies text to enter at the GUI or CLI

*italic text* Provides emphasis

Identifies variables

Identifies paths and Internet addresses

Identifies document titles

**code text** Identifies CLI output

Identifies syntax examples

For readability, command names in the narrative portions of this guide are presented in mixed lettercase: for example, switchShow. In actual examples, command lettercase is often all lowercase. Otherwise, this manual specifically notes those cases in which a command is case-sensitive.

Note and attention statements

The following note and attention statements are used in this documentation. They are listed below in order of increasing severity of potential hazards.
NOTE
A Note provides a tip, guidance or advice, emphasizes important information, or provides a reference to related information.

ATTENTION
An Attention indicates potential damage to hardware or data.

Documentation feedback

Because quality is our first concern at Data Dynamics, we have made every effort to ensure the accuracy and completeness of this document. However, if you find an error or an omission, or you think that a topic needs further development, we want to hear from you. Forward your feedback to: documentation@datdyn.com

Provide the title and version number of the document and as much detail as possible about your comment, including the topic heading and page number and your suggestions for improvement.

Contacting Support

If you encounter technical problems, send an email to support@datdyn.com or go to the Data Dynamics, Inc. Support site at www.datdynsupport.com.

Include the following information in your email or support request:

• Product name, version, and build number
• Operating system on which StorageX is installed
• Internet browser you use to access StorageX
• Your company name
• Your name and a phone number where we can reach you
• Your question or issue
About This Document

Documentation conventions ........................................ i
Documentation feedback ........................................... ii
Contacting Support ..................................................... ii

Chapter 1 Understanding analysis and archival in StorageX

In this chapter .......................................................... 1
Understanding the StorageX Management Portal ........... 2
Understanding gathering data ................................. 2
Understanding file analysis ........................................ 3
Understanding archiving to objects ...................... 4
Understanding workload tiering ......................... 4
Understanding application transforms .......... 4

Chapter 2 Getting started with the StorageX Management Portal

In this chapter .......................................................... 7
Understanding StorageX Management Portal requirements ...... 8
Supported operating systems ........................................ 8
Database and message channel requirements .................. 9
Minimum hardware requirements .................................. 9
Optimum hardware requirements .............................. 12
Permissions requirements ........................................ 13
Network port requirements ....................................... 13
Web service requirements ....................................... 14
Additional software requirements .......................... 14
Web browser requirements ................................ 15

Installing and configuring an Elasticsearch cluster ........ 15
Installing and configuring a Redis cluster .................. 17
Installing and configuring an Apache Cassandra cluster for index backup 20
Installing StorageX Management Portal and Metadata Service components ........................................ 24
Preparing to use the StorageX Management Portal ........ 28
Accessing the StorageX Management Portal ................ 29
Configuring monitoring for StorageX Management Portal components ........................................ 30
Upgrading the StorageX Management Portal ................ 32
Uninstalling the StorageX Management Portal ............ 35

Chapter 3 Using the StorageX Management Portal

In this chapter .......................................................... 37
Getting your storage resource data ........................ 37
Analyzing your storage resource data ...........................................40
Working with analysis results .........................................................43
Archiving your file data to an object storage resource .................45
Using Workload Tiering to migrate your file data to a different file storage resource .............................................................48
Viewing and modifying scans, datasets, analysis sets, and policies .................................................................50
Deleting scans, analysis sets, and policies .......................................52
Configuring Cassandra grooming settings .......................................53
Accessing Management Portal monitoring tools ............................54

Chapter 4 Installing and Using the StorageX Retrieval Portal

In this chapter .................................................................55
Understanding the StorageX Retrieval Portal ..................................55
Understanding StorageX Retrieval Portal requirements ..................55
Preparing to install the StorageX Retrieval Portal .........................57
Installing the StorageX Retrieval Portal .......................................58
Accessing the StorageX Retrieval Portal ......................................60
Retrieving files using the StorageX Retrieval Portal ......................60
Upgrading the StorageX Retrieval Portal .....................................62
Uninstalling the StorageX Retrieval Portal .................................63
The following Guide was created to provide information for installing, configuring, and using the Data Dynamics StorageX Management Portal and Data Dynamics StorageX Retrieval Portal.

This chapter provides overview information about the StorageX Management Portal and StorageX Retrieval Portal, as well as storage resource analysis, archiving file data, workload tiering, and application transforms using StorageX.

**NOTE**
This Guide does not cover StorageX functionality in the main StorageX Console, including migration, replication, DFS namespace management, and disaster recovery. For information about that StorageX functionality, see the *StorageX Administrator's Guide*.

**In this chapter**

- “Understanding the StorageX Management Portal” on page 2
- “Understanding gathering data” on page 2
- “Understanding file analysis” on page 3
- “Understanding archiving to objects” on page 4
- “Understanding workload tiering” on page 4
- “Understanding application transforms” on page 4
Understanding the StorageX Management Portal

The purpose of the StorageX Management Portal is to provide you with a useful view of your environment, to allow you to quickly and easily gather data, analyze that data, and – acting on that analysis – archive data that meets your criteria to a specified object store or perform workload tiering to copy data that meets your criteria to a different file storage resource.

The overall StorageX Management Portal process can be broken into four phases:

Set up your StorageX environment
Configure StorageX to manage the storage resources you want to scan, analyze, and archive in the Management Portal or using the StorageX API. For information about setting up StorageX and the StorageX Management Portal, see “Getting started with the StorageX Management Portal” on page 7.

Get data on your resources
Collect information on the storage resources in your environment, as necessary.

Analyze your file data
Analyze the collected file storage resource data to identify files that are interesting to you.

Archive your files to an object storage resource or tier your files to a different file storage resource
Transform a selected dataset from files to objects and store those objects in an object store or migrate a selected set of files from one file storage resource to another, based on criteria you determine.

Understanding gathering data

Before you can use the StorageX Management Portal to analyze your data, you must first gather data from the storage resources that are important to you.

The StorageX Management Portal lets you scan the storage resources in your environment and get data about every file, including all the metadata for each file, and store that information in a repository, which is a clustered Elasticsearch database, that can then be searched and queried to learn things about your environment. You can scan a single resource, an entire department’s worth of resources, or every server in your company, and pull together data about all of those files.

You can also tag your scanned data with custom tags, which enables you to better organize your resource data. For example, if you want to look specifically at your company’s Marketing resource data and usage, you could scan your Marketing department’s resources and tag all of the gathered data with the custom tag Department, with the value Marketing.

The Management Portal then lets you analyze only data with the Department tag and the value Marketing, or some further subset of that data set.

For more information about gathering file data in your environment, see “Getting your storage resource data” on page 37.
Understanding file analysis

After gathering data about the storage resources in your environment, the StorageX Management Portal enables you to analyze that data in a number of ways.

To use the Management Portal to analyze your data stored in the StorageX repository, you build an analysis set. An analysis set consists of two separate pieces, which can be re-used to build other analysis sets after being defined: a dataset and a query. These two components are defined as follows:

Dataset
The dataset is the overall set of files in the repository that you want to look at. You can create a dataset in one of the following ways:

- Select a dataset based on the name or description of a specific scan, using the isx:Scan Name or isx:Scan Description tags
- Select a dataset based on custom tags you specified when running one or more scans

For example, if you created and ran a scan named Marketing1, you could specify Marketing1 as the value for the isx:Scan Name tag.

Alternatively, if you created and ran scans where you assigned the custom tag Dept to all the scanned files, with the value Marketing, you could then specify Marketing as the value for the Dept tag in the Dataset Criteria. In this way, you could look at data collected over multiple scans of your resources but all belonging to the same department.

Query
The query is the set of criteria you want to use to filter specific files from the dataset. While the dataset is the overall data you want to search, the query narrows down the focus to the actual files you want.

In the Query Criteria, you can select one or more file attributes to look for or exclude, grouping those conditions together as necessary. For example, if you want to get a list of every file in your dataset owned by the user cbranan, you would select the Owner attribute, select the equal operator, and specify cbranan as the value.

You could also use the query to find old, unused data in your storage environment. For example, if you wanted to find files that no user has modified in more than two years, you could select the LastModified Time attribute, select the greater or equal operator, select 2 for the value, and then select years as the unit of time.

Additionally, you can also use tags during analysis, assigning specific tags to the files StorageX returns from the repository. This can be useful, for example, if you want to come back later and archive a particular subset of files. You could tag the returned files with the custom tag Archive, with the value Next Quarter, and then collect all files with that tag together at a later date for archival in an object storage resource.

After analyzing your data, you can then view your data in various formats. For more information about viewing analysis results, see “Working with analysis results” on page 43.

You can also archive file data using an existing analysis set. For more information about archiving file data to objects, see “Archiving your file data to an object storage resource” on page 45.

For more information about analyzing file data, see “Analyzing your storage resource data” on page 40.
Understanding archiving to objects

StorageX allows you to use the StorageX Management Portal to find and select file data that you want to archive from its original file format, using either the CIFS or NFS protocol, to a supported object storage resource.

For example, you could decide that you want the “cold data” in your company’s storage environment to be shifted off of your existing file storage resources and onto a object storage resource. Because the data is infrequently used or accessed, the files can be transformed into objects and stored in a private or public cloud, saving you storage space on your file-based systems, where you can keep data that users need to access regularly.

StorageX takes each file you want to archive and ingests the file, storing the file content, alternate data streams, and access control lists (ACLs) in one or more objects. The product automatically keys the objects so they can be found and retrieved in the object store.

You can configure StorageX to archive files on a scheduled basis, and you can tag your archived files in the object store so that you can more easily retrieve the files you want.

Users who need to access a specific file or files at a later date can then request that those files using the StorageX Retrieval Portal.

For more information about archiving files to objects, see “Archiving your file data to an object storage resource” on page 45.

Understanding workload tiering

You can also use StorageX to find file data that you do not necessarily want to archive to objects but that you want to shift to a different location on another file storage resource.

For example, you might have several file storage resources in your environment, each with different levels of speed and efficiency. In that situation, it may make sense to take “cold data” and move it to a lower-speed storage resource, and move “hot data,” files that are accessed often and need to be available quickly, to a higher-speed storage resource.

In StorageX, this practice is known as workload tiering. Files that need a faster response time are shifted to a higher-speed tier, while less-critical files are shifted to a lower-speed tier.

For more information about using workload tiering, see “Using Workload Tiering to migrate your file data to a different file storage resource” on page 48.

Understanding application transforms

In some situations, you may want to move your files from one or more file storage resources to an object storage resource all at one time, without needing to perform any kind of analysis or selection. You simply need to move a large number of files to the object store.

This may be necessary if, for example, you upgrade your company’s package-tracking system to a new system that operates in the cloud using object storage. You then have a system full of files that need to be transformed en masse to objects and put into the object storage resource. The same thing could also be necessary in reverse, where a large number of objects need to be transformed into files at one time.
StorageX enables you to perform this kind of application transform outside of the StorageX Management Portal user interface. Instead, you use the StorageX API, providing a list of files or objects to transform and the target for the transform and running file-to-object or object-to-file transformation API calls to transform your data.

For more information about performing application transforms, see the StorageX Administrator's Guide and the StorageX API Guide.
1. Understanding application transforms
Getting started with the StorageX Management Portal

This section explains how to install, configure, and access the StorageX Management Portal product in your environment.

In this chapter

- “Understanding StorageX Management Portal requirements” on page 8
- “Installing and configuring an Elasticsearch cluster” on page 15
- “Installing and configuring a Redis cluster” on page 17
- “Installing and configuring an Apache Cassandra cluster for index backup” on page 20
- “Installing StorageX Management Portal and Metadata Service components” on page 24
- “Preparing to use the StorageX Management Portal” on page 28
- “Accessing the StorageX Management Portal” on page 29
- “Configuring monitoring for StorageX Management Portal components” on page 30
- “Upgrading the StorageX Management Portal” on page 32
- “Uninstalling the StorageX Management Portal” on page 35
Understanding StorageX Management Portal requirements

This section describes the system requirements necessary for the StorageX Management Portal, StorageX Metadata Service components, Elasticsearch repository, Redis message channel, and Apache Cassandra database (optional, but recommended) to function properly.

This following topics in this section provide detailed information on system requirements for using the StorageX Management Portal:

- “Supported operating systems” on page 8
- “Database and message channel requirements” on page 9
- “Minimum hardware requirements” on page 11
- “Optimum hardware requirements” on page 12
- “Permissions requirements” on page 13
- “Network port requirements” on page 13
- “Web service requirements” on page 14
- “Additional software requirements” on page 14
- “Web browser requirements” on page 15

Supported operating systems

The StorageX Management Portal and StorageX Metadata Service must each be installed on a computer running one of the following operating systems with the latest service pack:

- Microsoft Windows Server 2012 Essentials, Standard, or Datacenter Edition
- Microsoft Windows Server 2012 R2 Essentials, Standard, or Datacenter Edition
- Microsoft Windows Server 2016 Essentials, Standard, or Datacenter Edition

Each node of the Elasticsearch cluster you want to use as a repository for file data can be installed on a computer running either Windows or Linux.

For more information about installing Elasticsearch and specific versions and distributions supported, see the following links:

- https://www.elastic.co/support/matrix

Each node of the Redis cluster you want to use for messaging between StorageX components can be installed on a computer running Linux.

**NOTE**

We do not recommend installing Redis on a Windows server in a production environment. You can only deploy a one-node Redis cluster on a single Windows server, and this type of installation does not provide sufficient scalability or redundancy. We recommend installing all Redis cluster nodes on Linux-based servers.

For more information about installing Redis, see the following links:

- https://redis.io/documentation
- https://redis.io/topics/quickstart
If you want to use Apache Cassandra as a backup system for the archival process, each node of the Cassandra cluster must be installed on a computer running Linux. For more information about installing Cassandra, see http://cassandra.apache.org/.

Database and message channel requirements

StorageX uses the following database- and message channel-related components in order for the StorageX Management Portal to function properly.

Elasticsearch

StorageX requires an Elasticsearch installation, which serves as the repository for information about all scanned file data, including all available metadata and custom tags. StorageX requires Elasticsearch version 6.4.2, 6.5.x, 6.6.x, or 6.7.x.

If you intend to use the StorageX Management Portal to scan more than a very minimal amount of data, we strongly recommend using a multi-node, clustered Elasticsearch installation with a minimum of 3 nodes.

A clustered Elasticsearch installation provides improved scalability, allowing StorageX to better distribute the load on the database as a whole and also provides high availability of your scanned file data should one or more nodes experience an outage.

NOTES:

• If you use a single-node, unclustered Elasticsearch installation for the StorageX Management Portal, you risk losing data stored in the repository in the event that the server hosting Elasticsearch experiences a failure. Using a clustered installation allows Elasticsearch to fail over from one node to another if a failure occurs.

• StorageX does not support Elasticsearch version 7.0.x. If you use a version 7.0.x installation of Elasticsearch with the product, the StorageX Management Portal will not function properly.

• You must install the StorageX Metadata Service and all Elasticsearch cluster nodes on separate computers.

• We recommend installing the most recent version of Elasticsearch.

The Elasticsearch installation can be local or remote, but the Management Portal server must have access to all nodes of the cluster.

Elasticsearch requires the following prerequisite:

• Java 8 or later

You can use either the Oracle Java Standard Edition or OpenJDK, version 8 or later. Elastic specifically recommends using Java version 1.8.0_131 or later.

You can download Java from the following Web pages:


http://openjdk.java.net/

To ensure that you have the correct version of Java installed, enter java -version in a command-line window.

For more information about Elasticsearch, see

Redis

StorageX also requires a Redis installation, which StorageX uses as a communication channel between the StorageX Metadata Service Processing Engine and the Web Service (API). StorageX requires Redis version 3.0 or later. We recommend using the latest available version of Redis.

NOTE
We do not recommend using any version of Redis Enterprise with your StorageX Management Portal installation.

We recommend using a multi-node, clustered Redis installation on Linux-based servers with a minimum of 3 nodes. We recommend installing Redis on servers running the Ubuntu Linux distribution, but other Linux distributions may work, as well.

When using StorageX in a large environment with large amounts of data, you may want to explore creating a larger Redis cluster installation with seven or more nodes.

NOTE
We do not recommend installing Redis on a Windows server in a production environment. You can only deploy a one-node Redis cluster on a single Windows server, and this type of installation does not provide sufficient scalability or redundancy. We recommend installing all Redis cluster nodes on Linux-based servers.

Redis requires the following prerequisites on Linux-based servers:

• make
• GCC (gcc)
• libc6-dev
• Tcl (tcl)

You can install the necessary prerequisites using APT. For more information about Redis and Redis clusters, see https://redis.io/documentation and https://redis.io/topics/cluster-tutorial.

Apache Cassandra

If you want to use StorageX to archive your scanned data or do workload tiering, we recommend installing and configuring Apache Cassandra in your environment. StorageX uses Cassandra as a backup for all indices created by Elasticsearch during the scan process.

In the event that something happens to your Elasticsearch cluster, any corrupted or deleted indices can be rebuilt from Cassandra. We recommend using the latest available version of Cassandra.

Cassandra requires the following prerequisites:

• Java 8
  You can use either the Oracle Java Standard Edition 8 or OpenJDK 8. You can download these versions of Java 8 from the following Web pages:
  http://openjdk.java.net/
  To ensure that you have the correct version of Java installed, enter java -version in a command-line window.
• Python 2.7
You can download Python from the following Web page:

https://www.python.org/downloads/

To ensure that you have the correct version of Python installed, enter `python --version` in a command-line window.

Because of these prerequisites, we recommend installing your Cassandra cluster on servers with Red Hat Enterprise Linux 7 or CentOS 7 installed.

After installing Cassandra, you must specify or create a keyspace for your installation. You can use the native Cassandra CQL shell, `cqlsh`, which is installed with StorageX Cassandra components, to create a new keyspace.

For basic information about installing and configuring an Apache Cassandra cluster, see “Installing and configuring an Apache Cassandra cluster for index backup” on page 20.

For more information about Apache Cassandra, see http://cassandra.apache.org/. For more information about creating a Cassandra keyspace, see https://docs.datastax.com/en/cql/3.3/cql/cql_reference/cqlCreateKeyspace.html.

### Minimum hardware requirements

The following requirements are the minimum supported for installing and running StorageX Metadata Service components, the Elasticsearch cluster, the Redis cluster, and the optional Apache Cassandra cluster.

**NOTE**

These minimum requirements may not be suitable for high-performance installations.

**StorageX Metadata Service components**

Each component of the StorageX Metadata Service has specific minimum requirements in terms of memory and storage space.

**StorageX Metadata Service Web Service**

The Web Service component requires a minimum of 14 MB of storage space, 4 GB of memory (RAM), and at least 4 cores.

**StorageX Metadata Service Processing Engine**

The Processing Engine component requires a minimum of 10 MB of storage space, 4 GB of memory (RAM), and at least 4 cores.

**Elasticsearch cluster node**

Each node in your Elasticsearch cluster requires a minimum of 16 GB of memory (RAM) and at least 2 cores. If possible, we recommend using at least 32 GB of memory and 8 cores.

**Redis cluster node**

Redis uses a relatively small amount of disk space but is RAM-intensive, using memory for its messaging operations. Each node in your Redis cluster requires a minimum of 3 MB of storage space, 15 GB of memory (RAM), and at least 4 cores. If possible, we recommend using at least 30 GB of memory and 8 cores.
In addition, we recommend installing your Redis cluster on Linux-based servers.

**NOTE**
We do not recommend installing Redis on a Windows server. You can only deploy a one-node Redis cluster on a single Windows server, and this type of installation does not provide sufficient scalability or redundancy.

For detailed information on Redis memory requirements and optimization, see the following Redis Websites:

- https://redis.io/documentation
- https://redislabs.com/redis-enterprise-documentation/administering/designing-production/hardware-requirements/

**Apache Cassandra cluster node**

Each node in your Cassandra cluster requires a minimum of 16 GB of memory (RAM) and at least 8 cores. If possible, we recommend using at least 32 GB of memory and 16 cores.

**Optimum hardware requirements**

When using StorageX to scan and analyze large numbers of files, we recommend using a distributed installation that is significantly more robust than the minimum requirements.

The following suggested specifications are based on scanning and analyzing a dataset of 1 million files.

**StorageX Metadata Service components**

For a dataset of this size, we recommend installing the StorageX Metadata Service components on separate servers, each with increased memory (RAM), number of CPU cores, and storage space.

**StorageX Metadata Service Web Service**

For the Web Service component in this scenario, we recommend a minimum of 10 GB of storage space, 8 GB of memory (RAM), and at least 8 cores.

**StorageX Metadata Service Processing Engine**

For the Processing Engine component in this scenario, we recommend a minimum of 10 GB of storage space, 8 GB of memory (RAM), and at least 8 cores.

**Elasticsearch cluster node**

For a dataset of this size, we recommend installing each Elasticsearch node in your cluster on a server with 32 GB of RAM and 8 cores, as well as at least 3 TB of storage space.

In addition, we recommend installing your Elasticsearch nodes on servers with SSD-based storage to improve indexing and querying performance.

**Redis cluster node**

For a dataset of this size, we recommend installing each Redis node in your cluster on a server with 30 GB of RAM and 8 cores, as well as at least 20 GB of storage space.
In addition, we recommend installing all Redis cluster nodes on Linux-based servers.

**NOTE**

We do not recommend installing Redis on a Windows server. You can only deploy a one-node Redis cluster on a single Windows server, and this type of installation does not provide sufficient scalability or redundancy.

For detailed information on Redis memory requirements and optimization, see the following Redis Websites:

- https://redis.io/documentation
- https://redislabs.com/redis-enterprise-documentation/administering/designing-production/hardware-requirements/

**Apache Cassandra cluster node**

For a dataset of this size, we recommend installing each Cassandra node in your cluster on a server that meets the minimum requirements above (at least 32 GB of RAM and 8 cores) but which also includes at least 1 TB of storage space.

On each replica in the Cassandra cluster, we recommend reserving at least 520 MB of storage space for the Cassandra keyspace. For more information about keyspaces and replication in a Cassandra cluster, see the Apache Cassandra documentation at http://cassandra.apache.org/.

**Permissions requirements**

The account used to install the StorageX Management Portal must have Domain Administrator, Local Administrator, or Power User access on all file systems, servers, and Cassandra nodes.

In addition, the account used to run the StorageX Metadata Service must be the main StorageX service account, and must have the permissions required to log on as a service on any computer where you want to install StorageX Management Portal and StorageX Metadata Service components.

**NOTES:**

- You must ensure the account you want to use as the StorageX Metadata Service service account has the Log on as a service privilege on a computer before installing any Management Portal or Metadata Service components on that computer.
- Specific levels of access may be required for some types of storage resource. For more information about permissions required to manage your resources, see the StorageX Administrator’s Guide.

**Network port requirements**

The StorageX Management Portal uses the following TCP ports for communication:

- The StorageX Management Portal and REST API listen on TCP port 9777.
- The StorageX Metadata Service listens on TCP port 9464.
- The StorageX Metadata Service Web Service/REST API listens on TCP port 9462.

These ports must be open in any firewalls located between the Management Portal and the StorageX server.
Understanding StorageX Management Portal requirements

The Elasticsearch cluster uses the following TCP ports, by default:
- 9200 for communication using the Elasticsearch API
- 9300 for communication between cluster nodes and with the Java client

The following Web page provides more detailed information about Elasticsearch port requirements:

The Redis cluster uses the following TCP ports, by default:
- 6379 for client communication
- 16379 for communication between cluster nodes

The following Web page provides more detailed information about Redis port requirements:
https://redis.io/topics/cluster-tutorial.

The Cassandra cluster uses the following TCP ports, by default:
- 7000/7001 (if SSL is enabled) for communication between cluster nodes
- 7199 for Java Management Extensions (JMX) monitoring
- 9042 and 9160 for client communication
- 9142 for client communication if SSL is enabled

The following Web pages provide more detailed information about Apache Cassandra port requirements:

Web service requirements

The computer where you want to install the StorageX Management Portal must have the HTTP service enabled.

Additional software requirements

Any computer where you want to install one of the StorageX Metadata Service components must have the 64-bit version of the Microsoft Visual C++ Redistributable Package for Visual Studio 2013 installed.

NOTES:
- If you install StorageX Metadata Service components on more than one server, you must install the Redistributable Package on each server.
- Ensure you install the Redistributable Package for Visual Studio 2013 and not a different version of Visual Studio. Later versions of the Redistributable Package are not supported.

You can download the redistributable package from the following location:
Web browser requirements

Because the StorageX Management Portal is Web-based, users must use the latest version of one of the following Web browsers to access the product:

- Mozilla Firefox
- Google Chrome
- Microsoft Edge
- Microsoft Internet Explorer

NOTE
If using Microsoft Internet Explorer, the StorageX Management Portal only supports Microsoft Internet Explorer 11.

Installing and configuring an Elasticsearch cluster

To use StorageX, including the StorageX Management Portal and StorageX Metadata Service, to scan and analyze file data in your environment, you must install an Elasticsearch cluster.

NOTES:

- We recommend using a minimum of 3 separate nodes for an Elasticsearch cluster.
- The installation steps below are not meant to be comprehensive, and Data Dynamics, Inc. Support does not support the Elasticsearch product. For complete information about the product, refer to the Elasticsearch documentation or contact Elastic support.
- StorageX requires Elasticsearch version 6.4.2, 6.5.x, 6.6.x, or 6.7.x. StorageX does not support Elasticsearch version 7.0.x.

The following Web pages provide detailed information about installing and configuring Elasticsearch:

- https://www.elastic.co/downloads/elasticsearch

To install and configure a new Elasticsearch cluster:

1. If you want to install an Elasticsearch cluster node on a Windows server, complete the following steps:
   a. Log on to a Windows server where you want to install a node of your Elasticsearch cluster, using an account with administrator-level access.
   b. Ensure the server has Java 8 installed.
   c. Ensure that all necessary ports are open in the server firewall.
   d. Download the most recent Windows version of Elasticsearch from https://www.elastic.co/downloads/elasticsearch.

NOTE
Elastic recommends using the MSI installation package when installing on Windows servers.
Installing and configuring an Elasticsearch cluster

e. Navigate to the location where you saved the Elasticsearch installation package.
f. Double-click the Elasticsearch .msi file.
g. If prompted to run the file, click Run or Yes to confirm.
h. If you want to install Elasticsearch in the default locations specified, select Use default directories.
i. If you want to install Elasticsearch in specific locations on your server, select Use a custom installation directory, then specify the directory location or locations you want to use.
j. Click Next.
k. If you want to install Elasticsearch as a service, select Install as service and provide account information and other details.
l. If you do not want to install Elasticsearch as a service, select Do not install as a service (start manually when needed).
m. Specify the account you want to run Elasticsearch.
n. Click Next.
o. In the Network host field, ensure you specify the fully-qualified domain name or IP address for the Elasticsearch node.

**NOTE**
If you do not specify the domain name or IP address for the node, you will not be able to access the Elasticsearch remotely from the StorageX server.

p. Specify other configuration options for your Elasticsearch node, as necessary, and then click Next.

**NOTE**
If installing the first node in your cluster, ensure you use the same cluster name for all subsequent node installations. For detailed information about configuration options, see the Elasticsearch documentation.

q. Select any Elasticsearch plugins you want to install, then click Next.

**NOTE**
For information about Elasticsearch plugins, see the Elasticsearch documentation.

r. Select the Basic license level, then click Install.
s. If prompted to allow the Elasticsearch installation package to proceed, click Yes.
t. When the installation program finishes, click Open Elasticsearch in the browser to automatically launch the Elasticsearch status page to confirm that Elasticsearch has been installed correctly.

2. If you want to install an Elasticsearch cluster node on a Linux server, complete the following steps:
a. Log on to a Linux server where you want to install a node of your Elasticsearch cluster, using an account with root- or administrator-level access.
b. Ensure the server has Java 8 installed and that all applications are up-to-date.
c. Ensure that all necessary ports are open in the server firewall.

d. Download the most recent Linux RPM version of Elasticsearch from https://www.elastic.co/downloads/elasticsearch.

**NOTE**
Elastic recommends using the MSI installation package when installing on Windows servers.

e. Open a terminal and run the following commands to verify your Java installation:

   ```
   java -version
echo $JAVA_HOME
   ```

f. In the terminal, navigate to the location where you saved the Elasticsearch installation package.

g. In the terminal, run the following commands:

   ```
   rpm -Uvh elasticsearch-X.X.X.rpm
   sudo systemctl start elasticsearch.service
   ```

h. To confirm that Elasticsearch has been installed correctly, open a browser and navigate to http://localhost:9200.

3. Repeat steps step 1 through step 2 for each server where you want to install an Elasticsearch cluster node.

---

**Installing and configuring a Redis cluster**

StorageX uses the third-party Redis product to enable communication between the different components of the StorageX Metadata Service.

We recommend installing a stand-alone Redis cluster on multiple Linux-based servers. We recommend using a minimum of 3 separate nodes for a Redis cluster. The following instructions are for setting up a 3-node Redis cluster with two Redis instances running on each node, using the following configuration:

**Server 1:**
- Master Node A, Port 6379
- Slave Node C, Port 6381

**Server 2:**
- Master Node B, Port 6380
- Slave Node A, Port 6379

**Server 3:**
- Master Node C, Port 6381
- Slave Node B, Port 6380

In this configuration, each node has one master node instance and one slave node instance, with the slave node instance used for replication from one of the other three nodes.
StorageX requires Redis version 3.0 or later. The steps below are tailored to Redis version 5.0.

We do not recommend using any version of Redis Enterprise with StorageX.

We recommend installing Redis on a Linux server running the most recent version of Ubuntu, and the following instructions are specific to Ubuntu. For information about installing Redis on other Linux distributions, see the Redis documentation.

We do not recommend installing Redis on a Windows server in a production environment. You can only deploy a one-node Redis cluster on a single Windows server, and this type of installation does not provide sufficient scalability or redundancy. We recommend installing all Redis cluster nodes on Linux-based servers.

The installation steps below are not meant to be comprehensive, and Data Dynamics, Inc. Support does not support the Redis product. For complete information about the product, refer to the Redis documentation.

Each Linux server must be configured to allow access to all necessary ports, as well as a port higher than 10000. For example, on the first server, both port 6379 and 16379 must be open.

The following Web pages provide detailed information about installing and configuring Redis:

- https://redis.io/download
- https://redis.io/topics/quickstart
- https://redis.io/topics/cluster-tutorial
- https://codeflex.co/configuring-redis-cluster-on-linux/

To install and configure a new Redis cluster on Ubuntu Linux:

1. Log on to a Linux server where you want to install a node of your Redis cluster, using an account with root- or administrator-level access.

2. Open a terminal and run the following commands to install prerequisites necessary for Redis:
   ```
   sudo add-apt-repository "deb http://us.archive.ubuntu.com/ubuntu/ bionic universe"
   sudo apt-get update && sudo apt-get upgrade
   sudo apt install make gcc libc6-dev tcl
   ```

3. Run the following commands to download and un-tar the most recent stable version of Redis:
   ```
   wget http://download.redis.io/redis-stable.tar.gz
   tar xvzf redis-stable.tar.gz
   ```

4. Navigate to the redis-stable directory and run the following commands to install Redis and then test that Redis was installed correctly:
   ```
   sudo make install
   sudo make test
   ```
5. In the redis-stable directory, run the following commands to create separate master and slave configuration files:
   
   ```
   cp redis.conf c_slave.conf
   mv redis.conf a_master.conf
   ```

6. Open the `a_master.conf` file using a text editor and comment the following line:
   ```
   # bind 127.0.0.1
   ```

7. Uncomment the following lines in the `a_master.conf` file:
   ```
   cluster-enabled yes
   cluster-config-file nodes-6379.conf
   cluster-node-timeout 15000
   ```

8. In the `a_master.conf` file, modify the `protected-mode` option as follows:
   ```
   protected-mode no
   port 6379
   pidfile /var/run/redis_6379.pid
   ```

9. Save and close the `a_master.conf` file.

10. Open the `c_slave.conf` file using a text editor and comment the following line:
    ```
    # bind 127.0.0.1
    ```

11. Uncomment the following lines in the `c_slave.conf` file:
    ```
    cluster-enabled yes
    cluster-node-timeout 15000
    ```

12. In the `c_slave.conf` file, modify the following options:
    ```
    protected-mode no
    port 6381
    pidfile /var/run/redis_6381.pid
    cluster-config-file nodes-6381.conf
    ```

13. Save and close the `c_slave.conf` file.

14. Repeat step 1 through step 13 on each Linux server you want to use as a Redis cluster node, replacing the port numbers (6379, 6381) and letter designations (`a_master.conf`, `c_slave.conf`) as necessary for the configuration outlined above.

For example, on the second server, the master file would be `b_master.conf` and the port number for the master instance would be 6380, while the slave configuration file would be `a_slave.conf`, with the port number 6379.

15. On each Redis node server, run the following commands:
    ```
    redis-server redis-stable/X_master.conf
    redis-server redis-stable/Y_slave.conf
    ```

Where `X_master.conf` is the configuration file for the master node instance and `Y_slave.conf` is the slave node instance. For example, the file names could be `a_master.conf` and `c_slave.conf`. 
NOTE
You may need to start two separate terminals or terminal tabs to run both Redis instances simultaneously.

16. Verify that each Redis node is running in cluster mode and displays the correct port number.

17. On the first Redis node server, open a new terminal or terminal tab, and run the following command to create the actual Redis cluster:

```
redis-cli --cluster create SERVER1:MASTERPORT1 SERVER1:SLAVEPORT1
SERVER2:MASTERPORT2 SERVER2:SLAVEPORT2 SERVER3:MASTERPORT3 SERVER3:SLAVEPORT3
--cluster-replicas 1
```

Where `SERVER1`, `SERVER2`, and `SERVER3` are the respective IP addresses of each Redis node server, `MASTERPORT1`, `MASTERPORT2`, and `MASTERPORT3` are the respective port numbers for each Redis master instance, and `SLAVEPORT1`, `SLAVEPORT2`, and `SLAVEPORT3` are the respective port numbers for each Redis slave instance.

18. Enter yes to accept the suggested configuration.

19. To verify the Redis cluster is functional, run the following commands:

```
redis-cli -c -h SERVER1 -p MASTERPORT1
SERVER1:MASTERPORT1> cluster nodes
```

Where `SERVER1` is the IP address of the first Redis node server and `MASTERPORT1` is the port number for the master instance on that node.

*If the cluster was created successfully*, the `cluster nodes` command in redis-cli should display a list with six Redis cluster nodes, three of which are listed as master nodes and three of which are listed as slave nodes.

20. If you want to test that the cluster is functioning, complete the following steps:

a. On the master instance of the first Redis node, enter the following command in redis-cli:

```
SERVER1:MASTERPORT1> set test 'this key was defined on master A'
```

b. On the third Redis node server, run the following commands:

```
redis-cli -c -h SERVER3 -p MASTERPORT3
SERVER3:MASTERPORT3> get test
```

The Redis node should return the following response:

"this key was defined on master A"

Installing and configuring an Apache Cassandra cluster for index backup

If you want to use StorageX to archive file data to an object storage resource or perform workload tiering to move file data to a different file storage resource, we recommend installing and configuring an Apache Cassandra cluster, in addition to your Elasticsearch cluster.

StorageX uses Cassandra as a backup for all indices created by Elasticsearch during the scan process. If your Elasticsearch indices should become corrupted or deleted, you can rebuild those indices from the Cassandra backup.
NOTES:

- You must use a minimum of at least 3 separate nodes for a Cassandra cluster.
- The installation steps below are not meant to be comprehensive, and Data Dynamics, Inc. Support does not support the Apache Cassandra product. For complete information about the product, refer to the Cassandra documentation or contact the Cassandra developer community.

Each Cassandra node in the cluster needs a configuration YAML file (cassandra.yaml) and location properties file (cassandra-rackdc.properties).

When you install Cassandra, the installation package creates a default set of cassandra.yaml and cassandra-rackdc.properties files. You can use the StorageX YAML Assistant to create your own customized set of files for your Cassandra cluster. If you want to customize your Cassandra installation, you can modify additional settings within the cassandra.yaml file.

The following Web pages provide detailed information about installing and configuring Apache Cassandra:

- http://cassandra.apache.org/download/

To install and configure a new Apache Cassandra cluster:

1. If you want to use the StorageX YAML Assistant to create configuration files for your Cassandra cluster, complete the following steps before installing Cassandra:
   a. Log on to the StorageX server computer using an user account that is a member of the local Administrators group on the computer.
   b. Navigate to the Tools folder on your StorageX server, located in the \Program Files\Data Dynamics\StorageX folder by default.
   c. Run the YamlAssistant.exe application.
   d. Click When you have all the information gathered and ready, click here to start entering the information.
   e. In the Cassandra Cluster Name field, specify the name you want to use for the new Cassandra cluster.
   f. Click the Add (+) icon to add a data center.
   g. Select the data center in the data center list.
   h. In the Data Center Name field, specify the name you want to use for the data center.
   i. Click the Add (+) icon to add a new server/node.
   j. Provide the name, rack name, and IP address for the new node.
   k. If the new server uses an SSD, select Storage is SSD-based.
   l. If the new server is a Windows server, select This server is a Windows machine.
NOTE
By default, Cassandra nodes are Linux servers. We recommend installing your Cassandra cluster nodes on Linux servers rather than on Windows servers.

m. When finished adding all data centers and nodes, click Browse in the Output Directory field and specify the folder where you want to save all YAML configuration files and location files for your Cassandra cluster.

n. Click Export, then click OK.

o. Click Close to close the StorageX YAML Assistant.

p. Copy the files for each new node to the respective server where you want to deploy Cassandra.

2. Log on to a Linux server where you want to install a node of your Apache Cassandra cluster.

NOTE
Install any Cassandra nodes you want to use as seed nodes first, followed by other, non-seed node installations.

3. Ensure the server has both Java 8 and Python 2.7 installed and that all applications are up-to-date.

4. Ensure that all necessary ports are open in the server firewall.

5. If you want to install a Cassandra RPM package using yum, complete the following steps:
   a. Using a text editor, create a new cassandra.repo file in the following folder:
      /etc/yum.repos.d/cassandra.repo
   b. Add the following lines to the new cassandra.repo file:
      
      [cassandra]
      name=Apache Cassandra
      baseurl=https://www.apache.org/dist/cassandra/redhat/311x/
      gpgcheck=1
      repo_gpgcheck=1
      gpgkey=https://www.apache.org/dist/cassandra/KEYS
   c. Open a terminal and run the following command:
      sudo yum install cassandra
   d. If you receive a prompt about GPG keys or downloads, click y to confirm.
   e. Run the following command to clear out any default Cassandra data:
      sudo rm -rf /var/lib/cassandra/data/system/*
   f. If you used the StorageX YAML Assistant to create custom cassandra.yaml and cassandra-rackdc.properties files, copy those custom files to the /etc/cassandra/default.conf folder on the server, overwriting the default files.
   g. If you want to configure your cassandra.yaml and cassandra-rackdc.properties files manually, make edits as necessary to the files located in the /etc/cassandra/default.conf folder and save the modified files.
   h. Run the following command to start Cassandra:
Installing and configuring an Apache Cassandra cluster for index backup

service cassandra start

1. Run the following command to set Cassandra to start automatically when the server is rebooted:
   
   ```
   chkconfig cassandra on
   ```

6. If you want to install Cassandra from a binary tarball file, complete the following steps:

   a. Download the most recent version of Apache Cassandra from http://cassandra.apache.org/download/.

   b. Untar the downloaded Cassandra installation tarball file to a folder on your server where you want to install the application.

   c. Run the following command to clear out any default Cassandra data:
      
      ```
      sudo rm -rf /var/lib/cassandra/data/system/*
      ```

   d. If you used the StorageX YAML Assistant to create custom cassandra.yaml and cassandra-rackdc.properties files, copy those custom files to the /conf folder in your Cassandra installation folder, overwriting the default files.

   e. If you want to configure your cassandra.yaml and cassandra-rackdc.properties files manually, make edits as necessary to the files located in the /conf folder in your Cassandra installation folder and save the modified files.

   f. Open a terminal and navigate to the Cassandra bin folder.

      **NOTE**
      
      If you want to make the Cassandra bin folder more easily accessible on your server, add the folder to your path.

   g. Run the following command to start Cassandra in the foreground:
      
      ```
      cassandra -f
      ```

7. In the terminal, run the following command to verify that Cassandra is active:

   ```
   nodetool status
   ```

   The nodetool should display the new node, as well as any existing nodes, with a status of UN. This indicates that the node is up and in a Normal state. When multiple nodes in a specific data center are installed and running, all nodes should be displayed.

8. Repeat steps step 2 through step 7 for each server where you want to install a Cassandra cluster node.

9. After you finish installing all Cassandra nodes, create one or more keyspaces, as necessary for your cluster.

   A keyspace is the top-level namespace and CQL data structure of an Apache Cassandra cluster. This structure is similar to a SQL database file, and it defines how nodes in your cluster replicate data. Typically, a cluster will only have a single keyspace, but some Cassandra installations can have multiple keyspaces.

   You can create a keyspace on any Cassandra node by using the CQL shell interactive terminal, cqlsh, and the command **CREATE KEYSPACE**.
Installing StorageX Management Portal and Metadata Service components

The following Web pages provide detailed information about using cqlsh to create keyspaces in Cassandra:

- https://www.tutorialspoint.com/cassandra/cassandra_create_keyspace.htm

Alternatively, you can use a tool like DataStax DevCenter to run the CQL queries needed to create a keyspace on your cluster.

For more information about DataStax DevCenter, see http://docs.datastax.com/en/dse/5.1/dse-dev/datastax_enterprise/devcenter/dcToc.html.

The StorageX Management Portal is installed by default with the overall Data Dynamics StorageX product, using the StorageX Setup Wizard. For detailed information about installing StorageX, see the StorageX Administrator’s Guide.

NOTE

You cannot install the StorageX Management Portal by itself. The Management Portal must be on the StorageX server computer.

In addition to the main StorageX package, you must also install the Metadata Service components for the Management Portal to function correctly. The Metadata Service allows the StorageX server to collect file, share, and export data and store information about that data in the Elasticsearch repository and in your Apache Cassandra database, if applicable.

Before installing the StorageX Metadata Service components, you must install Elasticsearch and Redis. We recommend installing Elasticsearch and Redis each in a separate, clustered configuration for better performance, scalability, and redundancy.

For information about installing and configuring an Elasticsearch cluster, see “Installing and configuring an Elasticsearch cluster” on page 15. For information about installing and configuring a Redis cluster, see “Installing and configuring a Redis cluster” on page 17.

If using StorageX for archival purposes, we recommend you also install a Cassandra cluster before installing the Metadata Service, to act as a backup for indices created by Elasticsearch. For information about installing and configuring a Cassandra cluster, see “Installing and configuring an Apache Cassandra cluster for index backup” on page 20.

You can select one or more of the following StorageX Metadata Service components to install on a particular server:

StorageX Metadata Service Web Service (API)

The Web Service component receives metadata information from the data engine and responds to queries for metadata information. This component is a REST API service.

StorageX Metadata Service Processing Engine

The Processing Engine component processes metadata information and inserts that metadata into the search engine StorageX uses to find matching metadata for a query.
In relatively small-scale environments, you may want to install both StorageX Metadata Service components on the same computer. However, in larger-scale environments, you may want to separate out the different components on different servers for performance reasons.

If you want to install StorageX Metadata Service components separately, either one at a time on the same server or on different servers, you must install the components in the following order:

1. Elasticsearch cluster
2. Redis cluster
3. Cassandra cluster (optional but recommended)
4. StorageX Metadata Service Processing Engine
5. StorageX Metadata Service Web Service (API)

For specific hardware recommendations and requirements for these components, see “StorageX Metadata Service components” on page 11.

If you want to install StorageX Metadata Service components on a computer where only version 1.2 of TLS is enabled, the installation program may display an error.

**NOTE**
This issue only applies to computers with Microsoft Windows Server 2016 or later installed.

To work around this issue, Microsoft recommends modifying .NET Framework registry keys on the computer where you want to install the Metadata Service to allow the computer to use the strongest version of TLS that is available.

**ATTENTION**
Before making changes to a Windows registry key, ensure you back up or export the specific key, in case you need to later revert to the original.

Add the specified values to the following registry keys either manually using the Registry Editor or by creating and running or importing the text below as a .reg file:

Windows Registry Editor Version 5.00

[HKEY_LOCAL_MACHINE\SOFTWARE\WOW6432Node\Microsoft\.NETFramework\v4.0.30319]
"SystemDefaultTlsVersions"=dword:00000001
"SchUseStrongCrypto"=dword:00000001

[HKEY_LOCAL_MACHINE\SOFTWARE\Microsoft\.NETFramework\v4.0.30319]
"SystemDefaultTlsVersions"=dword:00000001
"SchUseStrongCrypto"=dword:00000001

When finished, restart the computer, open Windows PowerShell using an account that has Administrator access, and enter the following command:

```powershell
[Net.ServicePointManager]::SecurityProtocol
```

PowerShell should display `[SystemDefault]` in the output.

For more information about configuring TLS in Microsoft Windows environments, see https://docs.microsoft.com/en-us/dotnet/framework/network-programming/tls.
To install StorageX Management Portal components:

1. Log on to the StorageX server computer using an user account that is a member of the local Administrators group on the computer.
2. Open a command prompt, running the tool as an administrator.
3. Navigate to the location where you saved the StorageX installation package.
4. Within the installation package, navigate to the MetadataService\Windows folder.
5. Enter the MDSSetup.msi command.
6. On the Welcome to the StorageX Metadata Service Setup Wizard dialog box, review the welcome message and then click Next.
7. On the End-User License Agreement dialog box, review the terms of the license agreement, select the I accept the terms in the License Agreement check box, and then click Next.
8. On the Custom Setup dialog box, select which StorageX Metadata Service components you want to install by completing the following steps:
   a. If you want to install both StorageX Metadata Service components on the same computer, select all components. Typically, both components are selected by default.
   b. If you want to install only one StorageX Metadata Service component on the computer, deselect the other StorageX Metadata Service component by selecting Entire feature will be unavailable on the drop-down box.
   c. If you want to change the default location where the setup program installs files, click Browse to browse to and select a different folder. By default, the StorageX Setup program installs StorageX in the \Program Files\Data Dynamics\StorageX\MetadataService folder.
   d. If you want to see the total disk space required for all selected components, click Disk Usage. Click OK when finished.
9. Click Next.
10. Specify the StorageX server in the StorageX server field or click Browse to browse to and select the StorageX server computer that you want the StorageX Metadata Service to connect to.
11. Specify the port you want the StorageX Metadata Service to use. By default, the Metadata Service uses port 9463.
12. Click Next.
13. In the Redis Location dialog box, specify the name you want to use for the Redis cluster in the Cluster name field. This Redis cluster will be used as the communication channel for the Metadata Service.
14. In the Redis nodes/servers field, specify all nodes or servers in the Redis cluster you want to use, in a comma-separated list. You can include the fully-qualified domain name or IP address for each node.
15. In the Login and Password fields, specify the credentials used to connect to the Redis cluster, if applicable.

16. If your Redis installation requires SSL, click This cluster requires SSL.

17. Click Next.

18. On the Specify SSL Certificate dialog box, specify the certificate you want the StorageX Metadata Service to use to make SSL connections. You can either specify that you want StorageX to use the default certificate installed with the product or use a custom certificate generated outside of StorageX. Complete the following steps:
   a. If you want to use the default certificate, select Use self-signed certificate.
   b. If you want to use a custom certificate, select Use the following certificate, then select the certificate you want to use from the list of certificates displayed.
   
   **NOTE**
   The custom certificate must be valid, must not have expired, and must be stored locally. If a certificate is not installed locally, the certificate will not be displayed in the list.
   c. After you select an option, click Next.
   d. If prompted to continue, click Yes.

19. On the Service Log On Credentials dialog box, specify the login credentials for the StorageX Metadata Service service account. In the Account field, type the name of the domain user account you want to use for the StorageX Metadata Service service account, and in the Password field, type the password for the account.

20. Click Next.

21. In the Elasticsearch Location dialog box, specify all nodes for the Elasticsearch cluster in a comma-separated list in the Node contact URLs text box. You can include the fully-qualified domain name or IP address for each node.

22. Click Next.

23. If you want to use an Apache Cassandra cluster as a backup for Elasticsearch indices, complete the following steps in the Apache Cassandra Location dialog box:
   a. Select Use an Apache Cassandra cluster for archiving data with the StorageX Metadata Service installation.
   b. Specify the name of the cluster in the Cluster name field.
   c. In the Keyspace field, specify the name of the keyspace you want to use for the Apache Cassandra cluster. The keyspace must already be defined for the cluster. For more information about creating keyspaces for a Cassandra cluster, see “Installing and configuring an Apache Cassandra cluster for index backup” on page 20.
   d. In the Contact hosts field, specify all nodes for the Apache Cassandra cluster in a comma-separated list. You can include the fully-qualified domain name or IP address for each node.
   e. In the Login and Password fields, specify the credentials used to connect to the Apache Cassandra cluster, if applicable.
   f. If necessary for your Cassandra installation, modify other Apache Cassandra options, including the client port, SSL configuration, and compression type.
Preparing to use the StorageX Management Portal

Before using the StorageX Management Portal, you must first set up your StorageX environment. If you do not correctly set up all of your resources, universal data engines, and permissions before attempting to work in the Management Portal, you will not be able to successfully scan, analyze, or archive your data.

To set up StorageX for the Management Portal:

1. Log on to the StorageX Console computer using a user account that has Administrator access to your StorageX environment.
2. Click the Storage Resources tab.
3. In the Summary pane, click the File Storage Resources tab.
4. Ensure that all file storage resources you want to scan, analyze, or archive using the Management Portal are displayed in the list and are all properly configured, with all necessary validation checks in the Validated state.

NOTE
If StorageX cannot access a particular resource and display that resource in the StorageX Console, the Management Portal will also not be able to display that resource.

5. In the Summary pane, click the Object Storage Resources tab.
6. Ensure that all object storage resources you want to use as a destination for archival in the Management Portal are displayed in the list and are all properly configured, with all necessary validation checks in the Validated state.

NOTE
If StorageX cannot access a particular resource and display that resource in the StorageX Console, the Management Portal will not be able to archive data to that resource.

7. Click the Control Panel tab.
8. Click Deployed Data Engines.
9. In the right pane, select each universal data engine you want StorageX to use to scan, analyze, or archive data in the Management Portal and click Manage Credentials.
10. Verify that each universal data engine has sufficient permissions. If a universal data engine does not have access to the file storage resources you want to manage, modify the existing credentials.
11. Click OK to close.
12. Click File > Options.

13. Click Access Control.


15. Add each user you want to have access to the Management Portal and ensure that each has the Read Only access type selected, at a minimum. StorageX Administrator users have access to the Management Portal by default.

   For more information about role-based access control in StorageX, see the StorageX Administrator’s Guide.

16. Click OK.

17. Close the StorageX Console.

Accessing the StorageX Management Portal

You access the StorageX Management Portal using a Web browser.

**NOTE**

*If the browser displays a pop-up dialog box, do not select the Prevent this page from creating additional dialogs. checkbox before clicking OK. If you disable the browser’s pop-up dialog box capability, you cannot delete analysis groups, paths, or scans, and you will not be notified with any further message dialogs.*

To access the StorageX Management Portal:

1. Open a Web browser.

2. *If you use Microsoft Internet Explorer, complete the following steps to enable the browser to correctly display the StorageX Web interface:*
   a. In Internet Explorer, click Tools > Compatibility View settings.
   b. Clear Display intranet sites in Compatibility View and click Close.

3. *If you use a browser that automatically blocks pop-ups or an external pop-up blocker, disable any pop-up blocking functionality before accessing the Management Portal.*

   **NOTE**

   *If you do not disable all pop-up blockers, you cannot access the StorageX Retrieval Portal directly from the Management Portal.*

4. In the browser, navigate to the following URL:

   https://servername:9777/

   *Where servername is the name of the StorageX server where you installed and started the StorageX Management Portal.*

   **NOTE**

   *If the StorageX Management Portal server is attached to a domain, ensure you specify the full server name, including the domain. For example:*

   https://servername.domain.domain:9777/*
5. If you see an “Insecure Connection” warning while using Mozilla Firefox to access the StorageX Management Portal, complete the following steps:
   a. In the Web browser, open a new window or tab and enter about:config.
   b. If prompted by the browser, click to accept.
   c. In the list of preferences, find the security.enterprise_roots.enabled preference.
   d. If the preference does not exist, right-click the list, select New > Boolean, specify the name security.enterprise_roots.enabled, and click OK.
   e. Double-click the preference to set the value to true.
   f. Close the window or tab.
   g. Retry the StorageX Management Portal link.
6. When prompted, provide the user name and password you want to use to access the Management Portal, then click Log In.

**NOTE**
The user account must have at least Read Only access to the StorageX product as a whole to be able to access the Management Portal. For more information, see “Preparing to use the StorageX Management Portal” on page 28.

### Configuring monitoring for StorageX Management Portal components

After you install the StorageX Management Portal, you can set up one or more third-party tools to monitor the health of the Cassandra database cluster. We recommend installing the Prometheus, Grafana, and Reaper tools to enable monitoring of your Cassandra cluster.

**NOTE**
These tools are third-party products and are not supported by the Data Dynamics, Inc. Technical Support team. If you encounter issues while installing or using Prometheus, Grafana, or Reaper, please contact the product owner for the product.

Prometheus listens to a stream of metrics that Cassandra sends out and parses those metrics into a usable format.

Grafana lets you add Prometheus as a data source, channeling those Cassandra metrics, and allows you to see detailed graphics and charts of that data.

Reaper is a tool for repairing your Cassandra cluster if you run into any issues, and for grooming old, unnecessary data from the cluster. Reaper allows you to easily view the status of your Cassandra cluster and ensure all cluster nodes are healthy.

For more information about Prometheus, Grafana, and Reaper, see their respective Web sites:
- https://prometheus.io/
- https://grafana.com/
- http://cassandra-reaper.io/

When these tools have been installed on the appropriate servers in your environment, you can view the status of each tool and the most recent results returned by each tool, if applicable, within the StorageX Management Portal.
Configure the endpoints for Prometheus, Grafana, and Reaper in the Maintenance Portal area of the Management Portal.

To configure the StorageX Management Portal to access Prometheus, Grafana, and Reaper:

1. In a browser, open the StorageX Management Portal Web interface.
2. Click the Maintenance Portal tab.
3. If you want to configure the Management Portal to access Prometheus, complete the following steps:
   a. Click Prometheus, then click Configure.
   b. In the Endpoint Host Name field, specify the name of the server where Prometheus is installed.
   c. If Prometheus is installed on the local Management Portal server, enter the following: localhost

   **NOTE**
   You can only specify English characters from A to Z, dots (.), dashes (-), or digits (1-9).
   d. In the Endpoint Port Number field, specify the port number used to access Prometheus. By default, the port number value is 9090.
   e. Click Save.

4. If you want to configure the Management Portal to access Grafana, complete the following steps:
   a. Click Grafana, then click Configure.
   b. In the Endpoint Host Name field, specify the name of the server where Grafana is installed.
   c. If Grafana is installed on the local Management Portal server, enter the following: localhost

   **NOTE**
   You can only specify English characters from A to Z, dots (.), dashes (-), or digits (1-9).
   d. In the Endpoint Port Number field, specify the port number used to access Prometheus. By default, the port number value is 3000.
   e. Click Save.

5. If you want to configure the Management Portal to access Reaper, complete the following steps:
   a. Click Reaper, then click Configure.
   b. In the Endpoint Host Name field, specify the name of the server where Reaper is installed.

   **NOTE**
   You can only specify English characters from A to Z, dots (.), dashes (-), or digits (1-9).
   c. In the Endpoint Port Number field, specify the port number used to access Reaper. By default, the port number value is 8080.
Upgrading the StorageX Management Portal

When you upgrade your StorageX installation as a whole, you must also upgrade your StorageX Management Portal installation.

For information on upgrading the StorageX Retrieval Portal, see “Upgrading the StorageX Retrieval Portal” on page 62.

NOTES:

• After you upgrade StorageX, ensure you clear your browser cache before reopening the StorageX Management Portal.
• As part of the upgrade process from StorageX 8.0 to StorageX 8.1, you must uninstall the existing StorageX Metadata Service, reboot the server, and reinstall the Metadata Service components on the server using the new version of the installation package.
• You must also run the included Repository Upgrade Tool during the upgrade process. This tool will automatically move any data in your existing Cassandra installation into the new structure created during the StorageX 8.1 upgrade.

To upgrade an existing StorageX Management Portal installation:

1. Before upgrading the StorageX Management Portal and StorageX Retrieval Portal, first upgrade the StorageX server and database. For information about upgrading the StorageX server, see the StorageX Administrator’s Guide.

2. Log on to the StorageX server where the StorageX Management Portal is installed using a user account that is a member of the local Administrators group on the computer.

3. Close any running browsers or instances of the StorageX Console.

4. Open a command prompt, running the tool as an administrator.

5. In the command prompt, navigate to the location of the Repository Upgrade Tool (RepositoryUpgrade). By default, this file is located in the following folder:

   C:\Program Files\Data Dynamics\StorageX\Tools\RepositoryUpgrade

6. Run the following command:

   RepositoryUpgrade.exe

   NOTE
   If you do not run the Repository Upgrade Tool, you cannot reuse your existing Cassandra keyspace for StorageX 8.1.

7. When the tool finishes running, open the Windows Control Panel.

8. In the Control Panel, click Programs > Uninstall a program.

9. Select StorageX Metadata Service, and then click Uninstall.

10. If a confirmation dialog box displays asking if you are sure you want to uninstall the StorageX Metadata Service, click Yes, then click Yes again.
11. When the Metadata Service is completely uninstalled, reboot the server.

12. After the server finishes restarting, log back on.

13. Reopen a command prompt, running the tool as an administrator.

14. In the command prompt, navigate to the location where you saved the new StorageX installation package.

15. Within the installation package, navigate to the MetadataService\Windows folder.

16. If you want to use your existing Cassandra installation as a backup for indices created by Elasticsearch, enter the following command:

   MDSSetup.msi

17. If you do not want to use your existing Cassandra installation, enter the following command:

   MDSSetup.msi ReuseCassandraKeyspace=1

18. On the Welcome to the StorageX Metadata Service Setup Wizard dialog box, review the welcome message and then click Next.

19. On the End-User License Agreement dialog box, review the terms of the license agreement, select the I accept the terms in the License Agreement check box, and then click Next.

20. On the Custom Setup dialog box, select which StorageX Metadata Service components you want to install by completing the following steps:

   a. If you want to install both StorageX Metadata Service components on the same computer, select all components. Typically, both components are selected by default.

   b. If you want to install only one StorageX Metadata Service component on the computer, deselect the other StorageX Metadata Service component by selecting Entire feature will be unavailable on the drop-down box.

   **NOTE**

   If you want to install StorageX Metadata Service components on different computers, first install the StorageX Metadata Service Processing Engine and then the StorageX Metadata Service Web Service.

   c. If you want to change the default location where the setup program installs files, click Browse to browse to and select a different folder. By default, the StorageX Setup program installs StorageX in the \Program Files\Data Dynamics\StorageX\MetadataService folder.

   d. If you want to see the total disk space required for all selected components, click Disk Usage. Click OK when finished.

21. Click Next.

22. Specify the StorageX server in the StorageX server field or click Browse to browse to and select the StorageX server computer that you want the StorageX Metadata Service to connect to.

23. Specify the port you want the StorageX Metadata Service to use. By default, the Metadata Service uses port 9463.

24. Click Next.
25. In the Redis Location dialog box, specify the name you want to use for the Redis cluster in the Cluster name field. This Redis cluster will be used as the communication channel for the Metadata Service.

For more information about installing your Redis cluster, see “Installing and configuring a Redis cluster” on page 17.

26. In the Redis nodes/servers field, specify all nodes or servers in the Redis cluster you want to use, in a comma-separated list. You can include the fully-qualified domain name or IP address for each node.

27. In the Login and Password fields, specify the credentials used to connect to the Redis cluster, if applicable.

28. If your Redis installation requires SSL, click This cluster requires SSL.

29. Click Next.

30. On the Specify SSL Certificate dialog box, specify the certificate you want the StorageX Metadata Service to use to make SSL connections. You can either specify that you want StorageX to use the default certificate installed with the product or use a custom certificate generated outside of StorageX. Complete the following steps:
   a. If you want to use the default certificate, select Use self-signed certificate.
   b. If you want to use a custom certificate, select Use the following certificate, then select the certificate you want to use from the list of certificates displayed.

   **NOTE**
   The custom certificate must be valid, must not have expired, and must be stored locally. If a certificate is not installed locally, the certificate will not be displayed in the list.

   c. After you select an option, click Next.
   d. If prompted to continue, click Yes.

31. On the Service Log On Credentials dialog box, specify the login credentials for the StorageX Metadata Service service account. In the Account field, type the name of the domain user account you want to use for the StorageX Metadata Service service account, and in the Password field, type the password for the account.

32. Click Next.

33. In the Elasticsearch Location dialog box, specify all nodes for the Elasticsearch cluster in a comma-separated list in the Node contact URLs text box. You can include the fully-qualified domain name or IP address for each node.

For more information about installing your Elasticsearch cluster, see “Installing and configuring an Elasticsearch cluster” on page 15.

34. Click Next.

35. If you want to use an Apache Cassandra cluster as a backup for Elasticsearch indices, complete the following steps in the Apache Cassandra Location dialog box:
   a. Select Use an Apache Cassandra cluster for archiving data with the StorageX Metadata Service installation.
   b. Specify the name of the cluster in the Cluster name field.
c. In the Keyspace field, specify the name of the keyspace you want to use for the Apache Cassandra cluster. The keyspace must already be defined for the cluster. For more information about creating keyspaces for a Cassandra cluster, see “Installing and configuring an Apache Cassandra cluster for index backup” on page 20.

**NOTE**

If you want to use your previous Cassandra installation, ensure you specify the same cluster name, keyspace name, and contact hosts used in the StorageX 8.0 installation.

d. In the Contact hosts field, specify all nodes for the Apache Cassandra cluster in a comma-separated list. You can include the fully-qualified domain name or IP address for each node.

e. In the Login and Password fields, specify the credentials used to connect to the Apache Cassandra cluster, if applicable.

f. If necessary for your Cassandra installation, modify other Apache Cassandra options, including the client port, SSL configuration, and compression type.

36. Click Next.

37. On the Ready to install StorageX Metadata Service dialog box, click Install to begin the installation.

38. On the Completed the StorageX Metadata Service Setup Wizard dialog box, click Finish to close the setup program.

39. Close the command prompt.

40. In a browser, open the StorageX Management Portal Web interface.

---

**Uninstalling the StorageX Management Portal**

*If you want to remove the StorageX Management Portal from an installed computer, we recommend using the Windows Control Panel to uninstall the product.*

**NOTES:**

- Before you uninstall, we recommend you ensure no scans are currently running in your StorageX environment.
- When uninstalling StorageX as a whole, first uninstall all instances of the StorageX Retrieval Portal, then uninstall all instances of the StorageX Metadata Service and StorageX Management Portal, and then uninstall all other StorageX components. For more information about uninstalling the Retrieval Portal, see “Uninstalling the StorageX Retrieval Portal” on page 63. For more information about uninstalling other StorageX components, see the *StorageX Administrator’s Guide*.

**To uninstall the StorageX Management Portal:**

1. Log on to a computer where the StorageX Management Portal is installed.

2. *If the Data Dynamics StorageX product is also installed on the computer where the StorageX Management Portal is installed*, close any instances of the StorageX Console on the local computer before uninstalling StorageX.

3. Open the Windows Control Panel.
4. Click Add or Remove Programs or Programs > Uninstall a program, depending on the version of Windows installed on the computer.
5. Select StorageX.
6. Click Uninstall.
7. Click Yes to confirm.
8. If a Windows User Account Control dialog box displays, click Yes to confirm you want to run the uninstallation program.
9. When finished, close the Programs and Features dialog box.
Using the StorageX Management Portal

This section explains how to use the StorageX Management Portal to gather information on the storage resources in your environment, analyze that data, and archive data to object storage or use workload tiering to migrate data to a different share or export.

In this chapter

- “Getting your storage resource data” on page 37
- “Analyzing your storage resource data” on page 40
- “Working with analysis results” on page 43
- “Archiving your file data to an object storage resource” on page 45
- “Using Workload Tiering to migrate your file data to a different file storage resource” on page 48
- “Viewing and modifying scans, datasets, analysis sets, and policies” on page 50
- “Deleting scans, analysis sets, and policies” on page 52
- “Configuring Cassandra grooming settings” on page 53
- “Accessing Management Portal monitoring tools” on page 54

Getting your storage resource data

In order to use the StorageX Management Portal, you first need to gather data about your existing file storage resources and the data stored by those resources.

You can do this by scanning storage resources managed by StorageX, storing data about those resources in Elasticsearch, and tagging the scanned data as necessary. If you want to delineate your data by region, for example, you can tag the data from each group of scanned resources with a custom Region tag, with the values APAC, NorthAmerica, EMEA, or SouthAmerica.

When tagging data during a scan, keep the following restrictions in mind:

- You can only use a specific tag once per scan.
- You can specify a maximum of 10 tags per scan.
- Tag names and values are case-insensitive. If you create a dataset with the custom tag Region, you cannot add a second tag with the lower-case name region.
- Tag key names can use either alphanumeric characters or underscores (_) and cannot start with a number. You cannot specify a key name that starts with a number or includes an invalid non-alphanumeric character.
• Tag values can use a combination of alphanumeric characters and specific valid non-alphanumeric characters. Valid non-alphanumeric characters you can include in a tag value are: underscore (_); hyphen (-); plus (+); equal (=); colon (:); white space ( ); and forward slash (/). You cannot specify a tag value that includes an invalid non-alphanumeric character.

• You can specify a tag name with a maximum of 128 characters, and a tag value with a maximum of 256 characters.

• You cannot specify a tag name or value that starts with a number or includes an invalid non-alphanumeric character.

If you later decide to archive your file data, StorageX pushes any custom tags you used when scanning your data to the destination object storage resource, where they are converted into object tags. This allows you to later retrieve specific archived files based on their object tags.

ATTENTION
When specifying the name for a new scan, ensure you do not reuse the name of a previously run, modified, or deleted scan. If you reuse a scan name, any analysis sets including the new scan may include data from the earlier scan.

For more information about scanning storage resources, see “Understanding gathering data” on page 2.

To scan your data:

1. In a browser, open the StorageX Management Portal Web interface.

2. Click the Discover tab, then click Create.

3. In the Create Scan dialog box, specify the name you want to use for the new scan, as well as an optional description.

   NOTE
   During analysis of your data, you can use the scan name like a tag, allowing you to only analyze data gathered from that specific set of resources.

4. Click the Storage Resources tab.

5. Select the checkbox beside the name of each resource you want to scan.

6. In the Shares or Exports column, click the drop-down menu for each selected resource and select Shares, Exports, or Both, depending on the resource and what data you want to get.

7. Click Search.

8. In the Paths to Scan tab, select each export or share you want to scan.

   NOTE
   You can also select the checkbox at the top of the list to select all shares and exports listed, or use the text boxes at the top of the Export or Share and Local Path columns to filter for specific names and paths.

9. Click the Scan Options tab.

10. If you want StorageX to automatically select a universal data engine to scan your resources, select Automatically select a data engine.
11. If you want StorageX to use a specific universal data engine group to scan your resources, select Use this data engine group, then select the universal data engine group from the drop-down menu.

12. If you want StorageX to use a specific universal data engine to scan your resources, select Use this specific data engine, then select the universal data engine from the drop-down menu.

13. If you want to tag the data returned by the scan when StorageX stores the data in the repository, complete the following steps:
   a. Click the Tags tab.
   b. For each tag you want to add, specify a tag name and value.
   c. If you want to add a new tag, click the Add a new tag icon (+).
   d. If you want to remove a tag, click the Remove this tag icon (-).

14. If you want StorageX to run the scan on an existing schedule, complete the following steps:
   a. Click the Schedule tab.
   b. Select Run scan policy on a scheduled basis.
   c. Select the existing schedule you want from the Select Schedule drop-down menu.

15. If you want StorageX to run the scan on a new schedule, complete the following steps:
   a. Click the Schedule tab.
   b. Select Run scan policy on a scheduled basis.
   c. Click Add New.
   d. Select the date you want the schedule to start running in the Start Date field.
   e. Select the time you want the schedule to start and recur in the Schedule Time field.
   f. In the Schedule Frequency drop-down menu, select the frequency on which you want StorageX to run the scan. Possible options are Once, Hourly, Daily, and Weekly.
   g. If you selected the Hourly frequency, specify the number of hourly interval at which the scan will run in the Schedule task hourly field.
   h. If you selected the Daily frequency, specify the number of days after which the scan will run in the Schedule task daily field.
   i. If you selected the Weekly frequency, specify the number of weeks after which the scan will run in the Schedule task weekly field, then select one or more days of the week on which the scan will run.

   **NOTE**
   *If you select the Weekly frequency, you must select at least one day of the week on which the scan will run.*

16. Click Save when finished.

17. In the Resource Data Scans grid, click the checkbox next to each scan you want to run.

18. Click Run. StorageX sends the scan to be run to the universal data engine(s) specified in the scan Properties.

19. If you want to cancel a running scan, complete the following steps:
a. Select the scan in the Resource Data Scans grid.

b. Click Cancel.

c. Click Yes to confirm.

20. If you want to see detailed information about a scan, click the link in the State column for the scan to view the scan manifest.

21. If you want to see the status of specific paths included in the scan, click the arrow beside the scan name to expand the scan into the list of paths the scan contains.

22. If you want to see detailed information about each specific path in a scan, either expand the scan into the list of paths or open the Paths tab of the scan manifest, then click the link in the State column for the path to view the scan path manifest.

If you need to change an existing scan, you can modify a scan in the Management Portal by selecting the scan and clicking Edit.

NOTE
If you deselect one or more paths in a scan, ensure you click Save and then click Edit again before searching for or adding any new paths. If you deselect a path and then click Storage Resources and search for a new path before clicking Save, the path you want to remove will still be selected.

Analyzing your storage resource data

Once you have scanned your resources to get your file data, you can use the StorageX Management Portal to analyze the data stored in the repository.

To analyze data using the Management Portal, create an analysis set with a dataset containing the resources you want to analyze and a query containing criteria to find the specific data you want to analyze.

To specify the data you want to analyze, you can create a new dataset, use an existing dataset, or use an existing analysis set.

The analysis set returns a list of either files or shares and exports, based on the analysis set properties, which you can view in the Analysis Results grid. You can use the Raw Data view or select a specific graphical or tabular Report, depending on how you want to view the filtered data. You can also use the analysis set to create Archive or Workload Tiering policies.

If you want to use an analysis set to create an Archive policy, you must specify that the analysis set return a list of files.

If you want to use an analysis set to create a Workload Tiering policy, you must specify that the analysis set return a list of shares and exports.

For more information about analyzing file data, see “Understanding file analysis” on page 3. For more information about viewing analysis results, see “Working with analysis results” on page 43.

NOTES:

• You must specify at least one criterion for a new dataset before you can save the analysis set.
• If you create an analysis set that includes a query matching the File Name attribute, you must include the file extension in the File Name value. For example, if you want the analysis set to return all files with the name Test.txt, you must include the .txt extension in order to see those files in the analysis results.

• If you create an analysis set that includes a query matching the File Extension attribute, you must include the leading period of the file extension in the File Extension value. For example, if you want the analysis set to return all JPEG-format files, you must specify .jpg, including the period, in order to see those files in the analysis results.

• To create or run an analysis set, your user account must be assigned at least the Read and Execute access type in the Access Control settings in the StorageX Console.

• If you create an analysis set that includes more than one run of a dataset, you can only use that analysis set for analytical purposes or to create a Workload Tiering policy. You cannot use the analysis set to create an Archive policy, and graphic-format reports in the Management Portal of the analysis results will display incorrect information.

• If you select the Include data from previous runs of the dataset option when you create an analysis set, the analysis set uses two scan runs by default. To allow the StorageX Metadata Service to function efficiently, StorageX is configured by default to groom all but the two most recent runs of any scan from the repository. If you need to modify StorageX grooming and dataset settings, please contact Data Dynamics, Inc. Support.

To create an analysis set:

1. In a browser, open the StorageX Management Portal Web interface.
2. Click the Analyze tab, then click Create.
3. In the Create Analysis Set dialog box, specify the name you want to use for the analysis set, as well as an optional description.
4. If you want StorageX to return a list of files, select List of Files under the Results type drop-down menu.
   
   NOTE
   If you create an analysis set that returns a list of files, you cannot use that analysis set to create a Workload Tiering policy.

5. If you want StorageX to return a list of shares and/or exports, select List of Shares/Exports under the Results type drop-down menu.
   
   NOTE
   If you create an analysis set that returns a list of shares and exports, you cannot use that analysis set to create an Archive policy.

6. Click the Dataset tab.
7. If you want to use an existing dataset for the analysis, click Select an existing dataset and select the dataset you want to use.
8. If you want to use an existing analysis set as the dataset for the analysis, click Select an existing dataset and select the analysis set you want to use.
9. **If you want to create a new dataset to analyze**, complete the following steps:
   a. Specify the name you want to use for the dataset.
   b. For each criterion you want to use to determine the dataset StorageX will analyze, specify a criterion name, operator, and value.
   c. **If you want to add a new criterion**, click Add rule.
   d. **If you want to add a new criteria group**, click Add group.
   e. **If you want results that meet all criteria in a group**, click AND for that group.
   f. **If you want results that meet one of a set of criteria in a group**, click OR for that group.
   g. **If you want to remove a criterion**, click Delete next to the criterion.
   h. **If you want to remove a criteria group**, click Delete next to the criteria group.

10. **If you want to include previous runs of the dataset in the analysis set**, select Include data from previous runs of the dataset. StorageX includes two runs of the dataset in the new analysis set, by default.

**ATTENTION**

*If you select this option, you cannot use the analysis set to create an Archive policy. You can only use the analysis set for analytical purposes or to create a Workload Tiering policy. In addition, any graphic-format reports in the Management Portal of the analysis results will display incorrect information.*

11. **If you want the analysis set to return a list of files**, complete the following steps:
   a. Click the Query tab.
   b. **If you want to use an existing query for the analysis**, click Select a saved query and select the query you want to run on the dataset.
   c. **If you want to create a new query for the analysis**, complete the following steps:
      i. Specify the name you want to use for the query.
      ii. For each criterion you want to use to build the query StorageX will run on the dataset, specify a criterion name, operator, and value.
      iii. **If you want to add a new criterion**, click Add rule.
      iv. **If you want to add a new criteria group**, click Add group.
      v. **If you want results that meet all criteria in a group**, click AND for that group.
      vi. **If you want results that meet one of a set of criteria in a group**, click OR for that group.
      vii. **If you want to remove a criterion**, click Delete next to the criterion.
      viii. **If you want to remove a criteria group**, click Delete next to the criteria group.

12. **If you want the analysis set to return a list of shares and/or exports**, complete the following steps:
a. Click the Statistics tab.

b. In the Statistics Criteria field, specify the statistical criterion you want to use to determine which shares and exports are included in the analysis set. StorageX examines the dataset to see which files match the query, then uses the statistics to return matching shares or exports.

For example, you can use a query that looks for all files that have not been accessed in more than 24 months, and then select Percentage of Files, greater or equal, and 50 to set the statistics criteria to match shares or exports where more than 50% of the files those shares or exports contain match the query.

StorageX returns only those files where more than half of the files are older than 24 months and ignores any shares or exports where fewer than half of the files match.

Possible options for statistics criteria are Number of Files, Number of Bytes, Number of Total Bytes, Percentage of Bytes, Percentage of Files, or Percentage of Total Bytes.

c. If you want to use an existing query for the analysis, click Select a saved query and select the query you want to run on the dataset.

d. If you want to create a new query for the analysis, complete the following steps:
   i. Specify the name you want to use for the query.
   ii. For each criterion you want to use to build the query StorageX will run on the dataset, specify a criterion name, operator, and value.
   iii. If you want to add a new criterion, click Add rule.
   iv. If you want to add a new criteria group, click Add group.
   v. If you want results that meet all criteria in a group, click AND for that group.
   vi. If you want results that meet one of a set of criteria in a group, click OR for that group.
   vii. If you want to remove a criterion, click Delete next to the criterion.
   viii. If you want to remove a criteria group, click Delete next to the criteria group.

13. Click Save when finished.

14. Select the analysis set and click Run. StorageX runs the analysis set and displays the results in the Analysis Results grid.

---

**NOTE**

If you run an analysis set that includes a large number of results and then click a different tab in the Management Portal, the analysis set run is canceled. The Analysis Results grid only displays data for a selected analysis set while the Analyze tab is active.

---

## Working with analysis results

By default, when you select an analysis set in the StorageX Management Portal, the Management Portal displays a list of all files, shares, or exports returned in the analysis in the Analysis Results grid, depending upon the options selected when the analysis set was created.

You can view the detailed file, share, or export data in its raw form in the Raw Data grid format, which allows you to use the grid controls to minutely sort or filter the data displayed, or you can click Report to view more high-level graphical-format reports.
The Management Portal includes the following canned report types, each of which can be used for the analyzed data:

**Activity Analysis Report**
Displays information about the most recently-accessed files in the analysis set results, both by file count and file size. This report is based on the Last Access Time file attribute.

**NOTE**
*If the analysis set results include a file that exactly matches the lowest value of a column on the X axis of the report, which will also be the highest value of the previous column, the file is counted in that previous column. For example, if the Last Access Time of a file is exactly 1 year in the past, that file is counted in the 6 months - 1 year column in the report, rather than in the 1 year - 2 year column in the report.*

**Size Analysis Report**
Displays information about the sizes of the files in the analysis set results, with files categorized by size range and the number of files displayed for each range.

**NOTE**
*If the analysis set results include a file that exactly matches the lowest value of a column on the X axis of the report, which will also be the highest value of the previous column, the file is counted in the higher of the two columns. For example, if the size of a file is exactly 1 MB, that file is counted in the 1 MB - 10 MB column in the report, rather than in the 256 KB - 1 MB column in the report.*

**Age Analysis Report**
Displays information about the overall age of the files in the analysis set results, both by file count and file size. This report is based on the Creation Time attribute for files located on CIFS resources and on the Last Modified Time attribute for files located on NFS resources.

**NOTE**
*If the analysis set results include a file that exactly matches the lowest value of a column on the X axis of the report, which will also be the highest value of the previous column, the file is counted in that previous column. For example, if the Creation Time of a file on a CIFS resource is exactly 60 days in the past, that file is counted in the 30 days - 60 days column in the report, rather than in the 60 days - 90 days column in the report.*

**Type Analysis Report**
Displays information about the most common file types in the analysis set results, both by file count and file size.

**Top Consumers Report**
Displays information about the user accounts who own most of the files in the analysis set results, both by file count and file size.

**File Storage Savings Report**
Displays information about your possible savings in terms of storage based on the percentage of files in the repository that match both the dataset and query criteria. *If you want to archive your files to an object storage resource using the current analysis set, this report displays the percentage of files StorageX would archive.*
You can use the Modified Time slider to see how the percentage of files to be archived changes when you change the time criterion.

For example, if the query included in your analysis set looks for files that have not been modified for more than 12 months, and the File Storage Savings report tells you that 66% of the files returned would be archived, you can use the slider to see what the percentage would be if you looked for files older than 18 or 24 months.

**NOTE**

If you modify the Modified Time criterion using the File Storage Savings report slider, the analysis set query does not reflect the changes made in the report. To modify the analysis set query, you must select the analysis set and click Edit, then make your changes in the Query tab.

The Raw Data grid can be exported in comma-separated value (CSV) format, and the individual graphical-format reports can be printed or exported in PNG, PDF, or JPEG format.

When exporting data from the Raw Data grid in CSV format, you can export all data in the analysis set or export only the data currently visible.

**NOTE**

Once you have created at least one analysis set, you can also view graphical-format reports for your analysis sets on the Home tab of the Management Portal. You can select a different analysis set for each display area.

### Archiving your file data to an object storage resource

For many users, the end goal of analyzing file data in their storage environment is to determine which files can be safely archived to non-file storage. Specifically, they may want to archive unused data to an object storage resource, so that the data no longer occupies valuable storage space on a company's file resources.

You can take the analysis set created in the Analyze tab of the Management Portal and archive that collected, filtered set of files to an object storage resource managed by StorageX, based on whatever criteria you used when you built the analysis set itself.

For example, if you wanted to archive all files that none of your users have modified in the last two years, you would create an analysis set that includes a dataset with all of your servers in it and a query that looks for files with a LastModified Time attribute that is two years or more in the past. That analysis set returns a list of files, and you can then tweak the analysis set to return the exact list you want StorageX to archive. For more information about creating and seeing the results of analysis sets, see “Analyzing your storage resource data” on page 40 and “Working with analysis results” on page 43.

With the analysis set finalized, go to the Archive tab and click Create to begin creating the Archive policy that will transform your old files into objects.

**ATTENTION**

If you create and run an Archive policy, StorageX automatically removes files from the source once the policy has archived those files to the object storage resource. Do not run an Archive policy unless you are ready to delete files from the source.
Archiving your file data to an object storage resource

NOTE
StorageX does not support archiving CIFS or NFS symbolic links to object storage resources.

*If you added custom tags to the file data you want to archive when you initially scanned your files,* StorageX pushes those custom tags to the destination object storage resource when it runs the Archive policy. After the archival process is complete, if you decide you want to retrieve some of your archived data, you can use your custom tags as criteria for retrieving specific files that you tagged during scanning.

You can also choose to add additional custom tags to the archived data on the object storage resource, using the Tags tab of the Create Archive Policy dialog box, and use those tags as criteria for retrieving your files.

When tagging archived data, keep the following restrictions in mind:

- You can only use a specific tag once per Archive policy.
- You can specify a maximum of 10 tags per policy.
- Tag names and values are case-insensitive.
- Tag key names can use either alphanumeric characters or underscores (_) and cannot start with a number. You cannot specify a key name that starts with a number or includes an invalid non-alphanumeric character.
- Tag values can use a combination of alphanumeric characters and specific valid non-alphanumeric characters. Valid non-alphanumeric characters you can include in a tag value are: underscore (_); hyphen (-); plus (+); equal (=); colon (:); white space ( ); and forward slash (/). You cannot specify a tag value that includes an invalid non-alphanumeric character.
- You can specify a tag name with a maximum of 128 characters, and a tag value with a maximum of 256 characters.
- You cannot specify a tag name or value that starts with a number or includes an invalid non-alphanumeric character.

To create an Archive policy:

1. In a browser, open the StorageX Management Portal Web interface.
2. Click the Archive tab, then click Create.
3. In the Create Archive Policy dialog box, specify the name you want to use for the archive, as well as an optional description.
4. Click the Object Store tab.
5. Click Select Endpoint/User and select the object storage resource where you want to archive your files.
   
   **NOTE**
   You can only select an object storage resource already being managed by StorageX. You cannot add a new object storage resource in the StorageX Management Portal.

6. Click Select a bucket and select the bucket where you want to archive your files.
7. Click the Archive Options tab.
8. *If you want StorageX to automatically select a universal data engine to archive your files,* select Automatically select a data engine.
9. If you want StorageX to use a specific universal data engine group to archive your files, select Use this data engine group, then select the group from the drop-down menu.

10. If you want StorageX to use a specific universal data engine to archive your files, select Use this specific data engine, then select the universal data engine from the drop-down menu.

11. Click the Analysis Set tab.

12. Click Analysis Set and select the analysis set you want to use for archiving your file data.

**NOTE**
You cannot create or modify an analysis set in the Create Archive Policy dialog box. If you want to create or modify an analysis set, click Cancel, make your analysis set changes in the Analyze tab of the Management Portal, and return to the Archive tab to create a new Archive policy for that analysis set.

13. If you want StorageX to run the Archive policy on an existing schedule, complete the following steps:
   a. Click the Schedule tab.
   b. Select Run Archive policy on a scheduled basis.
   c. Select the existing schedule you want from the Select Schedule drop-down menu.

14. If you want StorageX to run the Archive policy on a new schedule, complete the following steps:
   a. Click the Schedule tab.
   b. Select Run Archive policy on a scheduled basis.
   c. Click Add New.
   d. Select the date you want the schedule to start running in the Start Date field.
   e. Select the time you want the schedule to start and recur in the Schedule Time field.
   f. In the Schedule Frequency drop-down menu, select the frequency on which you want StorageX to run the Archive policy. Possible options are Once, Hourly, Daily, and Weekly.
   g. If you selected the Hourly frequency, specify the number of hourly interval at which the Archive policy will run in the Schedule task hourly field.
   h. If you selected the Daily frequency, specify the number of days after which the Archive policy will run in the Schedule task daily field.
   i. If you selected the Weekly frequency, specify the number of weeks after which the Archive policy will run in the Schedule task weekly field, then select one or more days of the week on which the policy will run.

**NOTE**
If you select the Weekly frequency, you must select at least one day of the week on which the Archive policy will run.

15. If you want to tag your archived files as objects on the object storage resource, complete the following steps:
   a. Click the Tags tab.
   b. For each tag you want to add, specify a tag name and value.
   c. If you want to add a new tag, click the Add a new tag icon (+).
Using Workload Tiering to migrate your file data to a different file storage resource

There are several ways to make your storage environment more efficient and cost-effective. Outside of archiving files to object storage, you can also optimize your storage usage, shifting less-frequently-used files to lower-tier file storage resources while migrating in-demand files to higher-tier file storage resources.

When you create a Workload Tiering policy, StorageX automatically generates one or more Phased Migration policies, one for each source/destination pair you select from an analysis set. You must specify an existing Phased Migration policy template when you create a Workload Tiering policy.

NOTES:
Using Workload Tiering to migrate your file data to a different file storage resource

- You can only create a Workload Tiering policy using an analysis set that returns a list of shares and exports. For more information about creating analysis sets, see “Analyzing your storage resource data” on page 40.

- You cannot create a new policy template using the Management Portal. Create or modify Phased Migration policy templates in the StorageX Console. For more information about Phased Migration policy templates, see the StorageX Administrator’s Guide.

- After you create a Workload Tiering policy, you cannot modify the generated Phased Migration policies in the Management Portal. You can only manage the generated policies in the StorageX Console. For more information about Phased Migration policies, see the StorageX Administrator’s Guide.

- If you create a Workload Tiering policy in the Management Portal, that policy is affected by any installation-wide setting changes in the StorageX Console. Specifically, if you click File > Options in the StorageX Console, click the Data Movement > Policy Validation Options tab, check the Do not perform path accessibility validation option, and then run the Workload Tiering policy, StorageX will not validate that the paths included in the policy exist and can be accessed from the StorageX server. This may cause the StorageX Console to display errors for the Workload Tiering policy run. For more information about configuring path validation for policies, see the StorageX Administrator’s Guide.

- After you create a Workload Tiering policy, you cannot change the analysis set used by the policy or change any of the specified sources or destinations included in the policy. If you want to use a different analysis set or migrate to or from a different source or destination, you must create a new Workload Tiering policy.

To create a Workload Tiering policy:

1. In a browser, open the StorageX Management Portal Web interface.
2. Click the Workload Tiering tab.
3. Click Create.
4. In the Create Workload Tiering Policy dialog box, specify the name you want to use for the policy, as well as an optional description.
5. In the Template field, select the Phased Migration policy template you want to use for all Phased Migration policies generated by the Workload Tiering policy.
6. Click the Analysis Set tab.
7. Click Analysis Set and select the analysis set you want to use for migrating share or export data. You can only select analysis sets that returns lists of shares and/or exports.

   **NOTE**
   You cannot create or modify an analysis set in the Create Workload Tiering Policy dialog box. If you want to create or modify an analysis set, click Cancel, make your analysis set changes in the Analyze tab of the Management Portal, and return to the Workload Tiering tab to create a new Workload Tiering policy for that analysis set.

8. Click Search.
9. On the Resources tab, select each source share or export you want to migrate and complete the following steps:
Viewing and modifying scans, datasets, analysis sets, and policies

After you create a scan, dataset, analysis set, or policy in the StorageX Management Portal, you can later view or modify any of those objects as necessary using the portal.

NOTES:

- If you modify a dataset included in an analysis set, the analysis set results will not change unless you re-run the scan and then re-run the analysis set.
- If you run a scan, remove a path from the scan, and then re-run the scan, any analysis set that includes the scan will display data for the removed path. The removed path information is retained in the StorageX repository, and the StorageX Metadata Service sees no new data for the path from the second scan run and therefore does not update the information for the path. We recommend that when you modify a scan, you specify a new custom tag for the modified scan. That way you can use the new tag for the analysis set and retrieve only the data you need.

To view or modify a scan, dataset, analysis set, or policy:

1. In a browser, open the StorageX Management Portal Web interface.

2. If you want to view or modify a scan, complete the following steps:
   a. Click the Discover tab.
   b. Select the scan you want to view or modify.
   c. Click Edit.
   d. Click the tab you want to view or modify.
   e. Make changes, if necessary.
   f. Click Save or Cancel when finished.

3. If you want to view or modify an analysis set, complete the following steps:
   a. Click the Analyze tab.
   b. Select the analysis set you want to view or modify.
   c. Click Edit.
d. Click the tab you want to view or modify.
e. Make changes, if necessary.
f. Click Save or Cancel when finished.

4. If you want to view or modify the dataset for an analysis set, complete the following steps:
   a. Click the Analyze tab.
   b. Select the analysis set you want to view or modify.
   c. Click the name in the Dataset column for the analysis set.
   d. Make changes, if necessary.
   e. Click Save or Cancel when finished.

5. If you want to view or modify the query for an analysis set, complete the following steps:
   a. Click the Analyze tab.
   b. Select the analysis set you want to view or modify.
   c. Click the name in the Query column for the analysis set.
   d. Make changes, if necessary.
   e. Click Save or Cancel when finished.

6. If you want to view or modify an Archive policy, complete the following steps:
   a. Click the Archive tab.
   b. Expand the Archive policy and select the version of the policy you want to view or modify.
   c. Click Edit.
   d. Click the tab you want to view or modify.
   e. Make changes, if necessary.
   f. Click Save or Cancel when finished.

7. If you want to view the analysis set for an Archive policy, complete the following steps:
   a. Click the Archive tab.
   b. Select the Archive policy you want to view or modify.
   c. Click the name in the Analysis Set column for the Archive policy.
   d. Make changes, if necessary.
   e. Click Save or Cancel when finished.

8. If you want to view or modify a Workload Tiering policy, complete the following steps:
   a. Click the Workload Tiering tab.
   b. Select the Workload Tiering policy you want to view or modify.
   c. Click Edit.
   d. Make changes to the name or description for the policy, if necessary.
Deleting scans, analysis sets, and policies

If you no longer need a scan, analysis set, or policy, you can delete the object in the StorageX Management Portal.

NOTES:
- You cannot delete an analysis set that is used as the dataset for a second analysis set.
- You cannot delete an analysis set that includes a dataset or query that is in use by a second analysis set.
- You cannot delete an analysis set that is included in an Archive or Workload Tiering policy.

To delete a scan, analysis set, or policy:

1. In a browser, open the StorageX Management Portal Web interface.
2. If you want to delete a scan, complete the following steps:
   a. Click the Discover tab.
   b. Select the scan you want to delete.
   c. Click Delete.
   d. Click Yes to confirm.
3. If you want to delete an analysis set, complete the following steps:
   a. Click the Analyze tab.
   b. Select the analysis set you want to delete.
   c. Click Delete.
   d. Click Yes to confirm.
   e. If you also want to delete any datasets or queries included in the analysis set, click Yes.
   f. If you do not want to delete datasets or queries included in the analysis set, click No.
4. If you want to delete an Archive policy, complete the following steps:
   a. Click the Archive tab.
   b. Expand the Archive policy and select the version of the policy you want to delete.
   c. Click Delete.
   d. Click Yes to confirm.
5. If you want to delete a Workload Tiering policy, complete the following steps:
a. Click the Workload Tiering tab.

b. Select the policy you want to delete.

c. Click Delete.

d. Click Yes to confirm.

Configuring Cassandra grooming settings

When Cassandra is installed in your environment for use in backing up Elasticsearch indices, StorageX automatically grooms old, unused data out of the configured Cassandra repository.

By default, StorageX stores data in the Cassandra repository for one year (365 days) before grooming that data, and runs a Cassandra grooming job once per week to look for data more than a year old. However, you can configure those Cassandra grooming settings in the Maintenance Portal tab of the StorageX Management Portal.

In addition to the default Cassandra grooming provided by StorageX, you can use native Cassandra tools or other third-party tools to remove old data from your repository. For more information about Cassandra, see http://cassandra.apache.org/.

To configure Cassandra repository grooming settings in StorageX:

1. In a browser, open the StorageX Management Portal Web interface.

2. Click Data Grooming Options > Cassandra Grooming.

3. In the Grooming Cutoff tab of the Edit Cassandra Grooming Options dialog box, specify the number of days, weeks, months, or years you want to retain data in your Cassandra repository. Use the Count field to specify the value and the Unit field to select Day, Week, Month, or Year.

   For example, if you want to retain your data for three months, you would enter 3 in the Count field and select Month in the Unit field.

4. If you want to change the default Cassandra grooming schedule, complete the following steps:
   a. Click the Schedule tab.
   b. In the Start Date field, specify the date you want the grooming schedule to start.
   c. In the Schedule Time field, specify the time of day you want the grooming job to run.
   d. In the Schedule Frequency field, select the frequency at which you want StorageX to groom old Cassandra data. Possible options are Once, Hourly, Daily, or Weekly.
   e. If you select Hourly, Daily, or Weekly for the grooming schedule frequency, specify the number of hours, days, or weeks you want to use as an interval.

      For example, if you want to groom old Cassandra data every three weeks, you would select Weekly in the Schedule Frequency field and then specify 3 in the Schedule task weekly field.

5. When finished, click Save.
Accessing Management Portal monitoring tools

If you have configured access to third-party monitoring tools for the Management Portal or database, you can access those tools in the Maintenance Portal tab, including Prometheus, Grafana, and Reaper. When you click on each tool sub-tab, the Management Portal opens the separate endpoint in a new browser window or tab.

NOTE
If you use a browser that automatically blocks pop-ups or an external pop-up blocker, the Management Portal cannot open new browser windows or tabs for your configured monitoring tools. Ensure you disable any pop-up blocking functionality before trying to access monitoring tools from the Management Portal.

If you have not configured the StorageX Management Portal to access a monitoring tool, no window or tab will be opened when you click the sub-tab for that tool. You can click Configure to update your tool endpoints, if necessary.

For more information about configuring StorageX Management Portal monitoring tools, see “Configuring monitoring for StorageX Management Portal components” on page 30.

For more information about Prometheus, Grafana, and Reaper, see their respective Web sites:

- https://prometheus.io/
- https://grafana.com/
- http://cassandra-reaper.io/

To access StorageX Management Portal monitoring tools:

1. In a browser, open the StorageX Management Portal Web interface.
2. If you want to access your Prometheus installation from the Management Portal, click Prometheus.
3. If you want to access your Grafana installation from the Management Portal, click Grafana.
4. If you want to access your Reaper installation from the Management Portal, click Reaper.
5. When finished, close the new browser windows/tabs opened for each tool.
Chapter 4

Installing and Using the StorageX Retrieval Portal

This section explains how to install, configure, and access the StorageX Retrieval Portal product in your environment.

In this chapter

• “Understanding the StorageX Retrieval Portal” on page 55
• “Understanding StorageX Retrieval Portal requirements” on page 55
• “Preparing to install the StorageX Retrieval Portal” on page 57
• “Installing the StorageX Retrieval Portal” on page 58
• “Accessing the StorageX Retrieval Portal” on page 60
• “Retrieving files using the StorageX Retrieval Portal” on page 60
• “Upgrading the StorageX Retrieval Portal” on page 62
• “Uninstalling the StorageX Retrieval Portal” on page 63

Understanding the StorageX Retrieval Portal

Having your data archived to an object storage resource, whether in the cloud or your company’s private lab, is not useful unless you can get that archived data back when you need it. The purpose of the StorageX Retrieval Portal is to allow you to locate specific files archived on an object storage resource to which you have access and retrieve those files for further use.

Using the Retrieval Portal, you can find the files you need and request that StorageX retrieve those files. The StorageX File Recovery Service then retrieves the files, stores the files in a configured Staging Area where you can access them, and emails you to let you know your files are ready.

If you are an administrator, you can view all of your users’ file retrieval requests and their current status and make retrieval requests on behalf of any of the users in your environment.

Understanding StorageX Retrieval Portal requirements

This section describes the minimum system requirements necessary for the StorageX Retrieval Portal to function properly.

Supported operating systems

The StorageX Retrieval Portal must be installed on a computer running one of the following operating systems with the latest service pack:

• Microsoft Windows Server 2012 Essentials, Standard, or Datacenter Edition
Understanding StorageX Retrieval Portal requirements

- Microsoft Windows Server 2012 R2 Essentials, Standard, or Datacenter Edition
- Microsoft Windows Server 2016 Essentials, Standard, or Datacenter Edition
- Microsoft Windows 8 Pro or Enterprise Edition (64-bit version)
- Microsoft Windows 8.1 Pro or Enterprise Edition (64-bit version)
- Microsoft Windows 10 Pro or Enterprise Edition (64-bit version)

**NOTE**
This computer can be a server, a workstation, or a virtual machine (VM).

**Permissions requirements**

The account used to install the StorageX Retrieval Portal must be a member of the local Administrators group on the installation computer.

In addition, the account used to run the StorageX Retrieval Portal must also be a member of the local Administrators group on the installation computer, must be added to the Administrators role in the Access Control settings in the StorageX Console, and must have the permissions required to log on as a service on the installation computer.

**NOTE**
You must ensure the account you want to use as the StorageX Retrieval Portal service account has the Log on as a service privilege on a computer before installing the Retrieval Portal on that computer.

**Network port requirements**

Before installing the StorageX Retrieval Portal, ensure that TCP port 9888 is open in any firewalls between the portal and the StorageX server. The StorageX Retrieval Portal and File Retrieval Service REST API use this port to communicate with the StorageX server.

**Web service requirements**

The computer where you want to install the StorageX Retrieval Portal must have the HTTP service enabled.

**Web browser requirements**

Because the StorageX Retrieval Portal is Web-based, users must use the latest version of one of the following Web browsers to access the product:

- Mozilla Firefox
- Google Chrome
- Microsoft Edge
- Microsoft Internet Explorer

**NOTE**
*If using Microsoft Internet Explorer,* the StorageX Management Portal only supports Microsoft Internet Explorer 11.
Preparing to install the StorageX Retrieval Portal

Staging Area storage space requirements

In order to install and use the StorageX Retrieval Portal, you must configure a Staging Area on a CIFS share accessible from the Retrieval Portal computer. The Staging Area should have sufficient storage space to store a large number of files retrieved from one or more object storage resources, at least temporarily.

Preparing to install the StorageX Retrieval Portal

Before installing the StorageX Retrieval Portal, you must first configure both the Staging Area you want to use to store files retrieved for users and at least one notification profile.

NOTE
The Staging Area and notification profile are both required by the StorageX Retrieval Portal installation package. If you do not configure both before running the installer, you cannot complete the installation process.

StorageX uses the Staging Area to store retrieved files until the requesting user can download them, and the notification profile to email users who request files through the Retrieval Portal when their files are ready to be downloaded.

To set up your environment for the Retrieval Portal:

1. Set up a CIFS share on an accessible server in your environment to be used as the Staging Area for the Retrieval Portal.

   NOTE
   We recommend allocating a significant amount of storage space to this Staging Area, so that multiple files requested by multiple users can be stored there until they are downloaded.

2. Log on to the StorageX Console computer using a user account that has Administrator access to your StorageX environment.

3. In the StorageX Console, click File > Options.

4. Click the Notification Profile Properties tab.

5. Click New.

6. Specify the SMTP settings you want to use for the notification profile. For more information about the SMTP settings, click the Help button on the dialog box.

7. Specify the message settings you want to use for the new profile. For more information about the message settings, click the Help button on the dialog box.

   NOTE
   The StorageX Retrieval Portal does not utilize the email address provided in the Message Settings > To: field. Instead, the Retrieval Portal sends the email notification to the email address provided by the user or administration in the Portal itself.

8. Click Send Test Message to verify the profile settings are correct.

9. If you want to use the notification profile for all notifications, click Enable all usage of this notification.
10. Click OK.
11. When finished, click OK.
12. Click Access Control.
13. Under Role-based Access Control, verify that Enable is selected.
   For more information about role-based access control in StorageX, see the StorageX
   Administrator’s Guide.
14. Click OK.
15. Close the StorageX Console.

Installing the StorageX Retrieval Portal

When users want to retrieve their archived data from an object storage resource, they can use the
StorageX Retrieval Portal.

The Retrieval Portal is installed on a Microsoft Windows server and can be accessed from any
computer with an up-to-date Web browser. You must install the Retrieval Portal separately from the
main StorageX installation.

If you want to install the StorageX Retrieval Portal on a computer where only version 1.2 of TLS is
enabled, the installation program may display an error.

NOTE
This issue only applies to computers with Microsoft Windows Server 2016 or later installed.

To work around this issue, Microsoft recommends modifying .NET Framework registry keys on the
computer where you want to install the Metadata Service to allow the computer to use the
strongest version of TLS that is available.

ATTENTION
Before making changes to a Windows registry key, ensure you back up or export the specific key, in
case you need to later revert to the original.

Add the specified values to the following registry keys either manually using the Registry Editor or
by creating and running or importing the text below as a .reg file:

Windows Registry Editor Version 5.00

[HKEY_LOCAL_MACHINE\SOFTWARE\WOW6432Node\Microsoft\.NETFramework\v4.0.30319]
"SystemDefaultTlsVersions"=dword:00000001
"SchUseStrongCrypto"=dword:00000001

[HKEY_LOCAL_MACHINE\SOFTWARE\Microsoft\.NETFramework\v4.0.30319]
"SystemDefaultTlsVersions"=dword:00000001
"SchUseStrongCrypto"=dword:00000001

When finished, restart the computer, open Windows PowerShell using an account that has
Administrator access, and enter the following command:

[Net.ServicePointManager]::SecurityProtocol

PowerShell should display [SystemDefault] in the output.
For more information about configuring TLS in Microsoft Windows environments, see https://docs.microsoft.com/en-us/dotnet/framework/network-programming/tls.

To install the StorageX Retrieval Portal:

1. Log on to the StorageX server computer using an user account that is a member of the local Administrators group on the computer.
2. Open a command prompt, running the tool as an administrator.
3. Navigate to the location where you saved the StorageX installation package.
4. Within the installation package, navigate to the FileRecoveryService\Windows folder.
5. Enter the FRSSetup.msi command.
6. On the Welcome to the StorageX Retrieval Portal Setup Wizard dialog box, review the welcome message and then click Next.
7. On the End-User License Agreement dialog box, review the terms of the license agreement, select the I accept the terms in the License Agreement check box, and then click Next.
8. On the StorageX Retrieval Portal Account dialog box, specify the account you want to use to run the StorageX Retrieval Portal. In the Account field, type the name of the domain user account you want to use for the Retrieval Portal service account, and in the Password field, type the password for the account.
9. Click Next.
10. On the StorageX Retrieval Portal Settings dialog box, specify the location you want to use as the Staging Area for the Retrieval Portal in UNC format.

\textbf{ATTENTION}
After installation, these settings cannot be modified.

11. Select the email notification profile you want StorageX to use to notify the end user when their files have been retrieved.

\textbf{NOTE}
You can only select email notification profiles that are already configured in the StorageX Console. You cannot create or configure profiles in this installer. \textit{If you do not see the profile you want to use in this list}, exit the installer, use the StorageX Console to create the profile, and then rerun the installer. For more information about creating email notification profiles, see the StorageX Administrator's Guide.

12. Click Next.
13. On the Ready to install StorageX Retrieval Portal dialog box, click Install to begin the installation.
14. On the Completed the StorageX Retrieval Portal Setup Wizard dialog box, click Finish to close the setup program.
15. Close the command prompt.
Accessing the StorageX Retrieval Portal

You access the StorageX Retrieval Portal using a Web browser.

NOTES:

- If the browser displays a pop-up dialog box, do not select the Prevent this page from creating additional dialogs checkbox before clicking OK. If you disable the browser’s pop-up dialog box capability, you cannot delete analysis groups, paths, or scans, and you will not be notified with any further message dialogs.

- In addition, if you block pop-up windows or dialog boxes, you cannot access the StorageX Management Portal directly from the Retrieval Portal.

To access the StorageX Retrieval Portal:

1. Open a Web browser.

2. If you use Microsoft Internet Explorer, complete the following steps to enable the browser to correctly display the StorageX Web interface:
   a. In Internet Explorer, click Tools > Compatibility View settings.
   b. Clear Display intranet sites in Compatibility View and click Close.

3. If you use a browser that automatically blocks pop-ups or an external pop-up blocker, disable any pop-up blocking functionality before accessing the Retrieval Portal.

   NOTE
   If you do not disable all pop-up blockers, you cannot access the StorageX Management Portal directly from the Retrieval Portal.

4. In the browser, navigate to the following URL:

   https://servername:9888/index.html

   Where servername is the name of the computer where you installed and started the StorageX Retrieval Portal.

5. If the browser displays a warning about invalid certificates, ignore the warning and continue.

6. When prompted, provide the user name and password of the account you want to use to access the Retrieval Portal, then click Log In.

Retrieving files using the StorageX Retrieval Portal

If your data has been previously archived to object storage, and you decide you need to access some or all of that data, you can use the StorageX Retrieval Portal to retrieve the files you want and download those retrieved files. You can request that StorageX retrieve archived files from any object storage resource to which you have access.

When you find the files you want retrieved, you can either browse a tree of archived files and select the files you want to request or search for the specific files you want to request using custom tags or file metadata.

For example, if you know you need a specific Microsoft PowerPoint presentation created in January 2015, you can create a search that looks for the name of the .ppt file, looks for all files with a Creation Time older than 24 months, or looks for both criteria combined.
After you submit a request to have files retrieved, StorageX queues the retrieval request, running the retrieval process as soon as resources are available, and then emails a UNC path with the location of the retrieved files to the email address you specify when you submit the request.

StorageX creates a unique folder within the Staging Area, stores your retrieved files in the new folder, and assigns Full Control access for that folder to you as the requester and to the local administrator. No other user can access your retrieved files.

Alternatively, administrators can request archived files belonging to any users in the StorageX environment. An administrator can submit a retrieval request on behalf of a specific user and have the file location sent to that user.

**NOTE**
When you retrieve files from an object storage resource, StorageX does not rebuild the original file structure of the retrieved files, but instead returns a large collection of files.

To request the retrieval of archived files using the StorageX Retrieval Portal:

1. In a browser, open the StorageX Retrieval Portal Web interface.
2. Log in using either your user account or an account with administrator permissions, as applicable.
3. If you are a user and want to search for a set of your archived files to retrieve, complete the following steps:
   a. Click the Search tab.
   b. For each Dataset and Query criterion you want to use to search for your files, specify a criterion name, operator, and value.
   c. If you want to add a new criterion, click + Condition.
   d. If you want to add a new criteria group, click + Group.
   e. If you want results that meet all criteria in a group, click AND for that group.
   f. If you want results that meet one of a set of criteria in a group, click OR for that group.
   g. If you want to remove a criterion, click the Remove this criterion icon (-) next to the criterion.
   h. If you want to remove a criteria group, click the Remove this criteria group icon (-) next to the criteria group.
   i. When finished configuring the search criteria, click Search for My Files.
   j. Select the files you want to retrieve.
   k. In the right-side column, click Specify notification email address and enter the email address you want the Retrieval Portal to use to notify you when the files have been retrieved from object storage.
4. If you are a user and want to browse for a set of your archived files to retrieve, complete the following steps:
   a. Click the Select tab.
   b. Double-click each tree node you want to browse to expand that node.
   c. Navigate to the nodes that contain the specific files you want to retrieve.
d. Select the files you want to retrieve.

e. In the right-side column, click Specify notification email address and enter the email address you want the Retrieval Portal to use to notify you when the files have been retrieved from object storage.

5. **If you are an administrator and want to search for a set of archived files to retrieve for another user**, complete the following steps:

a. Click the Search tab.

b. For each Dataset and Query criterion you want to use to search for the files, specify a criterion name, operator, and value.

c. **If you want to add a new criterion**, click + Condition.

d. **If you want to add a new criteria group**, click + Group.

e. **If you want results that meet all criteria in a group**, click AND for that group.

f. **If you want results that meet one of a set of criteria in a group**, click OR for that group.

g. **If you want to remove a criterion**, click the Remove this criterion icon (-) next to the criterion.

h. **If you want to remove a criteria group**, click the Remove this criteria group icon (-) next to the criteria group.

i. When finished configuring the search criteria, click Search for My Files.

j. Select the files you want to retrieve.

k. In the right-side column, click Specify fully-qualified user name or group and enter the user or group name to which you want to assign ownership of the retrieved files.

l. Click Specify notification email address and enter the email address you want the Retrieval Portal to use to notify you when the files have been retrieved from object storage.

6. Click Get My Data.

7. When the Success dialog box appears, review the text and click X to close.

8. Check the notification email address until you see an email notification that the requested files were retrieved.

9. In Windows Explorer, navigate to the path included in the email notification. You can then copy the files to a different location, open the files, or send the files to another user, as necessary.

---

**Upgrading the StorageX Retrieval Portal**

When you upgrade StorageX and the StorageX Management Portal, you must also upgrade your StorageX Retrieval Portal installation.

For information on upgrading the StorageX Management Portal, see “Upgrading the StorageX Management Portal” on page 32.

---

**NOTE**

After you upgrade StorageX, ensure you clear your browser cache before reopening the StorageX Retrieval Portal.
To upgrade an existing StorageX Retrieval Portal installation:

1. Before upgrading the StorageX Retrieval Portal, first upgrade the StorageX server, StorageX database, StorageX Management Portal, and all StorageX Metadata Service components. For information about upgrading the StorageX server, see the StorageX Administrator’s Guide.

2. Log on to the StorageX server where the StorageX Management Portal is installed using a user account that is a member of the local Administrators group on the computer.

3. Close any running browsers or instances of the StorageX Console.

4. Open a command prompt, running the tool as an administrator.

5. Navigate to the location where you saved the StorageX installation package.

6. Within the new installation package, navigate to the FileRecoveryService\Windows folder.

7. Enter the FRSSetup.msi command.

8. On the Welcome to the StorageX Retrieval Portal Setup Wizard dialog box, review the welcome message and then click Next.


10. On the Completed the StorageX Retrieval Portal Setup Wizard dialog box, click Finish to close the setup program.

11. Close the command prompt.

12. In a browser, open the StorageX Retrieval Portal Web interface.

If you want to remove the StorageX Retrieval Portal from an installed computer, we recommend using the Windows Control Panel to uninstall the product.

NOTE
When uninstalling StorageX as a whole, first uninstall all instances of the StorageX Retrieval Portal, then uninstall all instances of the StorageX Metadata Service and StorageX Management Portal, and then uninstall all other StorageX components. For more information about uninstalling the Management Portal, see “Uninstalling the StorageX Management Portal” on page 35. For more information about uninstalling other StorageX components, see the StorageX Administrator’s Guide.

To uninstall the StorageX Retrieval Portal:

1. Log on to a computer where the StorageX Retrieval Portal is installed.

2. If the Data Dynamics StorageX product is also installed on the computer where the StorageX Retrieval Portal is installed, close any instances of the StorageX Console on the local computer before uninstalling StorageX.

3. Open the Windows Control Panel.

4. Click Add or Remove Programs or Programs > Uninstall a program, depending on the version of Windows installed on the computer.

5. Select StorageX Retrieval Portal.

6. Click Uninstall.
7. Click Yes to confirm.

8. *If a Windows User Account Control dialog box displays,* click Yes to confirm you want to run the uninstallation program.

9. When finished, close the Programs and Features dialog box.