HPE Storage Optimizer

Software Version: 5.5

Administration Guide

Document Release Date: July 2017
Software Release Date: June 2017

Legal notices

Warranty

The only warranties for Hewlett Packard Enterprise Development LP products and services are set forth in the express warranty statements accompanying such products and services. Nothing herein should be construed as constituting an additional warranty. HPE shall not be liable for technical or editorial errors or omissions contained herein.

The information contained herein is subject to change without notice.

Restricted rights legend

Confidential computer software. Valid license from HPE required for possession, use or copying. Consistent with FAR 12.211 and 12.212, Commercial Computer Software, Computer Software Documentation, and Technical Data for Commercial Items are licensed to the U.S. Government under vendor's standard commercial license.

Copyright notice

© Copyright 2015-2017 Hewlett Packard Enterprise Development LP

Trademark notices

Adobe™ is a trademark of Adobe Systems Incorporated.

Microsoft® and Windows® are U.S. registered trademarks of Microsoft Corporation.

UNIX® is a registered trademark of The Open Group.

This product includes an interface of the 'zlib' general purpose compression library, which is Copyright © 1995-2002 Jean-loup Gailly and Mark Adler.

Documentation updates

The title page of this document contains the following identifying information:

- Software Version number, which indicates the software version.
- Document Release Date, which changes each time the document is updated.
- Software Release Date, which indicates the release date of this version of the software.

To check for recent software updates, go to https://softwaresupport.hpe.com/patches.

To verify that you are using the most recent edition of a document, go to https://softwaresupport.hpe.com/manuals.

This site requires that you register for an HPE Passport and sign in. To register for an HPE Passport ID, go to https://hpp12.passport.hpe.com/hppcf/login.do.

You will also receive updated or new editions if you subscribe to the appropriate product support service. Contact your HPE sales representative for details.

Support

Visit the HPE Software Support Online web site at https://softwaresupport.hpe.com.

This web site provides contact information and details about the products, services, and support that HPE Software offers.

HPE Software online support provides customer self-solve capabilities. It provides a fast and efficient way to access interactive technical support tools needed to manage your business. As a valued support customer, you can benefit by using the support web site to:

- Search for knowledge documents of interest
- Submit and track support cases and enhancement requests
- Download software patches
- Access product documentation
- Manage support contracts

- Look up HPE support contacts
- Review information about available services
- Enter into discussions with other software customers
- Research and register for software training

Most of the support areas require that you register as an HPE Passport user and sign in. Many also require a support contract.

To register for an HPE Passport ID, go to https://hpp12.passport.hpe.com/hppcf/login.do.

To find more information about access levels, go to https://softwaresupport.hpe.com/web/softwaresupport/access-levels.

Contents

Chapter 1: Introduction

This chapter provides an overview of HPE Storage Optimizer.

- HPE Storage Optimizer product suite
- HPE Storage Optimizer architecture
- Related documentation

HPE Storage Optimizer product suite

HPE Storage Optimizer software is an effective storage optimization solution for IT departments looking to cut the cost and complexity of storing and managing vast volumes of unstructured data. As an analytics-driven solution, HPE Storage Optimizer combines file analytics with policy-based data storage tiering and information optimization. This unique combination of technology allows you to intelligently reduce the total volume of data storage, shrink the cost and complexity of managing unstructured data, and intelligently distribute information across multiple storage repositories, including the cloud.

- **Increased cost containment:** HPE Storage Optimizer analyzes files based on metadata so that you can identify data currently stored in tier 1 infrastructure and move it to tier 2 storage. This makes it possible to materially lower the cost of primary storage and backup-related storage. You can make more effective and intelligent use of tiered storage, including the cloud, for added cost savings.
- **Better infrastructure management:** With HPE Storage Optimizer, storage optimization is no longer blind. Now you can bring together the power of file analytics and prioritized data backup in one cost-effective solution. This allows you to get more value from your existing infrastructure and significantly reduces OPEX.
- **Faster, simpler backup:** With HPE Storage Optimizer, you can reduce backup times by up to 50 percent while increasing application performance—with no noticeable impact on end users.
- **A truly holistic information governance strategy:** The analytics capabilities of HPE Storage Optimizer also enable you to optimize your governance and purchasing strategies. For example, an audit trail provides total visibility into the data you have, making it possible to know what you can defensibly dispose of. You can bridge the gap between legal and compliance, validate purchases are made with long-term objectives in mind, and put your IT team in a strategic position within the enterprise.

HPE Storage Optimizer architecture

HPE Storage Optimizer has a web application user interface. Functionality is available through several Dashboards in the user interface.

Components

HPE Storage Optimizer includes the following components.

- HPE Storage Optimizer Dashboard
- HPE Storage Optimizer Engine

- HPE Storage Optimizer Data Analysis
- HPE Storage Optimizer Connectors

HPE Storage Optimizer Dashboard

The HPE Storage Optimizer Dashboard interface allows users to view repositories, establish and review allocation of policies, administer categories, and monitor system activity and health, depending on their roles.

The following serviceis included in the HPE Storage Optimizer.

- **HPE Storage Optimizer Web Interface** is an IIS Web application that serves as the HPE Storage Optimizer user interface

HPE Storage Optimizer Engine

The HPE Storage Optimizer Engine provides the central capability to manage policy content within an organization.

The following services are included in the HPE Storage Optimizer Engine.

- **HPE Storage Optimizer Engine service** is a Windows service that executes all scheduled tasks
- **CallbackHandler** is an IIS Web application that receives notifications from HPE IDOL connectors
- **ControlPointLicenseService** is a Windows service that tracks the data usage details of your Storage Optimizer environment. The data populates the Usage Details page in the Storage Optimizer Dashboard.

HPE Storage Optimizer Data Analysis service

HPE Storage Optimizer Data Analysis allows your organization to analyze, understand, and deal with the unstructured data contained in legacy repositories.

HPE Storage Optimizer Connectors

The following connector types can be deployed from HPE Storage Optimizer IDOL Deploy Tool:

- The **HPE Storage Optimizer Exchange Connector** service scans and performs actions on items in Exchange repositories. This connector type has a connector framework deployed alongside.
- The **HPE Storage Optimizer FileSystem Connector** service scans and performs actions on items in file shares. This connector type has a connector framework deployed alongside.
- The **HPE Storage Optimizer Hadoop Connector** service scans and performs actions on items in Hadoop repositories. This connector type has a connector framework deployed alongside.
- The **HPE Storage Optimizer SharePoint 2007 Connector** service scans and performs actions on items in SharePoint 2007 sites. This connector type has a connector framework deployed alongside.
- The **HPE Storage Optimizer SharePoint 2010 Connector** service scans and performs actions on items in SharePoint 2010 sites. This connector type has a connector framework deployed alongside.
- The **HPE Storage Optimizer SharePoint 2013 Connector** service scans and performs actions on items in SharePoint 2013 sites. This connector type has a connector framework deployed alongside.

- The **HPE Storage Optimizer SharePoint Remote Connector** service scans and performs actions on items in SharePoint 2016 and SharePoint Remote sites. This connector type has a connector framework deployed alongside.

Related documentation

The following documents provide more detail on HPE Storage Optimizer.

- *HPE Storage Optimizer Installation Guide*
- *HPE Storage Optimizer Best Practices Guide*
- *HPE Storage Optimizer Administration Guide*
- *HPE Storage Optimizer Remote Analysis Agent Technical Note*
- *HPE Storage Optimizer Support Matrix*

The following documents provide more detail on HPE IDOL connectors.

- *HPE IDOL Distributed Connector Administration Guide*
- *HPE IDOL Exchange Connector (CFS) Administration Guide*
- *HPE IDOL File System Connector (CFS) Administration Guide*
- *HPE IDOL Hadoop Connector (CFS) Administration Guide*
- *HPE IDOL SharePoint 2007 Connector (CFS) Administration Guide*
- *HPE IDOL SharePoint 2010 Connector (CFS) Administration Guide*
- *HPE IDOL SharePoint 2013 Connector (CFS) Administration Guide*
- *HPE IDOL SharePoint Remote Connector (CFS) Administration Guide*

Chapter 2: Manage security

This chapter explains how to add users, how to apply role-based security and how file level security works.

- Introduction
- Enable HPE Storage Optimizer security
- User roles
- Set global role-based security
- Set security on categories, policies, and repositories
- File level security

Introduction

HPE Storage Optimizer supports a variety of role-based security settings that you can use to control user access to repositories, policies, categories, and administrative tasks.

Use the HPE Storage Optimizer Configuration Manager utility to identify a System Administrator and an LDAP server/base Distinguished Name. The System Administrator can then configure system-wide security settings.

You can apply role-based security settings either globally or at the policy, repository, or category level. Low-level security settings override the global settings.

Enable HPE Storage Optimizer security

You enable HPE Storage Optimizer role-based security in the HPE Storage Optimizer Configuration Manager.

To enable role-based security

1. Open the HPE Storage Optimizer Configuration Manager utility.
2. On the Security tab, select **Enable Security**.
3. In the **System Administrator Account** section, assign a HPE Storage Optimizer System Administrator by entering the **Domain** and the **Username**.
4. In the **Active Directory Settings** section, identify the LDAP server by entering the **Server** name and **Base DN**.
5. Click **Deploy**.

   The solution redeploys.

   > **NOTE:**
   > In addition to the LDAP server for the Active Directory Base DN, file-based security is also supported. See File level security, on page 21.

   Example

   For a Base DN entered as `'file:\\MACHINENAME\folder\userFile.xml'`, the sample file structure can be as follows:

```xml
<?xml version="1.0" encoding="utf-8"?>
<users>
    <user name="hpe\$Agnes.HPE Storage Optimizer" displayName="\$Agnes.HPE
Storage Optimizer">
        <group name="Test1" displayName="Test Group 1" />
    </user>
    <user name="hpe\$Billy.HPE Storage Optimizer" displayName="$Billy.HPE
Storage Optimizer" >
        <group name="Test1" displayName="Test Group 1" />
        <group name="Test2" displayName="Test Group 2" />
    </user>
    <user name="hpe\$Ciaran.HPE Storage Optimizer" displayName="$Ciaran.HPE
Storage Optimizer" >
        <group name="Test1" displayName="Test Group 1" />
    </user>
</users>
```

User roles

System Administrators can assign a combination of permissions to users depending on their roles in the organization. There are four major categories of user roles, and divisions within those categories.

- **Administration** permissions determine which features users can access in the user interface. Administrators can set permissions at the category, policy, or repository level, which override the global permission settings.

  - **Console Administrator** has access to the Administration dashboard, and can control HPE Storage Optimizer security settings.

- **Category** permissions apply to categories.

  - **Category Administrator** has full category permissions.

  - **Category Assigner** can view categories and assign them to policies.

  - **Category Editor** can edit, publish, create, and secure categories.

  - **Category Viewer** can view categories.

- **Policy** permissions apply to HPE Storage Optimizer policies.

  - **Policy Administrator** can create, view, edit, secure, and delete policies.

  - **Policy Approver** can view, manually assign, and approve policies, and can remove policies from documents.

  - **Policy Assigner** can view and manually assign policies.

  - **Policy Editor** can create, view, edit, secure, and delete policies.

  - **Policy Viewer** can view policies.

- **Repository** permissions apply to repositories.

  For individual repository content access, repository level roles take precedence over system level roles. For repository group content access, system level roles take precedence over repository level roles.

- ○ **Repository Administrator** has full repository permissions.
- ○ **Repository Coordinator** can tag analyzed repositories and manually assign policies to content.
- ○ **Repository Manager** can manually tag or perform actions on repositories.
- ○ **Repository Owner** can analyze the repository, view the analysis information and manually tag a document.
- ○ **Repository User** can manually assign policies to content.
- ○ **Repository Viewer** can view repositories.

Set global role-based security

Global role-based security settings determine the default permissions that users have. Administrators can combine user roles as desired to fit the profile of each user or user group.

> **NOTE:**
> Only administrators can set global role-based security settings.

To set global user roles

1. On the **Administration** dashboard, click **Security Management**.

   The All Security page opens. It lists the names and permissions of all users.

2. To filter the permissions by security type, select an item from the **All Security** list.

   Available security types include: Administration, Category, Policy, Repository, or the default, All Security.

To add a user or user group

1. Click **Add**.

2. Enter the name of the user or user group in the text box.

3. Click ☑ to verify the name against the LDAP server. You can only add valid user or group names.

   Add as many users or user groups as needed.

To edit permissions

1. Click **Edit** for the desired user or group.

   The Permissions dialog box appears.

2. Select one or more roles to assign to the user or group, and then click **Apply**.

   The selected roles are displayed in the **Permissions** column. If you assign more than one role to a user or group, the role with the highest permission level takes precedence.

   For more details on roles, see User roles, on the previous page

3. Click **Save**.

   The security settings are applied and are inherited by all categories, repositories, and policies.

Set security on categories, policies, and repositories

Categories, policies, and repositories inherit their security settings from the global settings. Sometimes it is necessary to override global permissions.

For example, you may want to allow an employee to view all repositories, yet only give permission to manage a single repository.

The System Administrator must set user permissions initially, but after the System Administrator assigns Category Administrators, Policy Administrators, and Repository Administrators permissions, those Administrators can set permissions on individual categories, policies, and repositories respectively.

Set security on categories

A category inherits the security settings from its parent category. Top-level categories inherit security settings from the All Category global security settings, which the System Administrator or any Category Administrator can set. Setting security on an individual category overrides the inheritance of settings from the parent category.

Set security for all categories

The HPE Storage Optimizer System Administrator must assign permissions initially. After one or more users is assigned the Category Administrator role, those users can also modify All Category settings.

You can set All Category security in two places:

- from the global security settings page. See Set global role-based security, on the previous page.
- from the Categories dashboard.

To set user permissions for all categories

1. On the **Categories** dashboard, click 🔒 above the category list.

   The Secure All Categories dialog box opens.

2. (Optional) **To add a user**

   a. Click **Add**.

   b. Enter the name of the user in the text box and click ☑ to verify the name against the LDAP server. You can only add valid user names.

   Add as many users as are needed.

3. To edit user permissions, click **Edit** by the desired user.

   A user role dialog box opens.

4. Select the category user role or roles to assign to the user. For details on user roles, see User roles, on page 17.

   The selected roles are displayed in the **Permissions** column. If you assign more than one role to a user, the role with the highest permission level takes precedence.

5. When you finish adding users and setting permissions, click **OK**.

The security settings apply and are inherited by all categories.

Set permissions for an individual category

Individual categories inherit security settings from their parents. In some cases you may want to override the inheritance.

For example, if a user has a Category Viewer role at the All Categories level, yet you want to give the user Category Editor privileges for one category.

HPE Storage Optimizer System Administrators and Category Administrators can set category-level security.

To set permissions on an individual category

1. Select a category from the taxonomy, and then click **Security** (🔒).

   The Secure <*Category Name*>dialog box opens.

2. (Optional) **To add a user**

   a. Click **Add**.

   b. Enter the name of the user in the text box, and then click ☑ to verify the name against the LDAP server. You can only add valid user names.

   Add as many users as needed.

3. To edit user permissions, click **Edit** by the desired user.

   A user role dialog box opens.

4. Select one or more Category user roles to assign to the user. For details on user roles, see .

   The selected roles are displayed in the **Permissions** column. If you assign more than one role to a user, the role with the highest permission level takes precedence.

5. When you finish adding users and setting permissions, click **OK**.

   The security settings apply and are inherited by any subcategories.

Set security on policies and repositories

Repositories and Policies inherit their security settings from the global security settings, however, you can set permissions on individual repositories or policies, which override the global settings.

To set security on an individual repository or policy

1. On the **Repositories** or **Policies** dashboard, click the menu icon (☰) on the repository or policy panel.

2. Click **Security**.

   The Security dialog box opens.

3. (*Optional*) **To add a user**

   a. Click **Add**.

   b. Enter the name of the user in the text box, and the click ☑ to verify the name against the

LDAP server. You can only add valid user names.

Add as many users as you require.

4. To edit user permissions, click **Edit** by the desired user.

A user role dialog box opens.

5. Select one or more Repository or Policy user roles to assign to the user. For details on user roles, see User roles, on page 17.

The selected roles are displayed in the **Permissions** column. If you assign more than one role to a user, the role with the highest permission level takes precedence.

6. Click **OK**.

The security settings apply.

File level security

There are differences in the file level security between Metadata repositories and HPE IDOL repositories.

- For individual repository content access, repository level roles take precedence over Storage Optimizer system level roles. See Individual repositories.

  Repository content display is based on NT security but policy assignment respects the Policy roles.

  > **NOTE:**
  > Repository owners of Analyzed repositories can view all documents in the repository. For repositories that are not in the Analyzed state, the Repository owner's view of the documents is based on NT security.

- For repository group content access, Storage Optimizer system level roles take precedence over individual repository level roles. See Repository groups, on page 23.

Individual repositories

> **NOTE:**
> Users and permissions are captured for File System repositories only and may not be available for any other type of repositories.

Metadata-based repository

Configuration setting (EnableSecureContent)	ControlPoint Deployment security enabled	Role	Has right to file/folder	View Contents	Access (result)
N	Y	Any role	Y	Y	Y (as expected)
	Y	Any role	N	Y	Y (as expected)
Y	Y	Repo Viewer	Y	Y	Y

Metadata-based repository, continued

Configuration setting (EnableSecureContent)	ControlPoint Deployment security enabled	Role	Has right to file/folder	View Contents	Access (result)
	Y	Repo User	Y	Y	Y
	Y	Repo Owner	Y	Y	Y
	Y	Repo Coordinator	Y	Y	Y
	Y	Repo Manager	Y	Y	Y
	Y	Repo Admin	Y	Y	Y
	Y	Repo User	N	N	N
	Y	Repo Viewer	N	N	N
	Y	Repo Owner	N	N	N
	Y	Repo Coordinator	N	Y	Y
	Y	Repo Manager	N	Y	Y
	Y	Repo Admin	N	Y	Y

Content-based repository

Configuration setting (EnableSecureContent)	ControlPoint Deployment security enabled	Role	IDOL security enabled	Has right to folder?	View contents	Access (result)
N	N		N	N	Y	Y (as expected)
	Y	Repo Viewer	Y	N	Y	Y (as expected)
	Y	Repo Admin	Y	N	Y	Y (as expected)
Y	Y	Repo	Y	N	N	N (as

Content-based repository, continued

Configuration setting (EnableSecureContent)	ControlPoint Deployment security enabled	Role	IDOL security enabled	Has right to folder?	View contents	Access (result)
		Viewer				expected)
	Y	Repo User	Y	N	N	N (as expected)
	Y	Repo Owner	Y	N	N	N (as expected)
	Y	Repo Coordinator	Y	N	Y	Y (as expected)
	Y	Repo Admin	Y	N	Y	Y (as expected)
	Y	Repo Manager	Y	N	Y	Y (as expected)
	Y	Repo Viewer	Y	Y	Y	Y
	Y	Repo User	Y	Y	Y	Y
	Y (Repo Admin)	Repo Admin	Y	Y	Y	Y (as expected)
	Y (Repo Viewer)	Repo Viewer	N	Y	Y	Y (as expected)
	Y (Repo Admin)		N	Y	Y	Y (as expected)

Repository groups

The following describes how documents are displayed in repository groups and subgroups for the various system level roles.

Repository users and repository viewers cannot see documents when they select a repository group level. To see documents, navigate to the sub-repositories.

Repository groups

System level role	Respect NT security?	Document displayed in repository group?	Document displayed in subgroup of repository?
Repo Coordinator Repo	N	Y	All

Repository groups, continued

System level role	Respect NT security?	Document displayed in repository group?	Document displayed in subgroup of repository?
Manager Repo Admin			
Repo Owner	Y - when the repository group is not Analyzed. No - when the repository group is Analyzed.	N	Same as Repository User/Viewer when analyzed Same as Repository Coordinator/Manager/Administrator when analyzed
Repo User Repo Viewer	Y	N	Based on sub repository's level role of the user.

Domain groups and Storage Optimizer

The Storage Optimizercomponents query Active Directory to get all of the domain groups a user belongs to, and uses that to determine if documents can be viewed.

In order to improve performance, Storage Optimizer caches the domain groups in **ControlPoint.dbo.CPCacheUserSecurity** for a certain amount of time. Storage Optimizer uses the cached groups for security instead of querying Active Directory every time.

- The `SecurityCacheTimeOut` setting in `\Dashboard\Web.config` defines the valid time period of the groups cache. The default is set to 1 hour.

  To force ControlPointStorage Optimizer to query Active Directory, you can delete the cache record in **ControlPoint.dbo.CPCacheUserSecurity** for a domain user so that the next time that user logs in, it must query Active Directory.

- The `EnableAddDomainAdmin` setting in `\Dashboard\Web.config` is used for when the domain user belongs to Administrators group, manually add DOMAIN ADMIN group to groups list.

  You can turn if off based on your Active Directory setup.

Chapter 3: HPE Storage Optimizer connectors

This section provides information on the supported HPE Storage Optimizer connectors.

> **NOTE:**
> Both Connector and CFS should be run by users with access to the data that needs to be analyzed. Furthermore, all access rights should be given to users running both these services. This is applicable to all connectors.

- Security mapping considerations for connectors
- Mapped security using OmniGroupServer (OGS)
- Enforce connector security
- HPE Storage Optimizer Exchange Web Service connector
- HPE Storage Optimizer Edge Filesystem connector
- HPE Storage Optimizer File System connector
- HPE Storage Optimizer Hadoop connector
- HPE Storage Optimizer Notes connector
- HPE Storage Optimizer SharePoint 2007 connector
- HPE Storage Optimizer SharePoint 2010 connector
- HPE Storage Optimizer SharePoint 2013 connector
- HPE Storage Optimizer SharePoint Remote connector

Security mapping considerations for connectors

The Securityinfo returned by Community and used by HPE Storage Optimizer only contains the user account.

If you are working with files where the permission is assigned by group name, it will not work by default. This is because the HPE IDOL index only stores the group names. When HPE Storage Optimizer uses the Securityinfo, the user name will not match with the group name, so nothing is returned.

The solution is to use the OmniGroupServer (OGS) to retrieve the security groups. OGS needs to be configured to refresh LDAP and combine on a 24 hour cycle.

For example, configure a task for OGS to refresh LDAP and combine on a 24 hours cycle using the following parameters

```
GroupServerRepeatSecs=86400
```

`GroupServerCycles` set to `-1` (cycles indefinitely).

Example

```
[Default]
//Default settings - these can also be set per repository
GroupServerStartTime=now
GroupServerCycles=0
```

```
GroupServerRepeatSecs=86400
GroupServerCaseInsensitive=TRUE
GroupServerMaxDataStoreQueue=5000
GroupServerIncremental=TRUE

[Repositories]
//Comma separated list of repositories which are queried when the repository
parameter not specified in action
GroupServerDefaultRepositories=Combine
Number=3
0=SharePoint2010
1=LDAP
2=Combine

[SharePoint2010]
GroupServerJobType=Connector
ConnectorHost=localhost
ConnectorPort=7024
ConnectorTask=GroupTask

[LDAP]
GroupServerLibrary=ogs_ldap.dll
LDAPServer=localhost
LDAPPort=389
LDAPBase=DC=home,DC=david, DC=local
LDAPType=MAD
LDAPBindMethod=NEGOTIATE
GroupServerCycles=-1

[Combine]
GroupServerJobType=Combine
GroupServerSections=SharePoint2010,LDAP
GroupServerStartDelaySecs=10
GroupServerCycles=-1
```

Mapped security using OmniGroupServer (OGS)

This section details the steps to capture the security groups for various HPE Storage Optimizer connectors using OmniGroupServer (OGS).

You must configure the following components:

1. **HPE IDOL Server.** You configure the HPE IDOL Server to process the security information contained in each document. You must also configure user security so that IDOL sends user and group information to the front-end application when a user logs in. See Configure HPE IDOL Server, on the next page.

2. **Connector (CFS).** You configure the connector to include security information, Access Control Lists (ACLs) in the documents that are indexed into HPE IDOL Server. You must also add a field to each document that identifies each security type. See Configure the HPE Storage Optimizer

3. **OmniGroupServer (OGS).** You configure OGS to retrieve and then combine connector and NT or LDAP group information. OGS retrieves connector information by sending the SynchronizeGroups action to the connector. OGS extracts NT security information directly from Active Directory. See .

Configure HPE IDOL Server

To integrate with the OGS, the IDOL Server must update the specific connector with an OGS Repository (LDAP, SharePoint2010 and so on) along with the OGS Server information.

The following is a sample IDOL configuration file where both the SharePoint and NT (file system) sections are updated with the OGS information.

The required settings are indicated by bold font.

ControlPoint IDOL.cfg:

```
[NT]
GroupServerHost=OGSHost
GroupServerPort=OGSPort
GroupServerRepository=LDAP

[LDAP]
LDAPServer=LDAPServerHost
LDAPPort=389

[SharePoint]
GroupServerHost=OGSHost
GroupServerPort=OGSPort
GroupServerRepository=Combine
```

> **NOTE:**
> After specifying these configuration settings, restart the HPE IDOL services.

Configure the HPE Storage Optimizer connector

OGS imports the security group information from the connectors by triggering a fetch action called SynchronizeGroups. As long as the connector has the OGS information which includes the hostname, port and repository, the connector can upload this information into the OGS Server.

The following are several samples of connector configuration files:

ControlPoint FileSystem Connector.cfg

```
[Ingestion]
IngestActions=META:ENFORCESECURITY=True,META:CPREPOSITORYTYPEID=3,META:SECURITYTYPE
=NT

[FetchTasks]
MappedSecurity=True
GroupServerHost=OGSHost
```

```
GroupServerPort=OGSPort
GroupServerRepository=NT
```

```
[TaskRepoTest]
IngestActions=
META:ENFORCESECURITY=True,Meta:SECURITYTYPE=NT
,META:CPREPOSITORYTYPEID=3,META:CPINDEXINGTYPE=1,META:AUTN_NO_
EXTRACT=true,META:AUTN_CATEGORIZE=false,META:AUTN_EDUCTION=false
```

SP 2010 Connector.cfg

```
[Ingestion]
IngestActions=META:ENFORCESECURITY=True,META:CPREPOSITORYTYPEID=2,META:SECURITYTYPE
=SharePointSecurity
```

```
[FetchTasks]
Number=1
GroupServerHost=OGSHost
GroupServerPort=OGSPort
GroupServerRepository=sharepoint2010
SynchronizeGroupsNumber=1
SynchronizeGroups0= Groups_TaskSP2010
EncryptACLEntries=False
MappedSecurity=TRUE
```

```
[Groups_TaskSP2010]
SecurityType=SharePointSecurity
IngestActions=META:CPREPOSITORYTYPEID=2,META:AUTN_NO_EXTRACT=true,META:AUTN_
CATEGORIZE=false,META:AUTN_EDUCTION=false
```

```
[TaskSP2010]
SecurityType=SharePointSecurity
IngestActions=
META:SecurityType=SharePointSecurity,META:ENFORCESECURITY=TRUE
,META:CPREPOSITORYTYPEID=2,META:AUTN_NO_EXTRACT=true,META:AUTN_
CATEGORIZE=false,META:AUTN_EDUCTION=false
```

> **NOTE:**
> After specifying these configuration settings, restart the Connector services.

Configure OmniGroupServer

OmniGroupServer (OGS) extracts the security group information based on the configurations defined for its repositories. Each repository has its security groups imported directly from the servers or through the connectors.

Security groups can be merged using the Combine parameter, as shown in the sample configuration file below. The LDAP groups can be merged with the SharePoint groups, which can be used by HPE Storage Optimizer to apply security at the repository level.

ControlPoint OGS.cfg

```
[Repositories]
```
GroupServerDefaultRepositories=Combine
```
Number=3
0=Sharepoint2010
```
1=LDAP
2=Combine

```
[LDAP]
GroupServerLibrary=ogs_ldap.dll
```
LDAPServer=LDAPServerHost
LDAPPort=389
```
LDAPBase=DC
LDAPType=MAD
LDAPBindMethod=NEGOTIATE
```
GroupServerAllUserGroups=NT AUTHORITY\AUTHENTICATED USERS
```
GroupServerCycles=-1
UseDomainPrefix=True
DomainPrefix=DOMAINPREFIX
```

```
[Sharepoint2010]
GroupServerLibrary=ogs_text.dll
GroupServerIncremental=TRUE
Textfile=SharePoint2010GS.txt
```
ConnectorHost=SharePointConnectorHost
ConnectorPort=SharePointConnectorPort
```
ConnectorTask= Groups_TaskSP2010
```

```
[Combine]
```
GroupServerJobType=Combine
GroupServerSections=LDAP,Sharepoint2010
```
GroupServerStartDelaySecs=10
GroupServerCycles=-1
```

> **NOTE:**
> After specifying these configuration settings, restart the OmniGroupServer services.

Enforce connector security

By default, all users in HPE Storage Optimizer are able to view the metadata of all items, regardless of HPE IDOL security permissions.

The `SecureMetaStoreContent` setting in `Dashboard\Web.config` controls the view and download options, depending on the IDOL security.

To enforce security

1. Navigate to the following location:

   ```
   \Program Files\Hewlett Packard Enterprise\Storage
   Optimizer\Dashboard\web.config
   ```

2. Locate the `<appSettings>` section.

3. Edit the "`SecureMetaStoreContent`" value from "`false`" to "`true`".

   Example

   ```
   <appSettings>
     <add key="SecureMetaStoreContent" value="true"/>
   </appSettings>
   ```

4. Save the file.

HPE Storage Optimizer Exchange Web Service connector

Summary

The Exchange Web Service (WS) connector can be used to analyze and execute policy on messages, appointments, contacts, and other items from an Exchange server.

The following versions of Exchange are supported:

- Microsoft Exchange 2007 SP1
- Microsoft Exchange 2010
- Microsoft Exchange 2010 SP1
- Microsoft Exchange 2010 SP2
- Microsoft Exchange 2013

Supported capability

The following policy types can be executed on content in an Exchange repository:

Dispose	Yes
Secure Leave	Yes
Secure Remove	Yes
Secure Shortcut	No

It is not possible to use an Exchange location as a target location.

DeployTool configuration

When selecting an Exchange Connector for inclusion in the deployment package, you must enter the following parameters:

Number of connectors in Group	The number of connectors to include in the deployment package. Each connector is configured to be in the same connector group.

Deployment Host(s)	The host(s) on which to deploy the connector(s). When more than one connector has been specified, the following are examples of valid entries for this field: • ServerA – all connectors are deployed to this server. • ServerA,ServerB – one connector is deployed to ServerA and the remainder to ServerB. • ServerA,ServerA,ServerB – two connectors are deployed to ServerA and the remainder to ServerB.
LDAP and Exchange Web Service User Domain	Domain for account to be used when accessing Active Directory and Exchange web services.
LDAP and Exchange Web Service User Username	Account to be used when accessing Active Directory and Exchange web services.
LDAP and Exchange Web Service User Password	Password for account to be used when accessing Active Directory and Exchange web services.

Configure Exchange WS Connector post deployment

When configuring an Exchange connector in HPE Storage Optimizer DeployTool, you must enter a user, domain, and password. This account is used by default to access both Active Directory and the Exchange web services. It is possible to configure different accounts manually by setting the following parameters in the connector configuration file:

`LDAPUsername` and `LDAPPassword` to specify the account to be used to access Active Directory.

`WSUsername`, `WSPassword`, and `WSDomain` to specify the account to be used to access Exchange web services.

Consider the following for the account used to access Exchange web services:

1. The account must have its own mailbox.

2. The account must have permission to retrieve information from other user's mailboxes using one of the following methods:

   • Enable impersonation rights. You must grant the user the permission to impersonate other users. In addition, you must set `ImpersonateMailboxOwner` to true in the connector configuration file.

   • Grant the user full access permission to each mailbox to be managed or analysed. You must set `ImpersonateMailboxOwner` to false in the connector configuration file.

   • Grant the user "Full Details" read access to each folder in each mailbox to be managed or analysed, including all folders below the root of the mailbox. You must set `ImpersonateMailboxOwner` to true in the connector configuration file.

Consider the following script as an example of how to change permissions and what permissions might be needed:

```
## save and run as createuser.ps1

## Read input from shell

$newusername = Read-Host "Enter New User Name"
$newemail = Read-Host "Enter New User Email Address"
$password = Read-Host "Enter Password For New User" -AsSecureString

## Create User and Mailbox

# Password can expire, change to not expire in user settings if corporate
policy allows

New-Mailbox -Name $newusername -Alias $newusername -UserPrincipalName $newemail
-SamAccountName $newusername -Password $password -DisplayName $newusername -
ResetPasswordOnNextLogon $false

## Add to groups (Some errors are expected for alternate version)
# Exchange 2010/2013 - add user to groups

Add-RoleGroupMember "Organization Management" -Member $newusername
Add-RoleGroupMember "Public Folder Management" -Member $newusername

#Exchange 2007 - add user to groups

Add-ExchangeAdministrator -identity $newusername -Role orgadmin
Add-ExchangeAdministrator -identity $newusername -Role publicfolderadmin

## Grant permissions and revoke denies if present

Get-ExchangeServer | Add-ADPermission -User $newusername -accessrights GenericRead,
GenericWrite -extendedrights Send-As, Receive-As, ms-Exch-Store-Admin -
Confirm:$False
Get-ExchangeServer | Remove-ADPermission -User $newusername -Deny -ExtendedRights
Receive-As -Confirm:$False
Get-MailboxDatabase | Add-ADPermission -User $newusername -AccessRights
ExtendedRight -ExtendedRights Receive-As, ms-Exch-Store-Admin -Confirm:$False

## For Forms registration

Get-PublicFolder -recurse | Add-PublicFolderClientPermission -User $newusername -
AccessRights Owner -Confirm:$False

## Some environments require additional security (Uncomment if needed)

# Get-Mailbox | Add-MailboxPermission -user $newusername -AccessRights FullAccess
```

The script may generate some errors, displayed in red or yellow text. Some errors are expected. The Mailbox Management user account is created using Exchange Management Shell.

The following settings can be adjusted manually by editing the connector configuration file. More information is provided in the HPE IDOL *Exchange Connector (CFS) Administration Guide* or by accessing the following URL:

`http://<Connector host>:7600/a=help`

Setting	Section	Description
DeleteMode	Default	Set to 2 if you want items removed by a Dispose policy to be moved to the user's Deleted Items folder. By default, items are permanently deleted.
ExchangeVersion	TaskName, FetchTasks or Default	Set to Exchange2010_SP1, Exchange2010, or Exchange2007_SP1 if you are using an early version of Exchange. The default setting is Exchange2010_SP2.
ImpersonateMailboxOwner	TaskName, FetchTasks or Default	Set to true if you are configuring the account used to access Exchange web services to have impersonation rights. Otherwise, do not set, or set to false (default).
LDAPPassword LDAPUsername	TaskName or FetchTasks	By default, the user specified by the Username setting (or the identity that the connector is running under when not set) is used when running LDAP queries against active directory. Set these parameters when a different user must be used for AD access. The password field can be encrypted.
Username, Password, Domain	TaskName, FetchTasks or Default	Specifies the user to be used to access both Exchange web services and Active Directory. Can be overridden by LDAPUsername or WSUsername. If no user is defined, the identity that the connector is running under is used.
WSDomain WSPassword WSUsername	TaskName, FetchTasks or Default	By default, the user specified by the Username setting (or the identity that the connector is running under when not set) is used when authenticating against the Exchange web service. Set these parameters when a different user must be used for authentication. The password field can be encrypted.

Add a new repository of type Exchange

When adding a new repository of type Exchange, the following parameters must be supplied:

Webservice URL	The URL of the Exchange web service
LDAP Path	The LDAP path to search for users with mailboxes to analyze

If Default Authentication is set to NO, the following additional parameters must be supplied. The credentials specified here are used for this repository in place of the details entered when the connector was configured in DeployTool.

Domain	The domain of the user specified by Username
Username	The user name to use to connect to LDAP and Exchange web services
Password	The password to use to connect to LDAP and Exchange web services

HPE Storage Optimizer Edge Filesystem connector

Summary

The Edge Filesystem connector is used to run Archive policies on documents and files held in Windows and Linux file shares.

Supported capability

Adding an Edge Filesystem repository allows you to select **No Analysis** as the Analysis type. This option is only available for Edge Filesystem repositories and is not supported for other File System repositories. For more information, see Add a repository, on page 55.

Edge Filesystem repositories do not support applying policies to archived content.

Policy summary screen status

For Edge Filesystem Connectors using Archive stub policies and the **Direct Policy Execution** setting, the Policy summary screen in the Console does not display policy status.

For policy status, check the Edge Filesystem Connector logs.

Define a Direct Target Location

To define a Direct Target Location for Edge Filesystem repositories, you can archive to a path.

Path	The path of the disk location to be used for documents secured to this target location. For Windows, the path is the UNC path.

HPE Storage Optimizer File System connector

Summary

The file system connector can be used to analyze and execute policy on documents and files held in Windows file shares.

File share type for indexing	Notes
NetApp volumes on the following shares: • CIFS • NFS	Configure the File System Connector to run with the domain accounts that have complete access to the NetApp volume for scanning documents. **NOTE:** The File System Connector CIFS mapping for UNIX shares on NetApp storage only captures the Windows-based permissions.
NTFS	

Supported capability

Repository policy types

The following policy types can be executed on content in a repository of type File System:

Dispose	Yes
Secure Leave	Yes
Secure Remove	Yes
Secure Shortcut	Yes

Target location policy types

A file system location can be established as a target location for relevant policy types. The following policy types can utilize target locations of type File System:

Secure Leave	Yes
Secure Remove	Yes
Secure Shortcut	Yes

DeployTool configuration

When selecting a File System Connector for inclusion in the deployment package, you must enter the following parameters:

Number of connectors in Group	The number of connectors to include in the deployment package. Each connector is configured to be in the same connector group.
Deployment Host(s)	The host(s) on which to deploy the connector(s). When more than one connector has been specified, the following are examples of valid entries for this field:
	ServerA – all connectors are deployed to this server.
	ServerA,ServerB – one connector is deployed to ServerA and the remainder to ServerB.
	ServerA,ServerA,ServerB – two connectors are deployed to ServerA and the remainder to ServerB.

Configure File System connector

The following settings can be adjusted manually by editing the connector configuration file. More information is provided in the *HPE File System Connector (CFS) Administration Guide.* or by accessing the following URL:

http://<Connector host>:7200/a=help

Setting	Section	Description
ForceDelete	*TaskName* or FetchTasks	Set to False to prevent deletion of read-only files.
IngestIfLastAccessChanged	*TaskName* or FetchTasks	Set to True to ensure that last access time is kept up-to-date.

Add new repository of type File System

When adding a new repository of type File System, you must supply the following parameters:

UNC Path	The UNC path of the file share to be registered and managed or analyzed.

Define a target Location of type File System

When adding a new target location of type File System, you must provide the following settings:

UNC Target Folder	The UNC path of the disk location to be used for documents secured to this target location.

Last access dates

> **NOTE:** Recording updates to last access dates is typically disabled in Windows Server through the Windows registry for performance reasons. This can be changed using the `fsutil` utility.

To ensure last access time updates are recorded

Run the following from the command line:

```
fsutil behavior set disablelastaccess 0
```

To turn off last access time updates

Run the following from the command line:

```
fsutil behavior set disablelastaccess 1
```

A reboot must be performed for any changes to take effect.

> **NOTE:** See Windows documentation for your specific version of Windows before making changes to the last access date behavior.

Enforce File System connector security

By default, security is off for content from the File System connector.

To enforce the File System connector security, edit the `Dashboard\Web.config` setting as follows:

```
<appSettings>
  <add key="SecureMetaStoreContent" value="true"/>
</appSettings>
```

For Repository browsing, IDOL document security is used when the user is a Repository User or Repository Viewer. For any higher permissions, such as Repository administrator, IDOL security is not used; the user can browse all items.

If a user does not have sufficient IDOL security permissions, they will still be able to view the Metadata, but not be able to view the Content of a file.

For example, a user with the correct IDOL security permissions will have the **View** and **Download** buttons available when browsing a document.

1. Perform a full scan of the repository for the first time only after the connector configuration has been set in step 1.

Retrieving security group information for the File System connector

See Security mapping considerations for connectors, on page 25.

For detailed steps and examples of configuring OmniGroupServer to retrieve security group information, see *HPE File System Connector (CFS) Administration Guide.*

HPE Storage Optimizer Hadoop connector

Summary

The Hadoop connector can be used to analyse and execute policy on documents and files held in a Hadoop Distributed File System (HDFS).

Supported capability

Repository policy types

The following policy types can be executed on content in a repository of type Hadoop:

Dispose	Yes
Secure Leave	Yes
Secure Remove	Yes
Secure Shortcut	No

A Hadoop location can be established as a target location for relevant policy types. The following policy types can utilise target locations of type Hadoop:

Secure Leave	Yes
Secure Remove	Yes

DeployTool configuration

When selecting a Hadoop Connector for inclusion in the deployment package, you must enter the following parameters:

Number of connectors in Group	The number of connectors to include in the deployment package. Each connector is configured to be in the same connector group.
Deployment Host(s)	The host(s) on which to deploy the connector(s). When more than one connector has been specified, the following are examples of valid entries for this field: • ServerA – all connectors are deployed to this server. • ServerA,ServerB – one connector is deployed to ServerA and the remainder to ServerB. • ServerA,ServerA,ServerB – two connectors are deployed to ServerA and the remainder to ServerB.
Hadoop Root	Enter the root URI of the Hadoop file system to connect to when securing

Uri	documents to a target location of type Hadoop.
Hadoop Path	Enter the location in the file system to be used by default when securing documents to a target location of type Hadoop.

Configure Hadoop connector

The following settings can be adjusted manually by editing the connector configuration file. For more information, see the Connector Guide or access the following URL:

`http://<Connector host>:13200/a=help`

Setting	Section	Description
FileSystemPath	TaskName, FetchTasks or Default	The location in the file system where the connector starts looking for files. The connector retrieves files from the specified folder and all of its subfolders. The path you specify must begin with a forward slash (/). To retrieve files from more than one folder tree, you can specify a comma-separated list of paths.
FileSystemRootUri	TaskName, FetchTasks or Default	The root URI of the Hadoop file system to connect to.

Adding new repository of type Hadoop

When adding a new repository of type Hadoop, you must supply the following parameters:

Filesystem Root URI	Enter the root URI of the Hadoop file system to connect to.
Filesystem Path	Enter the location in the file system where the connector starts looking for files. The connector retrieves files from the specified folder and all of its subfolders. The path you specify must begin with a forward slash (/). To retrieve files from more than one folder tree, you can specify a comma-separated list of paths.

Defining a Target Location of Type Hadoop

When adding a new target location of type Hadoop, you must provide the following settings:

Connector Config Section	The configuration setting in the connector configuration file that contains details needed to secure documents to the Hadoop target location. The default value for the section name is DefaultTargetLocationConfig. This section contains the details entered in HPE Storage Optimizer DeployTool.
Hadoop Target Folder	A target folder to be used for documents secured to this target location. The value supplied must start with `hdf://` or `hdfs://` and cannot end with a / character.

HPE Storage Optimizer Notes connector

Summary

The Notes connector can be used to analyse and execute policy on messages, appointments, contacts and other items from a Notes server.

Supported capability

Repository policy types

The following policy types can be executed on content in a repository of type Notes:

Dispose	Yes
Secure Leave	Yes
Secure Remove	Yes
Secure Shortcut	No

It is not possible to use a Notes location as a target location.

DeployTool configuration

When selecting a Notes Connector for inclusion in the deployment package, you must enter the following parameters:

Number of connectors in Group	The number of connectors to include in the deployment package. Each connector is configured to be in the same connector group.
Deployment Host(s)	The host(s) on which to deploy the connector (s). When more than one connector has been specified the following are examples of valid entries for this field: • ServerA – all connectors are deployed to this server. • ServerA,ServerB – one connector is deployed to ServerA and the remainder to ServerB. • ServerA,ServerA,ServerB – two connectors are deployed to ServerA and the remainder to ServerB.

Configure Notes connector

Some settings can be adjusted manually by editing the connector configuration file. For more information, see the Connector Guide or access the following URL:

`http://<Connector host>:13300/a=help`

Adding a new repository of type Notes

When adding a new repository of type Notes, you must supply the following parameters:

Notes Server	The name of the Notes server containing the repository.
Notes Database Directory	The folder that contains the database to be managed or analysed.
Notes Database	The name of the Notes database that is to be managed or analysed.
Notes User ID File Name	The Notes user ID file to be used to identify the user for connecting to the Notes server.
Notes User ID Password	The password for the user to be used to connect to the Notes server. The password can be encrypted for secure storage.

HPE Storage Optimizer SharePoint 2007 connector

Summary

The SharePoint 2007 connector can be used to analyse and execute policy on documents and files in SharePoint 2007 sites.

Supported capability

Repository policy types

The following policy types can be executed on content in a repository of type SharePoint 2007:

Dispose	Yes
Secure Leave	Yes
Secure Remove	Yes
Secure Shortcut	No

Target location policy types

A SharePoint 2007 location can be established as a target location for relevant policy types. The following policy types can utilize target locations of type SharePoint 2007:

Secure Leave	Yes
Secure Remove	Yes

DeployTool configuration

When selecting a SharePoint 2007 Connector for inclusion in the deployment package, you must enter the following parameters:

Number of connectors in Group	The number of connectors to include in the deployment package. Each connector is configured to be in the same connector group.
Deployment Host(s)	The host(s) on which to deploy the connector(s). When more than one connector has been specified the following are examples of valid entries for this field: ServerA – all connectors are deployed to this server. ServerA,ServerB – one connector is deployed to ServerA and the remainder to ServerB. ServerA,ServerA,ServerB – two connectors are deployed to ServerA and the remainder to ServerB.
SharePoint Host	The name of the web server or load balancer for the SharePoint farm or site.
SharePoint Port	The port number of the SharePoint web application.
SharePoint Credentials Username	The credentials to use to connect to the SharePoint web service (domain\username).
SharePoint Credentials Password	The password to use for the specified credentials. The password entered will be encrypted before addition to the configuration file.

Configure SharePoint 2007 connector

The settings can be adjusted manually by editing the connector configuration file. For more information, see the Connector Guide or access the following URL:

```
http://<Connector host>:7500/a=help
```

Adding New Repository of type SharePoint 2007

When adding a new repository of type SharePoint 2007, you must supply the following parameter(s):

SharePoint 2007 URL	The URL of the SharePoint location to be registered for analysis or management.

Defining a target location of type SharePoint 2007

When adding a new target location of type SharePoint 2007, you must provide the following settings:

Target URL	The URL of the SharePoint location to be used when securing documents to the target location.

HPE Storage Optimizer SharePoint 2010 connector

Summary

The SharePoint 2010 connector can be used to analyse and execute policy on documents and files in SharePoint 2010 sites.

Supported capability

Repository policy types

The following policy types can be executed on content in a repository of type SharePoint 2010:

Dispose	Yes
Secure Leave	Yes
Secure Remove	Yes
Secure Shortcut	No

Target location policy types

A SharePoint 2010 location can be established as a target location for relevant policy types. The following policy types can utilise target locations of type SharePoint 2010:

Secure Leave	Yes
Secure Remove	Yes

SharePoint 2010 connector Web service

After installing the SharePoint 2010 connector, the connector web service must be installed on a front-end SharePoint server. This web service provides the documents and metadata from SharePoint to the connector, returning more information than the default web services supplied with SharePoint.

> **NOTE:**
> An alternative to installing the connector web service is to use the SharePoint Remote connector. The SharePoint Remote connector requires no software to be deployed on the SharePoint server.

The connector web service is included in the webservice sub-folder of the installed SharePoint 2010 connector. Two versions of the web service are provided:

- `AutonomySharePoint2010Connector.wsp`

- `AutonomySharePoint2010ConnectorLight.wsp`

The Light version conforms with checks performed by Microsoft's SharePoint online code analysis framework (MSOCAF) tool. MSOCAF compatibility is a requirement when deploying solutions to Office 365.

> **NOTE:**
> If you deploy this version of the web service, you will be unable to configure it to return MHT files when collecting SharePoint pages.

Installing the Web service

The SharePoint connector web service must be installed manually.

To install the Web Service

1. Install the web service solution using the following commands from SharePoint 2010 Management Shell:

   `Add-SPSolution –LiteralPath <path to AutonomySharePoint2010Connector.wsp>`

   `Install-SPSolution -Identity AutonomySharePoint2010Connector.wsp -GACDeployment –Force`

2. Verify the successful deployment of the web service by accessing the following URL:

   `http://<site-collection server>/_vti_bin/autonomy/default.asmx`

Post-install steps

1. The user entered when configuring the SharePoint 2010 connector in HPE Storage Optimizer DeployTool must have Read access to all web applications.

2. The Application Pool identity under which the SharePoint web service is running must be the same as the Application Pool identity that SharePoint is running under.

3. If user profiles or activity feeds are to be analyzed, the application pool identity under which the SharePoint web service is running must also have full control permission to the User Profile

Service Application, in addition to the "Retrieve People Data for Search Crawlers" administrative right on the User Profile Service Application.

4. You must change the `web.config` manually to set `LocalStateDirectory` to an appropriate location for storing the connector web service state information. This location must be chosen based on the following:

   - Both the accounts running the connector

   - The application pool user needs to have full control to the local state directory

     The default location of the web.config for SharePoint 2010 is as follows:

     ```
     C:\Program Files\Common Files\Microsoft Shared\Web Server
     Extensions\14\ISAPI\Autonomy
     ```

     The section to be changed looks like this:

     ```
     <appSettings>

       <!-- Full path to a directory that can be used by the web service for
     storing state information.  -->

       <add key="LocalStateDirectory" value=" C:\Program
     Files\HPE\Sharepoint2010ConnectorCFS/localstate" />
     ```

     This default value is generally not appropriate due to Windows security restrictions. Therefore, select a value that is not under \Program Files.

DeployTool configuration

When selecting a SharePoint 2010 Connector for inclusion in the deployment package, you must enter the following parameters:

Number of connectors in Group	The number of connectors to include in the deployment package. Each connector is configured to be in the same connector group.
Deployment Host(s)	The host(s) on which to deploy the connector(s). When more than one connector has been specified, the following are examples of valid entries for this field: ServerA – all connectors are deployed to this server. ServerA,ServerB – one connector is deployed to ServerA and the remainder to ServerB. ServerA,ServerA,ServerB – two connectors are deployed to ServerA and the remainder to ServerB.
SharePoint Host	The name of the web server or load balancer for the SharePoint farm or site.
SharePoint Port	The port number of the SharePoint web application.

SharePoint Credentials Username	The credentials to use to connect to the SharePoint web service (domain\username).
SharePoint Credentials Password	The password to use for the specified credentials. The password entered will be encrypted before addition to the configuration file.

Configure SharePoint 2010 connector

The following settings can be adjusted manually by editing the connector configuration file. For more information, see the Connector Guide or access the following URL:

```
http://<Connector host>:7400/a=help
```

Setting	Section	Description
IndexActivityFeeds	TaskName or Default section	Set to false to disable analyzing of activity feeds.
IndexMinorVersions	TaskName or Default section	Set to true to enable analyzing of minor versions of list items and documents.
IndexUserProfiles	TaskName or Default section	Set to false to disable analyzing of user profiles.
IndexVersions	TaskName or Default section	Set to true to enable analyzing of previous major versions of list items and documents. By default, only the latest approved version of each list item and latest major version of each document are analyzed.
ProcessLists	TaskName or Default section	Set to false to disable analyzing of any list that is not a document library.

Each of the above settings can be added to an individual TaskName section to just influence the corresponding repository or to the Default section to be applied when scanning all repositories.

Configuring for MHTML file creation

SharePoint site pages can be collected in the MHTML (or MHT) format. This format combines the HTML code and all companion resources such as images into a single document and is a useful way to secure a faithful page rendition for file types such as .aspx.

By default, collecting such pages returns an IDX file, containing the metadata for the page, including visible text in fields such as DRECONTENT, LISTDESCRIPTION, or SP_WIKIFIELD, depending on the page type.

To return certain file types in MHTML format

- In the relevant TaskName section, set MhtFilterLua to the name of a LUA script. When a collect action is performed for a document, the connector calls the referenced Lua function passing it the collected document. If the Lua script returns true, the connector generates an MHT file with the

content of the web page.

Sample LUA script

```
function handler(document)
    sharePointObjectType =
     document:getFieldValue("BaseSharePointType");
    filetype = document:getFieldValue("FileType");
    --Use MHT files for all documents except where there is a
    --file that can be used instead (unless that file is a .aspx).
    if (sharePointObjectType == "Attachment"
        or sharePointObjectType == "ListItemVersion"
        or sharePointObjectType == "ListItem"
        and fileType ~= "aspx" and fileType ~= nil then
            return false
    end
    return true
end
```

Add a new repository of type SharePoint 2010

When adding a new repository of type SharePoint 2010 the following parameters must be supplied:

SharePoint 2010 URL	The URL of the SharePoint location to be registered for analysis or management.

Define a target location of type SharePoint 2010

When adding a new target location of type SharePoint 2010 the following settings must be provided:

Target URL	The URL of the SharePoint location to be used when securing documents to the target location.

Retrieve security group information for the SharePoint 2010 connector

See Security mapping considerations for connectors, on page 25.

For detailed steps and examples of configuring OmniGroupServer (OGS) to retrieve security group information, see *HPE Administration Guide for SharePoint 2010 Connector (CFS)*.

> **NOTE:**
> You will need multiple SharePoint server URLs in an OGS file if you have a different SharePoint farm. Each one of them needs to be a unique section in OGS.

The connector will also need to be configured to import group membership to OGS only on connecter service start. The SharePoint Connector pulls the group information from the entire farm. The StartURL has no impact. For a given WSHost defined in the Connector configuration file, you will only need one synchronize group call.

Example

```
[FetchTasks]
EncryptACLEntries=False
Number=0
GroupServerHost=localhost
GroupServerPort=4057
GroupServerRepository=Sharepoint2010
SynchronizeGroupsNumber=1

[Groups_TaskSPTest]
FetchMode=0
IncludeEmptyFields=True
ExtractSubfiles=True
MappedWebApplicationPolicies=True
IgnorePublishingPagesAspx=True
SecurityType=SharePointSecurity
IngestActions=META:ENFORCESECURITY=true,META:CPREPOSITORYTYPEID=2,META:AUTN_
CATEGORIZE=false,META:AUTN_EDUCTION=false
StartURL=http://localhost:90/
ScheduleStartTime=now
ScheduleCycles=1
ScheduleRepeatSecs=3600
IndexDatabase=SPTest
```

Troubleshooting

Troubleshooting the web service is covered in a chapter of the *HPE Administration Guide for SharePoint 2010 Connector (CFS).*

HPE Storage Optimizer SharePoint 2013 connector

Summary

The SharePoint 2013 connector can be used to analyse and execute policy on documents and files in SharePoint 2013 sites.

Supported capability

Repository policy types

The following policy types can be executed on content in a repository of type SharePoint 2013:

Dispose	Yes
Secure Leave	Yes

Secure Remove	Yes
Secure Shortcut	No

Target location policy types

A SharePoint 2013 location can be established as a target location for relevant policy types. The following policy types can utilize target locations of type SharePoint 2013:

Secure Leave	Yes
Secure Remove	Yes

SharePoint 2013 Connector Web Service

After installing the SharePoint web service, you must change the `web.config` manually to set LocalStateDirectory to an appropriate location for storing the connector web service state information. This location must be chosen based on the following:

- Both the accounts running the connector.

- The Application Pool user needs to have full control to the local state directory.

The default location of the `web.config` for SharePoint 2013 is as follows:

```
C:\Program Files\Common Files\Microsoft Shared\Web Server
Extensions\15\ISAPI\Autonomy
```

The section to be changed is as follows:

```
<appSettings>
!--  Full path to a directory that can be used by the web service for storing state
information.  -->
<add key="LocalStateDirectory" value=" C:\Program
Files\HPE\Sharepoint2013ConnectorCFS/localstate" />
```

This default value is generally not appropriate due to Windows security restrictions. Therefore, select a value that is not under `\Program Files`.

DeployTool configuration

When selecting a SharePoint 2013 Connector for inclusion in the deployment package, you must enter the following parameters:

Number of connectors in Group	The number of connectors to include in the deployment package. Each connector is configured to be in the same connector group.
Deployment Host(s)	The host(s) on which to deploy the connector(s). When more than one connector has been specified, the following are examples of valid entries for this field:

	ServerA – all connectors are deployed to this server. ServerA,ServerB – one connector is deployed to ServerA and the remainder to ServerB. ServerA,ServerA,ServerB – two connectors are deployed to ServerA and the remainder to ServerB.
SharePoint Host	The name of the web server or load balancer for the SharePoint farm or site.
SharePoint Port	The port number of the SharePoint web application.
SharePoint Credentials Username	The credentials to use to connect to the SharePoint web service (domain\username).
SharePoint Credentials Password	The password to use for the specified credentials. The password entered will be encrypted before addition to the configuration file.

Configure SharePoint 2013 connector

The following settings can be adjusted manually by editing the connector configuration file. For more information, see the Connector Guide or access the following URL:

```
http://<Connector host>:7100/a=help
```

Setting	Section	Description
IndexActivityFeeds	TaskName or Default section	Set to false to disable analyzing of activity feeds.
IndexMinorVersions	TaskName or Default section	Set to true to enable analyzing of minor versions of list items and documents.
IndexUserProfiles	TaskName or Default section	Set to false to disable analyzing of user profiles.
IndexVersions	TaskName or Default section	Set to true to enable analyzing of previous major versions of list items and documents. By default, only the latest approved version of each list item and latest major version of each document are analyzed.
ProcessLists	TaskName or Default section	Set to false to disable analyzing of any list that is not a document library.

Each of the above settings can be added to:

- An individual TaskName section to just influence the corresponding repository, or
- The Default section to be applied when scanning all repositories.

Adding a new repository of type SharePoint 2013

When adding a new repository of type SharePoint 2013, you must supply the following parameters:

SharePoint 2013 URL	The URL of the SharePoint location to be registered for analysis or management.

Defining a target location of type SharePoint 2013

When adding a new target location of type SharePoint 2013, you must provide the following settings:

Target URL	The URL of the SharePoint location to be used when securing documents to the target location.

HPE Storage Optimizer SharePoint Remote connector

Summary

The SharePoint Remote connector can be used to analyze and execute policy on documents and files in Microsoft SharePoint 2010, 2013 and 2016 sites. This connector also offers limited capability for documents and files in SharePoint Online.

Prerequisites

This section lists the prerequisites for installing the SharePoint Remote Connector.

- Windows Identity Foundation 3.5 feature on Windows 2012

Supported capability

Repository policy types

The following policy types can be executed on content in a repository of type SharePoint 2010, 2013 and 2016 using the Remote connector:

Dispose	Yes
Secure Leave	Yes
Secure Remove	Yes
Secure Shortcut	No

Target location policy types

A SharePoint location can be established as a target location for relevant policy types. The following policy types can utilize target locations of type SharePoint 2010, 2013 and 2016, or SharePoint Online using the SharePoint Remote connector:

Secure Leave	Yes [1]
Secure Remove	Yes[3]

DeployTool configuration

When selecting a SharePoint Remote Connector for inclusion in the deployment package, you must enter the following parameters:

Number of connectors in Group	The number of connectors to include in the deployment package. Each connector is configured to be in the same connector group.
Deployment Host(s)	The host(s) on which to deploy the connector(s). When more than one connector has been specified, the following are examples of valid entries for this field: ServerA – all connectors are deployed to this server. ServerA,ServerB – one connector is deployed to ServerA and the remainder to ServerB. ServerA,ServerA,ServerB – two connectors are deployed to ServerA and the remainder to ServerB.
SharePoint Credentials Username	The user name for connecting to the SharePoint web service.
SharePoint Credentials Password	The password for the specified credentials. The password entered will be encrypted before addition to the configuration file.
SharePoint Credentials Domain	The domain name for the credentials to connect to the SharePoint web service.

Configure SharePoint Remote connector

The settings can be adjusted manually by editing the connector configuration file. For more information, see the Connector Guide or access the following URL:

```
http://<Connector host>:7800/a=help
```

Adding a new repository of type SharePoint Remote

When adding a new repository of type SharePoint Remote, you must supply the following parameters:

SharePoint URL	The URL of the SharePoint location to be registered for analysis or management.

Defining a target location of type SharePoint Remote

When adding a new target location of type SharePoint Remote, you must provide the following settings:

Target URL	The URL of the SharePoint location to be used when securing documents to the target location.

1
SharePoint 2010 limits the size of files that you can upload to 3 MB. However, you can change this limit, for example, by running Powershell commands on the SharePoint Server:

```
$ws =
[Microsoft.SharePoint.Administration.SPWebService]::ContentService
$ws.ClientRequestServiceSettings.MaxReceivedMessageSize = 104857600 #100MB
```

Chapter 4: Manage Repositories

The HPE Storage Optimizer Repositories dashboard allows you to create and manage repositories.

- Repositories
- Add a repository
- XML repositories
- Search repositories
- Create a repository group
- Edit repository settings
- Change repository status
- Re-scan a repository
- Delete a repository
- Create a repository subset
- View repository compliance

Repositories

HPE Storage Optimizer manages content that you scan into HPE Storage Optimizer MetaStore. Repositories provide a view of the data that HPE Storage Optimizer manages.

The Register Repositories scheduled task discovers new repositories and registers them in HPE Storage Optimizer. You can also add repositories manually.

Repositories allow you to:

- browse content, and view and assign policies manually
- analyze and clean up data in legacy repositories
- identify a set of documents that you want to isolate for analysis or to promote to a higher analysis level

> **NOTE:**
> HPE Storage Optimizer is limited to managing documents in defined repositories.

Different repositories support different types of policy actions. For more information on policy phase actions, see Policy phases, on page 66

For more information on policy phases supported by the different target location types, see Target locations.

Supported policy phases by repository

Repository	Dispose	Secure (Remove)	Secure (Leave)	Secure (Link Shortcut)
Exchange	x	x	x	
File System	x	x	x	x
Hadoop	x	x	x	

Supported policy phases by repository, continued

Repository	Dispose	Secure (Remove)	Secure (Leave)	Secure (Link Shortcut)
Notes	x	x	x	
SharePoint 2007	x	x	x	
SharePoint 2010	x	x	x	
SharePoint 2013	x	x	x	
SharePoint Remote	x	x	x	

For a list of connector documents, see Related documentation, on page 15.

Repository status

Repositories can have one of three possible statuses: Registered, Analyzed, or Managed. The available information and the actions you can perform on the repositories are determined by the status.

- **Registered** repositories have been registered , but have not been analyzed. The Repositories page displays some basic repository statistics, however, you can browse the repository content.

- **Analyzed** repositories are ready for statistical analysis and cleanup. The summary page displays detailed statistical information about the repository contents, and you can take a number of actions to clean up legacy data.

- **Managed** repositories are being managed by HPE Storage Optimizer policies. Like Registered repositories, the Repositories page displays the number of documents and disk space for each repository. You also browse repository content. This status is required if you want to automatically assign policy to content. See Apply policies automatically, on page 79.

You can change the repository status manually. See Change repository status, on page 59.

Add a repository

You can manually add a repository on the Repositories dashboard. Alternatively, the Register Repositories scheduled task automatically adds repositories and maps them to individual databases.

To add a repository

1. Ensure that the appropriate connector is configured.

   Administrators can configure connectors on the Settings page in the Administration dashboard.

2. On the **Repositories** dashboard, click **+**.

   The Add New Repository page opens.

3. In the **Details** section, specify the following information:

   - **Name**. Enter the repository name.

   - **Description**. Enter a description of the repository.

   - **Type**. Select the repository type. You must provide additional information, which varies depending on the type you select.

- **Connector**. Select the connector to use for data scan. You can accept the default or choose an alternative, if one is configured, so that you can manually load balance.

  The necessary settings are dynamically loaded into the **Details** section after you select a connector.

  > **NOTE:**
  > To use Archive policies, you must select the Edge Filesystem Connector. You can specify the Archive policies in the **Direct Policy Execution** field in the **Settings** section.

4. In the **Settings** section, specify the following information:

   - **Network paths**. Enter one or more UNC paths for the repository.

     Click **Add (+)** to add more than one path.

   - **Include files of type**. Enter one or more file extensions, separated by commas, to include in the repository.

   - **Exclude files of type**. Enter one or more file extensions, separated by commas, to exclude from the repository.

5. In the **Analysis** section, set the following properties:

   - **Analysis Type**. Select one of the following analysis types:

     ○ **No Analysis** does not analyze any item.

        > **NOTE:**
        > This Analysis type is only available for Edge Filesystem Connector repositories.

     ○ **Repository Metadata Only** (default) analyzes metadata from the repository, but does not include document-level metadata. This is the fastest setting and builds the smallest analysis. It is useful to detect duplicate files.

     ○ **Metadata Only** analyses repository and document-level metadata. Processing time is slightly longer than the Repository Metadata Only setting because each document is opened.

   - **Capture Permissions and Ownership**. Select whether item permissions and ownership details should be captured.

   - **Analyze Subitems**. Select whether to assign a Policy to subitems. Examples of subitems include documents within a `.zip` or `.pst` file.

6. (*Optional*) In the **Properties** section, add any required properties (see Custom properties, on page 124).

   a. Click **Add**.

      The Add Property dialog box opens.

   b. From the **Property** list, select the property to add.

   c. From the **Value** list, select one or more values to apply to the repository.

   d. Repeat for as many properties as required.

   e. Click **Save**.

7. In the **Schedule** section, specify the following information to define the repository schedule.

- **Start Time**. Specify a start time. **Now** is selected by default.
- **Cycle**. Specify the number of times to run the schedule. **Run Once** is selected by default.
- **Recur Every**. Specify the recurrence period. The default is **1 hour**.

8. Click **Save**.

    You receive a prompt to restart all affected services. Depending on your selections, you may need to restart one or more of the following services: the selected connector and the associated Connector Framework Service.

    After you restart the HPE Storage Optimizer services, the new repository appears on the Repositories dashboard.

Search repositories

If you have a large number of repositories, you can use the Repositories dashboard to sort and filter the repository list to find repositories of interest. You can create custom properties to increase your sorting and filtering options (see Custom properties, on page 124).

You can switch between panel and grid displays. The Panel Display contains more information, while the Grid Display allows you to view more repositories at a time.

To switch between panel and grid display

- On the **Repositories** dashboard, click:
  - ◦ to view repositories in panel display. This is the default view.
  - ◦ to view a grid display.

To sort the repository list

1. Configure one or more custom properties that apply to repositories. See Create a custom property, on page 124.
2. On the Policies dashboard, select one of the criteria from the **Sort By** list.

    The repositories sorted by the selected criteria.

To filter the repository list

1. (Optional) Configure one or more custom properties that apply to repositories. See Create a custom property, on page 124.

    The Repositories dashboard's menu bar displays the properties on the left.
2. Select a value from one or more of the filters.

    By default, you can filter by repository type. You can also filter by custom properties.

    The repository list is filtered. If you filter by property value, the list displays only the repositories that have matching values. Filters are cumulative: you can filter by type, then by one property (for example, *Department*), then by another property (such as *Region*), and so on.

To filter the repository list by text

1. On the **Repositories** dashboard, click ▼.

    The Filter dialog box appears.
2. Enter text in the **Filter** box, and then click **Filter**.

The repository list updates.

Create a repository group

You can create a group of repositories for data analysis. The individual repositories remain accessible and available for analysis separately. Repository groups can be useful to analyze data by department, geographic region, repository type, or any other characteristic.

To create a repository group

1. On the **Repositories** dashboard, click **+**.

   The Add New Repository page opens.

2. Click **New Group**.

3. Under **Details**, enter the following information.

   - **Name** is the name of the repository group

     Allowed characters are A-Z, a-z, 0-9, and _.

   - **Description** is a description of the repository group

   - **Repositories**. Add as many repositories as you require.

     a. Click **Add**.

        The Add New Repository dialog box appears.

     b. Select repositories from the **Type** and **Connector** lists.

     c. Click **Save**.

4. (*Optional*) In the **Properties** group, add any required custom properties. See Custom properties, on page 124.

   a. Click **Add**.

      The Add Property dialog box appears.

   b. From the **Property** list, select the property to add.

   c. From the **Value** list, select one or more values to apply to the repository.

   d. Repeat for as many properties as you require.

   e. Click **Save**.

5. Click **Save**.

   The group appears on the Repositories dashboard. A link icon appears on the panel to indicate that it is a group.

Edit repository settings

You can edit repository settings on the Repositories dashboard. The options available for editing depend on the repository type.

To change repository settings

1. On the **Repositories** dashboard or any details page, click the menu icon (≡).

2. Click **Edit**.

The Edit Repository page opens.

3. In the **Details** group, you can:

    - Edit the **Description** of the repository.

    - Edit supplementary information, such as Network Paths for FileSystem repositories or Web service URLs for Exchange repositories.

4. Edit the **Capture Permissions and Ownership** and **Analyze Subitems** settings by selecting **Yes** or **No**.

5. Edit any of the **Properties** or **Scheduling** group settings, as required.

6. Click **Save**.

    Depending on your selections, you may receive a prompt to restart affected services, such as HPE Storage Optimizer connectors and the connector framework.

    The repository updates with the new settings.

Change repository status

You can change the repository status at any time on the Repositories dashboard. When you add repositories, the repository status is set to Registered by default.

Repositories must have a Managed status if you want HPE Storage Optimizer to apply policies to the repository content automatically. See .

To change the status of a repository

1. On the **Repositories** dashboard or any details page, click the menu icon (≡).

    Depending on the current status of the repository, the menu options vary. For example, if the repository is Registered, you can either **Analyze** or **Manage** the repository.

2. Click the desired repository change.

    A confirmation dialog box appears.

    If you move a repository to a Managed state, you can set the following options:

    - **Automatic Policy Assignment**

    - **Allow Policy Execution**

3. In the confirmation dialog box, click **Yes**.

    The repository status changes, and it moves to the appropriate Repositories tab.

    > **NOTE:**
    > You may also want to use the Analyze Density Indicator to know the Analyzed content for each Analyzed level.
    >
    > To view this information, click the menu button in the upper-right corner of the repository and select **Refresh Totals**.

Re-scan a repository

You can re-scan repositories from the Repositories dashboard.

When you manually scan a repository, you can see the progress of the scanning operation on the Repository panel.

To re-scan a repository

1. On the **Repositories** dashboard or any details page, click the menu icon (≡).
2. Click **Re-Scan Repository**.
3. In the Re-Scanning dialog box, click:
   - **Incremental Re-Scan** to process newly added, changed (since the last scan), or removed documents.

Create a repository subset

You can create a subset of analyzed repository data to view analysis metrics of a small portion of the repository contents. For example, you may want to analyze all files of a specific type, a specific size, created during a certain date range, and so on.

You can create subsets from a single repository or a repository group. You can create subsets of subsets. Also, you can promote the subset to a higher analysis level.

To create a repository subset

1. Create a filtered list of the files to analyze.

   There are several ways to do this:
   - view a file list and apply any desired filters (see Filter lists , on page 101
   - view data by statistical analysis, by tag, or by any other method
   - combine the resulting file list with filters
2. Click **Actions** > **Create Subset**.

   The Create Subset dialog box appears.
3. Enter a **Name** and **Description**.
4. Select the **Potential Set** to use to identify ROT (redundant, obsolete, or trivial) data. In addition, use the **Analysis Type** option to select the analysis type.

   If **Analysis Type** is changed, analysis is not automatically triggered. However, if it is not changed, analysis is automatically triggered. You cannot set this to a lower Analysis type than the parent Repository.
5. Click **Save**.
6. In the confirmation dialog box, click **OK**.

   The subset appears in the Subsets tab and resembles a repository. You can analyze the subset in a similar manner.

View repository compliance

If a repository is in a Managed state, you can view its overall level of compliance with all relevant policies.

To view repository compliance

1. On the **Repositories** dashboard, open the **Managed** tab.

2. To view more details, click the repository.

   The repository details page opens.

3. Click the **Policy** tab.

   The Policy Compliance section lists the percentage compliance with each policy, and the Policy Assigned/Executed Items lists the total number of affected items.

XML repositories

HPE Storage Optimizer provides the opportunity to add repositories for non-supported connectors, such as Documentum, FTP and so on.

- For more information on supported HPE Storage Optimizer connectors, see HPE Storage Optimizer connectors, on page 25.

- For more information on HPE IDOL connectors, see the *Administration Guide* for the specific IDOL connector.

By modifying configuration files and running a few processes from the Storage Optimizer Administration Console, these repositories can hold the ingestion information and can be analyzed just like any file system connector repositories.

When HPE Storage Optimizer does not provide support for a repository type (such as Documentum), its Register Repositories task creates an XML repository whenever it discovers a database that has not been analyzed by a HPE Storage Optimizer connector.

Limitations

- No repositories should be created from the Administration Console until the successful completion of Registering the XML repositories.

  This affects the repository ID counters and blocks any repository creations initiated from the Administration Console.

- The Register Repository task only works if there is at least one document to be ingested into HPE IDOL for every repository created through this method.

- If you add a large number of `Task` sections for a single connector, then each repository goes into an ingestion queue for processing the Register Repository tasks.

- The XML structure defines the tree structure displayed in the Administration Console.

  You define the tree structure in the XML file. It describes the tree structure that is displayed when users browse the content.

  - If you do not provide the XML structure, you can view the content, but there is no hierarchical navigation, and you must filter the full list of content in the repository.

  - The XML structure configuration is static. You must periodically update the XML structure for a repository where the contents are changing over time.

    For assistance in creating the XML structure, contact HPE Support.

Define the tree nodes

Administrators can use the XML elements described in the following table to define caption-filter pairs for each node in the tree.

> **NOTE:**
> You must provide `FieldText` or `CategoryId` for each `StructureItem`, but not both.

Element	Description
`Title`	The node title.
`FieldText`	The `FieldText` to run when tree node is selected. Matching items appear in the panel to the right of the navigation tree.
`CategoryId`	The ID of the category whose results you want to display when the tree node is selected. Matching items appear in the panel to the right of the navigation tree.

Add an XML repository

You can manually add an XML repository on the Repositories dashboard.

Alternatively, the Register Repositories scheduled task automatically adds repositories and maps them to individual databases. For additional information and limitations, see XML repositories, on the previous page.

To add an XML repository

1. On the **Repositories** dashboard, click **+**.

   The Add New Repository page opens.

2. In the **Details** section, specify the following information:

   a. **Name**. Enter the repository name.

   b. **Description**. Enter a description for the repository.

   c. **Type**. Select the repository type as **XML**.

   d. **XML Source**. Specify whether to structure the repository from XML input or using an existing XML file.

      > **NOTE:**
      > XML input takes precedence.

   e. **XML File Path** or **XML String**.

      - Enter the XML file path if you selected **File** as the option under **XML Source**.

      - Enter the XML string if you selected **Input** as the option under **XML Source**.

3. (*Optional*) In the **Properties** section, add any required properties. See Custom properties, on page 124.

   a. Click **Add**.

      The Add Property dialog box appears.

      b. From the **Property** list, select the property to add.

      c. From the **Value** list, select one or more values to apply to the repository.

        Repeat for as many properties as required.

    4. Click **Save**.

      You receive a prompt to restart all affected services.

      Depending on your selections, you may need to restart one or more of the following services:

- the selected connector
- the associated Connector Framework Service

      After you restart the HPE Storage Optimizer services, the new repository appears on the Repositories dashboard.

Sample XML repository

The following sample XML defines a simple tree structure for an XML repository.

```
<?xml version="1.0" encoding="utf-8">
<Structure>
  <StructureItemFilter>
    <FieldText>WILD{\\v-qa2-connector\FS\Start*}:DREREFERENCE</FieldText>
  </StructureItemFilter>
  <StructureItems xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
xmlns:xsd="http://www.w3.org/2001/XMLSchema">
  <StructureItem>
    <Title>Financial documents</Title>
    <ItemType>1</ItemType>
    <StructureItemFilter>
      <FieldText>WILD{\\v-qa2-connector\FS\Fin*}:DREREFERENCE</FieldText>
    </StructureItemFilter>
  </StructureItem>
  <StructureItem>
    <Title>Contracts</Title>
    <ItemType>1</ItemType>
    <StructureItemFilter>
      <FieldText>WILD{\\v-qa2-connector\FS\Contracts\*}:DREREFERENCE</FieldText>
    </StructureItemFilter>
    <StructureItems>
     <StructureItem>
      <Title>US Region</Title>
      <ItemType>1</ItemType>
      <StructureItemFilter>
        <FieldText>WILD{\\v-qa2-
connector\FS\Contracts\US\*}:DREREFERENCE</FieldText>
      </StructureItemFilter>
     </StructureItem>
     <StructureItem>
      <Title>Europe Region</Title>
      <ItemType>1</ItemType>
```

```
      <StructureItemFilter>
        <CategoryId>45323567564345</CategoryId>
      </StructureItemFilter>
    </StructureItem>
  </StructureItems>
  </StructureItem></StructureItems>
</Structure>
```

Delete a repository

When you no longer require a repository, for example, when you finish analyzing a repository or consolidating data, you can delete it from HPE Storage Optimizer. If Policies are applied to documents in a Repository, you cannot delete the Repository.

- If no items in the repository have policy assignments in the executing state, the repository can be deleted. In this case, any policy assignments are also deleted from HPE Storage Optimizer. They either completely execute or completely fail. The audit still keeps a record of the execution.

- If some items in the repository have policy assignments in the executing state, the repository cannot be deleted. The check box to delete is not available, and a warning message appears.

To remove a repository

1. On the **Repositories** dashboard or any details page, click the menu icon (≡).

2. Click **Delete**.

   A confirmation message opens.

3. Click **Delete** to remove the repository or cancel to abort the action.

   > **NOTE:**
   > Deleting the repository does not remove any policy associations from the files that it contains (unless you selected that option in the confirmation message box).

Chapter 5: Manage policies

A policy defines the rules and actions to perform on registered repositories. The HPE Storage Optimizer Policies dashboard allows you to create and manage policies and policy templates for enterprise information management.

- Policies
- Create a policy template
- Create a policy from a template
- Create an Archive policy
- Create a policy
  - Temporary locations for policy execution
- Edit a policy
- Policy execution rules
- Apply policies
- Remove a policy from an item
- Policy summary

Policies

A policy defines the rules and actions to perform on information content. HPE Storage Optimizer policies can be defined to address a variety of requirements including:

- information retention and disposal of content in repositories
- information categorization and capture of business records to record repositories
- information categorization and capture of important business information to secure storage for archiving

The following items are examples of typical HPE Storage Optimizer policies.

- delete project files in one or more file shares if they still exist three years after the project closes
- secure correspondence relating to supplier contracts in a SharePoint site and then delete it five years after the date of creation

HPE Storage Optimizer also offers you the ability to archive and stub a file using the Archive policy. Content of the specified source file is copied to the archive location. The copied file is not an exact copy, as it contains additional information. Therefore, you cannot access the archived version of the file as if it were the source file. Instead, you continue to access the local file and it will behave as if the file is still local. After the archive file is created, the source file is modified and a reparse point is placed on the file. In addition, the file is changed to a sparse file. This essentially removes the main data stream from the file. For information on creating a policy based on the Archive policy template, see Create a policy from a template, on page 69.

HPE Storage Optimizer also offers you the ability to delete the stub file using **Delete Archive Policy**. For more information on deleting the Archive policy, see Delete Archive policy, on page 71.

In addition, HPE Storage Optimizer offers you the Archiving command line utility, which can be used to recreate a file or directory stub, rehydrate a stubbed file or directory, dump the reparse data contents of a

stubbed file, or delete the stubbed source file. For information on the Archive command line utility, see Archiving command line utility, on page 144.

Policy phases

A policy consists of one or more phases. Each phase defines an action to take on a document that the policy is assigned to and that meets certain rule criteria. Policy phases can occur on repositories or target locations. For additional information, see Repositories, on page 54 or Target locations, on page 85.

You can perform the following actions using HPE Storage Optimizer policy phases.

- **Dispose**. Remove the item from the repository.
- **Secure**. Secure the item in a target location. Any conflicts are detected and prevent the item from being copied or moved. The three possible actions for the source files are:
  - **Leave**. Create a copy of the original file in the target location, and the original file remains in the repository.
  - **Remove**. The file moves from the repository to the target location.
  - **Shortcut**. The file moves from the repository to the target location, and a shortcut remains in the repository.

    > **NOTE:**
    > Shortcut policies apply to File System target locations only.

Policy templates

You can use a policy template to store a partial policy definition that you can then use to create a policy. Templates are useful when you need several similar policies. For example, several disposal policies have different disposal dates or declaration policies with different target repositories. Any mandatory parameters that are not supplied in a template must be provided in the policy you build using that template.

You can store as much or as little information as required in a policy template. At minimum, you must store the template name.

You use a default template, or create or modify your templates on the Administration dashboard under **Template Management**.

Default templates

The following templates are available by default.

- **Archive**. Archives and stubs a file.
- **Delete Archive**. Deletes Archive items.
- **Delete with review**. Sends items for review, and then deletes them if approved.
- **Delete without review**. Deletes items without first sending them for review.

Assign policies

You can assign policies:

- by assigning a category that is trained to match the content to which you want to assign the policy.
- from the HPE Storage Optimizer dashboard.

Assign Policies is a scheduled task that assigns policies based on category matches.

In certain cases, a category is retrained and content that initially matched the category may no longer match it. In such situations, any policies that were assigned based on the initial match are removed, however, this only applies to policies that have not executed or have executed actions that can be removed.

The Assign Policies task runs on a defined schedule. For more information, see Scheduled tasks, on page 111.

Execute policies

A policy can have one or more phases that execute in sequence or in parallel. Each phase has a name, action, execution rules, and a policy review definition.

You can apply policies to any document, however, if the policy phase has execution rules associated with it, the document must meet the criteria specified in the policy execution rules before the phase action executes. An example of an execution rule is: *five years after creation date*.

HPE Storage Optimizer checks for documents that are ready to execute (that is, that meet a policy phase execution rule) using a scheduled task that runs on a defined schedule. The execution rules are evaluated for each document and any rules that are satisfied start to execute.

There are multiple policy schedules that determine how frequently policies execute: Low, Medium, and High. For more information, see Schedule plans, on page 113.

In the first step of execution, HPE Storage Optimizer checks for policy conflicts. A policy conflict occurs when a policy phase is ready to execute on a document and one or more policies that have not executed are also assigned to that document. For example, a document may have a policy phase ready to execute with a disposal action *five years after creation date* and another policy has a disposal action *10 years after creation date*. All such conflicts must be resolved before execution continues.

When a policy phase executes, HPE Storage Optimizer performs the policy action on associated documents.

Create a policy template

You create a policy template from the Administration dashboard.

To create a policy template

1. On the **Administration** dashboard, click **Template Management**.

   The Template Management page opens.

2. Click **+**.

3. In the **Details** section, enter or select the following information:

   - **Name** of the template
   - **Description** is an optional description of the policy template
   - **Phases**. Specify a list of policy execution phases.

○ To specify one or more policy phases, click **Add**.

○ To group several policy phases together to run simultaneously, hover over the policy phase and click ▦. Select an item to group with this item.

○ To assign a different order to the policy phase sequence, hover over a policy phase and drag the entry to a new position.

Action	The action applies to the content when this phase executes.
	
	○ **Dispose** removes the document from the repository
	
	○ **Secure** secures the item in a target location. Any conflicts are detected and prevent the item from being copied or moved. There are three possible actions for the source files.
	
	  ■ **Leave** creates a copy in the target location and the original file remains in the repository
	
	  ■ **Remove** moves the file from the repository to the target location
	
	  ■ **Link Shortcut** moves the file from the repository to the target location and a shortcut remains in the repository

Name	Name of the phase.
Policy Review	Specifies whether items must be reviewed before HPE Storage Optimizer executes the associated Action. You can use the following values. ○ **System Default** (default) ○ **Review** ensures that HPE Storage Optimizer only applies the policy action after approval by an authorized user ○ **No Review**
Execution Rules	The criteria that the content must meet for HPE Storage Optimizer to apply the associated Action. ○ **Add Criteria** ○ **Begin Group** creates a group of conditions and specifies whether *all, any,* or *none* must be met. Click the down arrow to select an option. ○ **Repository Create Date** ○ **Repository Last Modified Date** ○ **Document Create Date** ○ **File Type**

4. In the **Settings** section, specify the following options.

- **Assign Policy** selects whether to enable the policy for assignment, and specifies when it will be available for assignment using the Date Options field.

- **Execute Policy** selects whether the HPE Storage Optimizer Engine checks the policy for items to execute.

- **Schedule Plan** selects how frequently HPE Storage Optimizer checks the policy for items to execute. The default values are:

- **High** every 10 minutes
- **Normal** every four hours. The default is Normal.
- **Low** every 24 hours

See Schedule plans, on page 113 for more details.

- **Priority** determines the priority of the policy that HPE Storage Optimizer uses during automatic conflict resolution. See Automatically resolve conflicts, on page 91

5. (*Optional*) In the **Properties** section, click **Add** to associate any properties and values that are appropriate for the policy. The properties are defined in **Administration > Settings > General > Properties**.

6. (*Optional*) In the **Assign To** section, click **Add** to select one or more categories. The policy will be assigned content associated with the selected categories.

7. Click **Save**.

Create a policy from a template

You can create a new policy from an existing policy template. A template provides some or all of the policy definition. You must provide any missing values in the policy definition.

To create a new policy from a policy template

1. On the **Policies** dashboard, click **+**.

   The Add Policy dialog box opens.

2. Select a template from the **Template** list, and then click **Continue**.

   The Add New Policy page opens.

3. In the **Details** section, specify the following information:

   - **Name** of the policy
   - **Description** for the policy
   - **Phases**. Specify a list of policy execution phases.
     - To specify one or more policy phases, click **Add**.
     - To group several policy phases together to run simultaneously, hover over the policy phase and click ⊞. Select an item to group with this item.
     - To assign a different order to the policy phase sequence, hover over a policy phase and drag the entry to a new position.

4. *(Optional)* In the **Settings** section, specify the following options.

   - **Assign Policy**
   - **Execute Policy** selects whether the HPE Storage Optimizer Engine checks the policy for items to execute.
   - **Schedule Plan**
   - **Compliance Policy**
   - **Priority** determines the priority of the policy. The policy with the highest priority executes on a document. The highest level of priority is 100. HPE Storage Optimizer uses this priority during automatic conflict resolution. See Automatically resolve conflicts, on page 91

5. (*Optional*) In the **Properties** section, click **Add** to associate any properties and values that are appropriate for the policy.

6. Click **Save**.

    The policy is saved.

Create an Archive policy

You can create a new Archive policy from an existing Archive policy template. A template provides some or all of the policy definition. You must provide any missing values in the policy definition.

To create a new Archive policy from the Archive policy template

1. On the **Policies** dashboard, click **+**.

    The Add Policy dialog box opens.

2. Select **Archive (Archives items)** from the Template list, and then click **Continue**.

    The Add New Policy page opens.

3. In the **Details** section, specify the following information:

    - **Name** of the policy
    - **Description** for the policy
    - **Archive Location** of the items to be archived.

        > **NOTE:**
        > This target location is defined in **Administration -> Target Locations**. See Target locations, on page 85.

4. In the **Execution Rules** section, specify the criteria that the content must meet for HPE Storage Optimizer to apply the associated Action.

    - **Add Criteria**
    - **Begin Group** creates a group of conditions and specifies whether *all, any,* or *none* must be met.
    - **Repository Create Date**
    - **Repository Last Modified Date**
    - **Document Create Date**
    - **File Type**

5. In the **Settings** section, specify the following options.

    - **Execute Policy** selects whether the HPE Storage Optimizer Engine checks the policy for items to execute.
    - **Priority** determines the priority of the policy. The policy with the highest priority executes on a document. The highest level of priority is 100. HPE Storage Optimizer uses this priority during automatic conflict resolution. See Automatically resolve conflicts, on page 91

6. (*Optional*) In the **Properties** section, click **Add** to associate any properties and values that are appropriate for the policy.

7. Click **Save**.

Delete Archive policy

There are two ways of applying a Delete Archive Policy. You can create a new repository or you can edit an existing archive repository.

Apply a Delete Archive policy while creating a new repository

1. Create a new **Delete Archive Policy**.

2. Create a new repository and select the **Edge File System Connector** from the list of connectors.

3. From the settings, select the newly created Delete Archive Policy.

4. Execute the policy.

Apply a Delete Archive policy while editing an archive repository

1. Create a new **Delete Archive Policy**.

2. Edit the existing archive repository.

3. From the settings, add the newly created Delete Archive Policy along with the existing archive policy.

4. Execute the policies.

Create a policy

For special cases, you can create a policy without using a policy template.

To create a policy

1. On the **Policies** dashboard, click **+**.

   The Add Policy dialog box opens.

2. Select **Blank (default)** from the Template list.

3. Click **Continue**.

   The Add New Policy page opens.

4. In the **Details** section, specify the following information:

   - **Name** of the policy
   - **Description** for the policy
   - **Phases**. Specify a list of policy execution phases.
     - To specify one or more policy phases, click **Add**.
     - To group several policy phases together to run simultaneously, hover over the policy phase and click ⊞. Select an item to group with this item.
     - To assign a different order to the policy phase sequence, hover over a policy phase and drag the entry to a new position.

Action	The action applies to the content when this phase executes.
	○ **Dispose** removes the document from the repository

	○ **Secure** secures the item in a target location. Any conflicts are detected and prevent the item from being copied or moved. There are three possible actions for the source files. ■ **Leave** creates a copy in the target location and the original file remains in the repository ■ **Remove** moves the file from the repository to the target location ■ **Link Shortcut** moves the file from the repository to the target location and a shortcut remains in the repository
Name	of the new phase.
Policy Review	specifies whether items must be reviewed before HPE Storage Optimizer executes the associated Action. This option ensures that HPE Storage Optimizer only applies the policy action after approval by an authorized user. You can use the following values. ○ **System Default** (default) ○ **Review** ○ **No Review**
Execution Rules	are the criteria that the content must meet for HPE Storage Optimizer to apply the associated Action ○ **Add Criteria** ○ **Begin Group** creates a group of conditions and specifies whether *all, any,* or *none* must be met. ○ **Repository Create Date** ○ **Repository Last Modified Date** ○ **Document Create Date** ○ **File Type**

5. In the **Settings** section, specify the following options:

- **Assign Policy** selects whether to enable the policy for assignment, and specifies when it will be available for assignment using the Date Options field.

- **Execute Policy** selects whether the HPE Storage Optimizer Engine checks the policy for items to execute.

- **Schedule Plan** selects how frequently HPE Storage Optimizer checks the policy for items to execute. The default values are:

  - **High** runs every 10 minutes

  - **Normal** runs every four hours. The default is Normal.

  - **Low** runs every 24 hours

  See Schedule plans, on page 113 for more details.

- **Priority** determines the priority of the policy, which HPE Storage Optimizer uses during automatic conflict resolution. See Automatically resolve conflicts, on page 91

- **Policy Approver Email Address** selects the email address of the policy approvers to be notified about the review before the policy execution.

- **Temp location** defines the shared network directory that is used for storing temporary files for Secure, Declare, or Declare in Place policy phases. For more information, see Define a temporary location for each policy, below

6. (*Optional*) In the **Properties** section, click **Add** to associate any properties and values that are appropriate for the policy.

7. (*Optional*) In the **Assign To** section, click **Add** to select categories. The policy will be assigned content associated with the selected categories.

8. When you finish adding phases, click **OK**.

Temporary locations for policy execution

Temporary locations for policy executions are shared network directories to store temporary files for the policy execution processes, such as Secure, Declare or Declare in Place policy phases. HPE Storage Optimizer release 5.5 introduces the feature of user-definable temporary locations for each policy.

Considerations

- When you save a policy, the Storage Optimizer Scheduler validates whether the temporary location is accessible from the Storage Optimizer Scheduler. The temporary location must be accessible from the Storage Optimizer Scheduler and the connectors.

  > **NOTE:**
  > The validation in the Add Policy page only checks for the accessibility from the Storage Optimizer scheduler. It does not check accessibility from the connectors. You will need to manually check the connector accessibility to the temporary locations.

  > **TIP:**
  > Using the same user account for Storage Optimizer scheduler service and for all connectors is a good way to ensure connector accessibility to temporary locations.

- As part of the policy execution, all temporary locations are regularly cleaned up. The cleaner deletes any subfolders that are older than the defined expiration period from the present for all temporary locations, including the default temporary location or the user-defined temporary location under any policies.

  The expiration time is defined under `Autonomy.ControlPoint.CollectCleanupTime` in the **GlobalSettings** table in the ControlPoint database. The default expiration time is 24 hours.

Define a temporary location for each policy

Unless otherwise defined on the Add Policy page, the Storage Optimizer Engine uses the default temporary location defined in the Settings page of the Administration Dashboard.

> **NOTE:**
> Adding a temporary location is only possible during policy creation.

> Once the policy is saved, the **Temporary Location** field is no longer modifiable. This restriction is to ensure that policy execution will not be affected by a change of temporary locations.

The temporary location is a field under the Settings area the Policy Execution page. For more information, see Create a policy, on page 71.

Search policies

If you have a large number of policies, you can use the Policies dashboard to sort and to filter the policy list to find the policies that you want. To sort and filter, you must configure custom properties that apply to policies. See Custom properties, on page 124.

To sort the policy list

1. Configure one or more custom properties that apply to policies. See Create a custom property, on page 124.
2. On the Policies dashboard, select one of the criteria from the **Sort By** list.

    The policies are sorted by the selected criteria.

To filter the policy list

1. Configure one or more custom properties to apply to policies. See Create a custom property, on page 124.

    The properties appear on the left of the menu bar of the Policy dashboard.
2. Select a value from one or more property filters.

    The policy list is filtered to display only the policies that have matching property values.

    Filters are cumulative, so you can filter by one property, for example, *Department*, and then filter the list by another property, such as *Region*, and so on.

To filter the policy list by schedule plan

-   On the Policies dashboard, select a Schedule Plan from the list.

    The policies are sorted by the selected schedule plan.

To filter the repository list by text

1. On the Policies dashboard, click ▼.

    A Filter dialog box opens.
2. Enter text in the **Filter** box, and then click **Filter**.

    The policy list updates.

Edit a policy

You can change the settings for a local policy from the Policies page.

To edit a policy

1. On the **Policies** dashboard, select a policy panel and click the menu icon (≡).

    The Edit Policy page opens.

2. In the **Details** section, change the following information as required.

   - **Name** of the policy

   - **Description** of the policy

   - **Phases**. Specify a list of policy execution phases or click the edit icon (✎) to edit a policy phase.

     ○ To specify one or more policy phases, click **Add**.

     ○ To group several policy phases together to run simultaneously, hover over the policy phase and click ▦. Select an item to group with this item.

     ○ To assign a different order to the policy phase sequence, hover over a policy phase and drag the entry to a new position.

Action	The action applies to the content when this phase executes. ○ **Dispose** removes the document from the repository ○ **Secure** secures the item in a target location. Any conflicts are detected and prevent the item from being copied or moved. There are three possible actions for the source files. ▪ **Leave** creates a copy in the target location and the original file remains in the repository ▪ **Remove** moves the file from the repository to the target location ▪ **Link Shortcut**. The file moves from the repository to the target location and a shortcut remains in the repository.
Name	of the new phase.
Policy Review	specifies whether items must be reviewed before HPE Storage Optimizer executes the associated Action. This option ensures that HPE Storage Optimizer only applies the policy action after approval by an authorized user. You can use the following values. ○ **System Default** (default) ○ **Review** ○ **No Review**
Execution Rules	The criteria that the content must meet for HPE Storage Optimizer to apply the associated Action. ○ **Add Criteria** ○ **Begin Group** creates a group of conditions and specifies whether *all, any,* or *none* must be met. ○ **Repository Create Date** ○ **Repository Last Modified Date** ○ **Document Create Date** ○ **File Type**

3. In the **Settings** section, change the following options as required.

- **Assign Policy** selects whether to enable the policy for assignment, and specifies when it will be available for assignment using the Date Options field.
- **Execute Policy** selects whether the HPE Storage Optimizer Engine checks the policy for items to execute.
- **Priority** determines the priority of the policy that HPE Storage Optimizer uses during automatic conflict resolution. See Automatically resolve conflicts, on page 91

4. (*Optional*) In the **Properties** section, edit any existing properties or values, or click **Add** to add new ones.

5. (*Optional*) In the **Assign To** section, edit any categories, or click **Add** to select new ones.

6. Click **Save**.

   Changes to some policy settings affect the documents to which HPE Storage Optimizer previously applied the policy.

Policy Setting	Effect
Execute Policy	If you deactivate a policy (that is, change the setting from *Yes* to *No*), HPE Storage Optimizer does not execute it again until you reactivate it. If the policy was previously active, the action taken by the policy remains.
Phases	If you add a new policy phase, it applies to all items that meet the rules associated with the policy from the time you add it. The new phase does not apply to existing items that HPE Storage Optimizer applied the previous phases to.
Execution Rules	Documents must meet the new rules for HPE Storage Optimizer to apply the action associated with the policy. The new rules do not change existing documents that HPE Storage Optimizer previously assigned to the policy.
Categories	HPE Storage Optimizer applies the policy to documents in the current category list. The policy association remains for documents in other categories that the policy previously applied to.

Edit a policy template

To create a policy template

1. On the **Administration** dashboard, click **Template Management**.

   The Template Management page opens.

2. Select a policy template panel.

   The Edit Policy Template page opens.

3. In the **Details** section, edit or select the following information:

- **Name** of the template
- **Description** is an optional description of the policy template
- **Phases**. Specify a list of policy execution phases.
  - To specify one or more policy phases, click **Add**.
  - To group several policy phases together to run simultaneously, hover over the policy phase and click ▦. Select an item to group with this item.
  - To assign a different order to the policy phase sequence, hover over a policy phase and drag the entry to a new position.

Action	The action applies to the content when this phase executes.
	○ **Dispose** removes the document from the repository
	○ **Secure** secures the item in a target location. Any conflicts are detected and prevent the item from being copied or moved. There are three possible actions for the source files.
	▪ **Leave** creates a copy in the target location and the original file remains in the repository
	▪ **Remove** moves the file from the repository to the target location
	▪ **Link Shortcut** moves the file from the repository to the target location and a shortcut remains in the repository
Name	Name of the phase.
Policy Review	Specifies whether items must be reviewed before HPE Storage Optimizer executes the associated Action. You can use the following values.
	○ **System Default** (default)
	○ **Review** ensures that HPE Storage Optimizer only applies the policy action after approval by an authorized user
	○ **No Review**
Execution Rules	The criteria that the content must meet for HPE Storage Optimizer to apply the associated Action.
	○ **Add Criteria**
	○ **Begin Group** creates a group of conditions and specifies whether *all, any,* or *none* must be met. Click the down arrow to select an option.
	○ **Repository Create Date**
	○ **Repository Last Modified Date**
	○ **Document Create Date**
	○ **File Type**

4. In the **Settings** section, specify the following options.
   - **Assign Policy** selects whether to enable the policy for assignment, and specifies when it will be available for assignment using the Date Options field.
   - **Execute Policy** selects whether the HPE Storage Optimizer Engine checks the policy for items to execute.

- **Schedule Plan** selects how frequently HPE Storage Optimizer checks the policy for items to execute. The default values are:
    - **High** every 10 minutes
    - **Normal** every four hours. The default is Normal.
    - **Low** every 24 hours

    See Schedule plans, on page 113 for more details.
- **Priority** determines the priority of the policy that HPE Storage Optimizer uses during automatic conflict resolution. See Automatically resolve conflicts, on page 91

5. (*Optional*) In the **Properties** section, click **Add** to associate any properties and values that are appropriate for the policy. The properties are defined in **Administration > Settings > General > Properties**.

6. (*Optional*) In the **Assign To** section, click **Add** to select one or more categories. The policy will be assigned content associated with the selected categories.

7. Click **Save**.

Policy execution rules

When you create a policy phase, you must create a set of rules for when to execute the policy phase action.

A policy phase that does not have execution rules executes immediately.

You can construct rules using field names, operators, and values.

Add Rule Builder fields

You add fields through the Rule Builder that is accessible through the Administration dashboard. The Rule Builder allows you to select fields and operators from lists. You can only select `Match`, `NumericDate`, or `Date` type fields. To use other fields, you must first define them as one of these types.

By default, the Rule Builder contains the following fields.

- Location (`CPLOCATION`) – This is shown on Category only.
- Repository Create Date (`AU_REPOSITORY_CREATEDDATE_EPOCHSECONDS`)
- Repository Last Modified Date (`AU_REPOSITORY_MODIFIEDDATE_EPOCHSECONDS`)
- Document Create Date (`CPDOCUMENT_CREATEDDATE_EPOCHSECONDS`)
- File Type (`IMPORTMAGICEXTENSION`)

To add a field

1. On the Administration dashboard, click the **Settings** panel.

    The Settings page opens.

2. On the General tab, click **Fields**.

    The Rule Builder page opens.

3. Under Details, click **Add**.

    The Add New Field dialog box appears.

4. Enter a **Display Name** for the field, click the **Field** box, and then select a field from the list.

5. Click **Save**.

   The rule appears in the Rule Builder Fields list.

Apply policies

You can apply policies to documents in HPE Storage Optimizer either manually or you can set up an automatic process to apply policies.

Apply policies automatically

You can automatically apply a policy by associating it with one or more server categories, either when you create the policy or when you edit it.

After you set up category associations, a scheduled task assigns policies to documents by category association. By default, this task executes every hour. You can change this frequency from the **Schedule Management** page that is accessible from the Administration dashboard.

You can also control policy assignment and execution at the repository level. Automatic policy assignment only applies to Managed repositories. To specify whether repositories support automatic policy assignment and execution, edit the repository, select the appropriate options, and then change the status to Managed.

Apply policies manually

You can apply policies to documents manually from any file list.

To apply a policy manually

1. Select one or more files to apply a policy to, and then click ⊠ .

2. A policy list opens, including all active and inactive policies.

3. Select one or more policies to apply to the documents, and then click **Apply**.

   HPE Storage Optimizer applies the policies to the documents.

Re-evaluate policy assignment based on category changes

Upon creating a new category, a policy associated with the new category, and a repository in the Managed state, you can run the Assign Policy scheduled task to assign the policy to documents based on the category matches. See Assign policies, on page 66 and Default scheduled tasks, on page 111.

Remove Policy Assignment scheduled task

When you need to change the category definition so that the new category matches a different set of files and you have assigned a policy based on the previous category definition, you run the Remove Policy Assignment scheduled task. It removes the policy assignments, based on the previous definitions, according to the following conditions:

- Documents are unassigned from the policy if the documents matching the previous category have not gone into the Executing state.

After you remove the policy assignment from documents using the Remove Policy Assignment scheduled task:

- You must wait 24 hours to use the Assign Policy scheduled task to assign the same policy to the same document again.

> **NOTE:**
> This limitation is to avoid documents being repeatedly assigned to the same policy based on category.

- You can still manually assign the same policy to the same document by using the HPE Storage Optimizer dashboard.

Expected behaviors

Consider the following interactions of the Remove Policy Assignment on different repository types.

1. If the repository type is **Content**, after you change the category definition:
   a. Run the Assign Policies task. This assigns the policy to documents that match the current category definition and does not match the previous category definition.
   b. Run the Remove Policy Assignment task after at least five (5) minutes. This removes the policy assignment from all documents that match the previous category definition.
   c. Run the Assign Policies task. This assigns the policy to the documents that were excluded in step 1a, which are documents that match both current and previous category definitions.

   For example:
   - Assume that an original category definition matches Documents A, B and C. Documents A, B, and C are assigned with a policy associated with the category.
   - An updated category definition matches Document B, C and D.
   - All documents are still in the Policy Assigned state.

   Results
   - Step 1a would assign to D.
   - Step 1b would remove the policy from Documents A, B and C.
   - Step 1c would assign to B, C.

   > **NOTE:**
   > This scenario also applies to Content repositories whose categories are trained by the following means:
   > - Boolean training text
   > - Training text
   > - Boolean training text and field text
   > - Training text and field text.

2. If the repository type is **Repository Metadata Only** or **Metadata Only**, after you change the category definition:

a. Run the Assign Policies task. This assigns the policy to documents that match the current category definition, regardless of whether the documents match previous category definitions.

b. Run the Remove Policy Assignment task after at least five (5) minutes. This removes the policy assignment from documents that match the previous category definition and does not match the new category definition.

For example:

- Assume that an original category definition matches Documents A, B, and C. Documents A, B, and C are assigned with a policy associated with the category.

- An updated category definition matches Document B, C and D.

- All documents are still in Policy Assigned state.

Results

- Step 2a would assign to D.

- Step 2b would remove the policy from A.

> **NOTE:**
> This scenario also applies to Content repositories whose categories are trained by the following means:
> - Field text

Limitations

- After you change the category definition, you must run the Assign Policies scheduled task once before you can run the Remove Policy Assignment scheduled task effectively.

  The time interval between running the Assign Policies and Remove Policy Assignment scheduled tasks must be at least five (5) minutes.

- The Remove Policy Assignment scheduled task is based on category versioning, not actual category matches. If you change the category definition such that the new category matches the same set of documents, the scheduled task will still remove the policy assignment because there is a new version of the category.

  After 24 hours, you can run the Assign Policy scheduled task to assign the same policy to the document again.

View the policies on items

The HPE Storage Optimizer Policies dashboard allows you to view:

- a summary of the policy information

- items that have a particular policy applied. This information can help you evaluate how widely and accurately a policy applies.

- policies that a particular item belongs to. This information can help you to identify whether a policy has been applied incorrectly or accidentally by a category match or a policy assigner.

View policy summaries

Policy summaries provide an overview of policy information.

To view a policy summary

- On the Policies dashboard, click the policy panel

  The summary page opens, displaying the date, number of policy items, the policy settings and phases, and the most common issues.

View items assigned to a policy

You can view all the items that a policy applies to. You can also apply a filter to view items that have a specific status.

To view items that are assigned to a particular policy

1. On the **Policies** dashboard, click the policy.

   The Policy summary page opens.

2. Click the **Policy Items** tab.

   The Policy Items page opens, displaying a list of policy items on the right of the page.

3. (Optional) To filter the view of items assigned to a particular policy

   - select one of the policy statuses from the box on the left of the page

     One or more of the following statuses may appear.

Status	Description
Policy Assigned	Items with a policy assigned to them, but whose execution criteria have not been checked.
Policy Executed	Items where the policy has executed.
Executing	Items that are currently being processed by the scheduler.
Awaiting Execution	Items that meet execution criteria, yet have not yet been executed.
Awaiting Review	Items that are ready to execute, but that require review before execution can proceed.
Awaiting Conflict Resolution	Items that are ready to execute, but that require you to resolve a conflict before execution can proceed.
Execution Rules not met	Items that do not meet the policy execution rules.
Execution Rejected	Items that were prevented from executing after review.
Prevented Due to Conflict	Items that were prevented from executing after resolution of a policy conflict.

View the policies that apply to an item

You can view a list of policies that have been assigned to an individual item, and see their current states.

To view a list of policies that apply to an item

1. Locate the item in the Repository dashboard view.
2. Select the check box beside the item.
3. Click **Actions > Properties**.

   The Properties dialog box opens.
4. Open the **Policies** tab.

View summary report of items processed by a policy

You can add a summary report of items processed by a policy to the Document Activity report generated by SQL Server Reporting Services.

> **NOTE:**
> To enable this report, your Storage Optimizer environment must have SQL Server Reporting Services configured with the Storage Optimizer data source.

Before you begin

1. Configure the Storage Optimizer data source in SQL Server Reporting Services.
2. During the ControlPoint database installation, select **Upload Reports** on the Audit Reports page of the HPE Storage Optimizer Database Installer wizard.

   For more information, see the *HPE Storage Optimizer Installation Guide*.

To enable the summary report

1. On the Administration dashboard, click **Settings**.

   The Settings page opens.
2. Click the **General** tab.
3. In the Details section, click **Yes** for the **Audit Activity** setting.
4. Specify any other settings as needed.
5. Click **Save**.
6. Navigate to the Storage Optimizer Reports site, select **English > Document Activity.rdl.**
7. Specify any necessary parameters, such as dates, and click **View Report.**

   The Document Activity report displays a summary report of items processed by a policy.

Remove a policy from an item

You can manually remove a policy from a document if, for example, it was incorrectly or accidentally assigned by a category match or a policy assigner.

To remove a policy from a document

1. On the **Policies** dashboard, click the policy.

   The Policy summary page opens.

2. Click the **Policy Items** tab.

   The Policy Items page opens, displaying a list of policy items on the right of the page.

3. Select the check box next to the item that you want to remove the policy from.

4. Click **Actions > Remove Policy**.

   If you are removing a policy that was automatically applied by a category, you can prevent the policy from being reassigned automatically.

Policy summary

To view a policy summary

1. On the **Policies** dashboard, double-click a policy.

   The Policy Summary page displays the following information:

   - **Date** that this version of the policy was published
   - **Items** number
   - **Executing** status
   - **Assigning** status
   - **Schedule Plan** type
   - **Priority** number
   - **Phases** list
   - **Policy Assignment Rate** shows a chart of the number of items that this policy has been assigned to by date
   - **Policy Execution Rate** shows a chart of the number of items that this policy has been executed on to by date

2. To change the granularity of the time axis on the policy activity charts, select an **Interval** option: *Hour, Day, Week, Month, Quarter, Year*.

Chapter 6: Manage target locations

You can create target locations to allow you to create policies that move, copy documents to the following locations.

- Target locations
- Add a target location
- Edit a target location
- Map repository fields to the target location metadata
- Define file naming conventions
- Direct target locations
- Remove a target location

Target locations

Target locations are repositories to which policies copy or move documents. When you create a policy to do so, you must specify the name of the target location. You can only specify defined target locations.

The documents in a target location repository are not necessarily imported , although in some cases you may want to scan them.

Example

> HPE Storage Optimizer may copy documents into a file system target location, and then apply disposal schedules to them in that location.

The **Manage Target Locations** on the Target Locations page is accessible through the Administration dashboard.

Different target locations support different types of policies. For more information on policy phase actions, see Policy phases, on page 66

For more information on policy phases by repository, see Repositories, on page 54.

Supported policy phases by target location

	Secure
File System	x
Hadoop	x
SharePoint 2007	x
SharePoint 2010	x
SharePoint 2013	x
SharePoint Remote	x

Add a target location

Use the following procedure to add a target location to your HPE Storage Optimizer system.

To add a target location

1. On the **Administration** dashboard, click **Target Locations**.

   The Target Locations page opens.

2. Click **+**.

   The Add New Target Location page opens.

3. In the **Details** section, specify the following information.

   > **NOTE:** Some options vary based on the type selected.

Name	Name of the target location.
Description	Description of the target location.
Connector Group	to use when sending work to the target location
Type	**Select the type of target location.**
via Connector	Select **via Connector** for a location that must be accessed by a HPE Storage Optimizer connector. Required for all policies that execute centrally.
Direct	Select **Direct** for a location that is accessed directly by the source repository connector. Required for Archive policies that execute on the source repository.
Direct Type	**Select the Direct Target Location type.**
Path	Select to archive using the Edge File System Connector.
Insert Configuration Settings	Predefined settings to use when sending items to the specified target location. Administrators must create Insert Configurations.

4. In the **Settings** section, specify the repository-specific target location values.

   > **NOTE:** Settings are not required in full for the target location. Any required parameters that are not supplied in the target location definition must be specified when a policy is created that references this target location.

5. Click **Save**.

Direct target locations

HPE Storage Optimizer uses direct target locations for archiving and stubbing files with the Edge Filesystem connector.

- **Path**. The archived file will be stored on a shared directory specified by the UNC path.

Edit a target location

You can alter the settings for a target location from the Target Locations page.

To edit a target location

1. On the Administration dashboard, click **Target Locations**.

   The Target Locations page opens.

2. On a target location panel, click the menu icon (≡), and then select **Edit**.

   The Edit Target Location page opens.

3. In the Details section, edit the following information as required.

Name	of the target location
Description	of the target location
Connector Group	name of the connector group that the required HPE IDOL connector registers with the distributed connector
Insert Configuration Settings	(*Optional*) Predefined settings to use when sending items to the specified target location. Administrators must create Insert Configurations. **NOTE**: Insert configuration settings are not required in full for the target location. Any required parameters that are not supplied in the target location definition must be specified when a policy is created that references this target location.

4. In the Settings section, edit the repository-specific target location value as required.

   - **File System**. Specify the UNC Target Folder.
   - **SharePoint 2007, 2010, 2013, or Remote**. Enter the Target URL.

5. Click **Save**, and then click **OK**.

Map repository fields to the target location metadata

If you want HPE Storage Optimizer-specific metadata to persist across target locations, you can map repository document fields to custom fields in target locations. Field mapping occurs at the group level.

For example, you configure custom field mapping for the Exchange connector, and the mapping applies to all Exchange repositories.

To map fields to target location metadata

1. On the Administration dashboard, click **Insert Configuration**.

   The Insert Configuration page opens.

2. Select the **Connector Group**.

   The available options depend on which connectors are active.

3. Select the **Insert Configuration** file to customize.

   You can modify the default file, create a new one, or duplicate a configuration file.

4. In the Field Mapping section, enter a part of the field name in the **Source Field** box.

   As you type, a list of matching fields is displayed. Select the field from the list.

5. (Optional) Click [*f*] to select a **Preprocessing** option.

   Any preprocessing options modify the final output.

   For example, you can map the date field and, by selecting one of the date options, change the date format from Epoch time to M/D/Y format. If you select an option, the icon darkens.

6. In the **Target Name** text box, type the target metadata field name.

7. (Optional) Repeat steps 4 to 6 to add as many custom mappings as required.

8. Click **Save**.

   The custom mappings take effect the next time a policy that includes the mapping runs.

Define file naming conventions

To organize data in target locations effectively, you can use naming conventions to ensure uniqueness and consistency in stored data. You can define a file naming convention in HPE Storage Optimizer to use when copying, moving, or declaring documents to target locations.

The HPE Storage Optimizer naming convention consists of a series of field names or text. The default naming convention is the field `AU_CP_TITLE`, underscore text, and a `UUID` (universal unique identifier) field. Target location files receive names in the following format by default.

`AU_CP_TITLE_UUID`

> **TIP:**
> HPE recommends that you use a unique identifier in your naming conventions to ensure that there are no duplicate file names.

To define a naming convention

1. On the Administration dashboard, click **Insert Configurations**.

   The Insert Configurations page opens.

2. Select the **Connector Group**.

   The available options depend on which connectors are active.

3. Select the **Insert Configuration** file to customize.

   You can modify the default file, create a new one, or duplicate a configuration file.

4. Open the Name Mapping section.

Each box in the name mapping section indicates a single field name or text string. Text boxes are marked with a **T**.

You can add or remove fields and text.

Click [+] to add field or text boxes, or click [×] to remove field or text boxes.

5. Define the naming convention as required.

   - In field boxes, enter a part of a field name. As you type, a list of matching fields is displayed. Select a field from the list.

   - In text boxes, enter the desired text string. You can only use characters that are allowed in Windows file names.

   - (*Optional*) Click [*f*] to the right of any field or text box to select a **Preprocessing** option.

     Preprocessing options modify the final output. For example, you can map the date field and, by selecting one of the date options, change the date format from Epoch time to M/D/Y format. If you select an option, the icon darkens.

6. Click **Save**.

   The naming convention updates and applies to any future documents that are copied, moved to target locations.

Remove a target location

When you no longer require a target location, you can remove it from HPE Storage Optimizer.

> **NOTE:**
> The target location that you are trying to remove must not exist as the target location for any current policy. Before you remove the target location, you must amend all policies that reference it to point to a different target location.

To remove a target location

1. On the Administration dashboard, click **Target Locations**.

   The Target Locations page opens.

2. On a target location panel, click the menu icon (≡).

3. Click **Delete…**

   A confirmation message opens.

4. Click **Delete**.

   The target location is removed.

Chapter 7: Manage policy conflicts

The Conflict Management page, which is accessible through the Administration dashboard, displays policy execution conflicts that HPE Storage Optimizer encounters as it applies and executes policies against content. The page allows an administrator to define the action to take for each policy conflict scenario encountered and lists each conflict resolution decision that was previously been defined.

- Policy conflicts
- Policy conflict set
- Resolve policy conflicts

Policy conflicts

A policy conflict occurs whenever a policy phase is ready to execute on a document and other policies were applied to that document.

HPE Storage Optimizer automatically reports policy conflicts when it encounters them. When HPE Storage Optimizer first encounters a conflict, it does not execute the policy for affected documents until the conflict is resolved.

Policy conflict set

The *Policy Conflict Set* is the combination of the policy phase that is running on the document (the executing policy), and other policies present on the document (conflicting policies).

For example, Documents A and B have three policies applied to them.

- **Policy1**. Dispose 5 years after creation.
- **Policy2**. Dispose 10 years after creation.
- **Policy3**. Secure Copy 1 year after date of last modification.

If HPE Storage Optimizer attempts to execute **Policy3** first on document A because it meets the policy execution rule, then the policy conflict set is:

- Executing policy — Policy3
- Conflicting policies — Policy1, Policy2

If HPE Storage Optimizer attempts to execute **Policy1** first on document B because it meets the policy execution rule, the policy conflict set is:

- Executing policy — Policy1
- Conflicting policies — Policy2, Policy3

Resolve policy conflicts

You can configure HPE Storage Optimizer to attempt to automatically resolve conflicts, or you can resolve them manually. There are advantages and disadvantages to each approach.

- Resolving conflicts automatically is fast, but may cause some undesirable resolutions
- Resolving conflicts manually ensures that conflicts are resolved the way you want, but is less efficient and more time-consuming

Automatically resolve conflicts

You can enable a configuration setting to allow HPE Storage Optimizer to automatically resolve conflicts based on the priorities of the conflicting policies.

- If the policy trying to execute has a higher priority than all other policies, HPE Storage Optimizer allows it to execute.
- If the policy has a lower priority than the others, HPE Storage Optimizer prevents it from executing.
- If the conflicting policies have the same priority, the conflict remains unresolved and you must resolve it manually.

To enable automatic conflict resolution

1. On the **Administration** dashboard, click **Settings**.
2. The Settings page opens.
3. On the **General** tab, click **Details**.
4. In the **Details** section, change the **Autoresolve Conflicts** option to **Yes**.
5. Click **Save**.

Manually resolve conflicts

You can manually resolve a policy conflict in two ways:

- **Allow** the Executing Policy phase to execute.
- **Prevent** the Executing Policy phase from executing.

Allowing or preventing a policy phase from executing does not impact the additional policies in the policy conflict set. These policies still execute when the policy execution rules are met.

HPE Storage Optimizer stores the conflict resolution decision (Allow or Prevent) and automatically applies this resolution to any documents that encounter the same Policy Conflict Set in the future.

To resolve a policy conflict

1. On the Administration dashboard, click **Conflict Management**.

   The Conflict Management page opens.
2. Select the policy conflict to resolve.

   Unresolved policy conflicts show an Action of **Undefined**.
3. In the menu bar, click **Actions**, and then click **Edit**.
4. If necessary, view the details of the **Executing Policy** and the **Additional Policies** by clicking their names in the relevant sections.
5. Select the required **Action** from the list.

   Available actions are:

- **Allow**. The Executing Policy is always allowed to execute.
- **Prevent**. The Executing Policy is prevented from executing.

6. (Optional) Update the automatically generated **Name** and add a **Description** for the current policy conflict.

7. (Optional) Click the policy conflict name to view the documents that match the associated policy conflict set to assist with policy conflict resolution decisions.

Chapter 8: Categories

Categories identify what content HPE Storage Optimizer policies apply to. Categories allow policies to be applied automatically to new content entering an organization. This section describes how to create categories, how to train them to match appropriate content, and how to measure their effectiveness.

- Taxonomy
- Categories

Taxonomy

Categories exist in a hierarchical structure called a *taxonomy*. The taxonomy has a single, top-level, root category. All category nodes have at least one parent category and can have zero or more children (sub-categories).

Categories

Most categories are used to find documents or files using metadata and concepts found within unstructured text.

A category definition is a mathematical rule against which each document can be evaluated for membership in that category. You can train a category by using:

- **Field Text** is a combination of field criteria that identifies a set of documents based on a property value match. The property value can either be from the document or from the storage location of the document.

Edit a category

When you edit an existing category, a draft version of the category is created, which allows you to edit the category and measure the impact of the adjustments without affecting the published category. The published category continues to be the version in use until you publish the draft category. You can also discard changes to the draft category to ensure that the published category continues to be the version in use.

To edit a category

1. Select the category in the taxonomy, and then click ✎.
2. Adjust the training or settings of the category, as required.

   The effect on the category results is indicated using a movement indicator to the right of the quality weight value.
3. Click **Save**.

   The Publish dialog box opens. You can add an optional comment.
4. Click **Publish**.

   The category is published in the Categories list.

View a category history

You can view the version history of a category in an Audit Report.

To view a category history

1. On the Administration dashboard, click **Audit Reports**.

   The Audit Reports page opens.

2. Select **Category Training Activity**.

3. Adjust the report settings as required.

4. Click **View Report**.

   The report opens and displays a history of the published versions of the selected categories, as well as any comments entered at the time of publication.

View the category details

On the Categories dashboard you can do the following:

- Browse categories.
- View category details.

To browse categories

- In the left panel category taxonomy, browse the category name in the **Categories** list.

  The right panel displays a file list for the category. You can perform several tasks on the file list:

  ○ Click **Display Summaries** 🗎 to toggle the display of file summaries in the file list.
  ○ Filter the file list.

    To filter a file list

    1. Click one of the filter icons to the right of the file list.

       When you click an icon, a dialog box opens where you can specify the filter criteria. You can filter by:

       ▪ **Title**
       ▪ **Age**, by date of creation, last access, or last modification. You can combine multiple selections to identify date ranges as required.
       ▪ **File Size**
       ▪ **File Type**
       ▪ **Users**
       ▪ **Group** allows you to filter by Active Directory group to see what documents are available to different groups
       ▪ **Tags** (Analyzed repositories) Select a tag, and then click **+** to include documents with the selected tag in the list, or click **-** to exclude documents with the selected tag from the list.
       ▪ **Custom Property** is only available if a HPE Storage Optimizer administrator configured any

- ▪ **Policies**
- ▪ **Potential** (Analyzed repositories) One or more of Redundant, Obsolete, and Trivial. You can select subsets of the Obsolete and Trivial criteria. For Redundant information, you can show all duplicates or only duplicates of specific repositories.

2. Click **Filter**.

    The filter applies and the file list refreshes. You can apply multiple filters to a file list as required.

    To clear all filters, click **X** above the filter icons.

To view category details

- From the category taxonomy, select the category, and then click .

    The View Category dialog box opens. It lists the basic category configuration settings.

Delete a category

You can delete an existing category if it is no longer required. Any HPE Storage Optimizer policy that uses this category no longer applies to the content after the category is deleted.

To delete a category

1. On the Categories dashboard, select the category from the taxonomy, and then click .

    A confirmation dialog box opens.

2. Click **Delete**.

Export individual categories

You can export individual categories and their children in XML format.

To export a category

1. On the Categories dashboard, select the category, and then do one of the following:

    - click
    - click **Actions > Export**

    The Export dialog box opens. To include the category contents, ensure that **Include Category Contents** is selected.

2. Click **Export**.

    The browser window offers the ability to select where to save the XML file, named `Category Export - CategoryName.xml`.

Export all categories

All categories under the top-level root node can be exported to an XML file.

To export all categories

1. On the Categories dashboard, select **Categories**, and then click ⬈.

   The Export All Categories dialog box appears. To include the category contents, ensure that **Include Category Contents** is selected.

2. Click **Export**.

   The browser window offers the option to select where to save the XML file, named `Category Export - All Categories.xml`.

Import a category hierarchy

You can import a previously exported category hierarchy.

To import a category hierarchy

1. Select the category under which to import the hierarchy, and then click ⬇.

   The Import Category page opens.

2. In the **File to Import** box, browse to the location of the category XML file, and then click **Open**.

3. In the **Child Categories** setting, select whether to **Keep** or **Remove** child categories.

4. In the **Duplicate Handling** setting, specify how to handle encountered duplicates:

   - **Keep existing**
   - **Merge**
   - **Overwrite**

5. Click **Import**.

Chapter 9: Clean up legacy data

This section describes how to manage legacy data in analyzed repositories.

- Introduction
- View repository data
- Clean up legacy data
- Configure potential ROT rule sets
- Re-analyze a repository
- Create and modify tags
- Modify analysis details
- Select a connector for manual scan

Introduction

When HPE Storage Optimizer analyzes a repository, it automatically identifies data appropriate for cleanup. You can then further refine the results by reviewing data in a number of ways. You can:

- view duplicates of master locations
- sort data by age, type, size, and other characteristics
- view data by tags that you have applied
- browse information in file lists

By reviewing legacy data, you can identify redundant, obsolete, or trivial information, and deal with it appropriately.

View repository data

There are several ways to view analyzed repository data.

- You can view repository lists by status on the Repositories dashboard. Click repositories to view more detailed information.
- For analyzed repositories, click items on the summary page, which redirect you to the appropriate tab.

By viewing content from different perspectives, you can identify which documents you want to clean up and how.

View a summary of repository data

The summary page of the repository details page displays statistical information about the data in Analyzed repositories.

To view the repository summary

- On the **Repositories** dashboard, click the **repository** under the **Analyzed** tab whose data you want to view.

The summary page displays the following statistical information. You can click shaded or colored areas in the various charts to drill down to another tab where you see a list of the selected files.

- Basic repository information: name, location, type (file system, Exchange, and so on), registration date, the total number of documents, and disk space appear in the menu bar.

- **ROT** (redundant, obsolete, or trivial) data, which shows potential and tagged redundant, obsolete, and trivial data, as well as the amount of disk space used by each and the total potential disk space savings. Potential ROT data appears as blue chart segments; tagged data is black.

  HPE Storage Optimizer automatically detects potential ROT data according to a default rule set. For example, it marks image files as trivial and duplicate information as redundant. You can also tag files as you review repository content. You can configure multiple rule sets to determine what HPE Storage Optimizer identifies as ROT data. See Configure potential ROT rule sets, on page 106.

- The **Addition Rate (Items)** and **Addition Rate (Disk Space)** display the amount of data added to the repository in each of the past ten years. This information gives you an idea of how quickly the repository is growing and how old the data is.

- **File Types** displays repository content by document type, such as text, video, audio, database files, spreadsheets, and so on. Depending on your repository and your organization's practices, you can use data type information to quickly identify documents for certain types of cleanup actions. For example, if you know that audio and video files are not relevant to your business needs, you can easily identify them for disposal.

- **Potential Risk Items** shows the number of documents considered to potentially represent risk since they could not be accessed during analysis. This can occur when documents are password protected, encrypted, or cannot be opened as the identity that the connector is running as cannot access it. The connector runs as a user on the NT/Network Service machine. That user may not have permission to open the file and inspect the contents.

View data details

You can view subsets of repository data by clicking segments on the Summary tab or by clicking different tabs. The Duplicates, Analysis, Tags, and Contents tabs display file lists that you can further refine by searching, filtering, or sampling. See Common file list operations for more details.

When you identify data that requires a cleanup action, you can tag it appropriately. See Clean up legacy data.

View duplicated data

Legacy repositories may contain multiple copies of the same data. *Master locations* contain *master documents*, or master copies of company records or other important items, however, there may be duplicates in other locations in the same repository.

It is rarely necessary to maintain duplicated data. It is likely that you will dispose of it during cleanup. HPE Storage Optimizer includes deduplication technology that detects duplicates of documents in master locations, as well as duplicates within individual repositories.

To view duplicate data

1. On the repository **Summary** page, perform one of the following actions:

   - Click the **Duplicates** tab, and then select how you want to view duplicate data.

     ○ **By Location**. The duplicate files appear on the right of the page. On the left, charts display the duplicate document count by master repository and within the repository (Internal), as well as the storage space used.

     ○ **By Duplicate Set**. The sets of duplicate documents are listed on the left of the page. When you select a duplicate data set, a list of all duplicate files in the set appears on the right. The oldest file in the set is marked with a red star, which indicates the master copy.

   - On the **Summary** page, click the **Potential** or **Tagged** shaded areas in the **Redundant** data chart, or the numerical total in the center of the chart. The tab displays the Potential or Tagged duplicates respectively.

     The duplicate files appear on the right of the page. On the left, charts display the number of duplicate files and the storage space used.

2. Refine the file list as required. See Common file list operations .

3. When you identify the duplicate data you want to clean up, take the appropriate cleanup action. See Clean up legacy data.

View data by statistical analysis

You can view data by a number of statistical analyses. Use the various options to isolate data by user, age, type, and so on.

To analyze data

1. On the repository **Summary** page, click the **Analysis** tab, and then select the statistic by which to analyze the data.

   - **by Age**
   - **by Custom Field** (if applicable)
   - **by Risk**. See View a summary of repository data , on page 97
   - **by Size**
   - **by Type**. See View data by file type, below
   - **by User**

   On the left side, charts display the number of files and the storage space used. The charts in the Count and Space Used display information relevant to your selection. For example, analyzing by User displays a graph that displays the number of documents by user name. Click the desired bar from the Count or Space Used graph to view the files.

2. Refine the file list as required. See Common file list operations .

3. When you identify the data to clean up, take the appropriate cleanup action. SeeClean up legacy data.

View data by file type

You can view repository data by general file type, and view all of the different file extensions that are included in the type. For example, you can view all files of the *Document* type, and then look within that

type to view the types of document, such as .DOC, .ODM, .PDF, and so on.

The *Other* file type includes all unknown extensions.

To view file type details

1. On the repository details page, click **Analysis > By Type**.

   On the left, charts display the number of files and the used storage space. By default, the analysis type is set to **All**, which displays information for general groups of data, such as *Document*, *Email*, *Image*, and so on.

2. Click **All**, and then select the data type to view.

   The Count and Space Used charts update and display totals by file extension.

View tagged data

If you have tagged files (see Clean up legacy data, on page 104 for more information on tagging), you can view a list of items divided by tag. Reviewers can easily identify items tagged for review by viewing tagged data.

To view tagged data

1. On the repository summary page, perform one of the following actions.
   - Click the **Tags** tab.

     The tab displays all tagged data in the repository.

     The tagged files appear on the right of the tab. On the left, a chart displays the number of files.

   - On the **Summary** tab, click the **Potential** or **Tagged** indicators in the **Redundant** data graphic.

     The tab displays the Potential or Tagged duplicates respectively.

     On the left, charts display the number of tagged files and the used storage space. Click the desired bar from the Count or Space Used graph to view the files.

2. Refine the file list as required. See Common file list operations , on the next page.

3. When you identify the data to clean up, take the appropriate cleanup action. SeeClean up legacy data, on page 104.

Browse data

To get a general idea of content in a repository or in a folder in the repository, you can browse the repository contents. You can browse repositories in any state: Registered, Managed, or Analyzed.

To browse repository data

1. Click the **Contents** tab.

   The tab displays a list of all files in the repository, as well as a collapsible Location box on the left side of the page that displays a tree structure of the repository. When you select a node in the tree, the file list displays only the contents of that node. You can select the **Including Subfolders** option to show contents of nodes below the current level.

2. Refine the file list as required. See Common file list operations , on the next page.

3. When you identify the data to clean up, take the appropriate cleanup action. SeeClean up legacy data, on page 104.

Common file list operations

A number of tasks common to all file lists in HPE Storage Optimizer allow you to refine your lists to identify data for cleanup. Refining a file list allows you to sort data by multiple criteria.

For example, you can view a list of text files, and then filter the list by date to identify text files created by a certain user or that contain certain keywords.

Search for files

You can search within a file list or across all repositories to identify documents that contain specific words. Search returns results only in Content Analyzed (HPE IDOL) repositories.

To search within a file list

1. To the right of the file list, click the search icon (T).

    The Title filter dialog box opens.

2. Enter the search text in the text box.

    To search document contents as well as in titles, select the **Contents** box. To search titles only, clear the box.

3. Click **Filter**.

    The file list displays the filtered search results.

To search across all repositories

1. In the navigation bar, click 🔍.

    The navigation bar changes to a search bar.

2. Enter the search text into the Search bar, and then press **Enter**.

    The search results open in a results dialog box.

Filter lists

Several standard filters are available that you can apply singly or in combination to refine a file list.

To filter a file list

1. Click one of the filter icons to the right of the file list.

    When you click an icon, a dialog box opens where you can specify the filter criteria. You can filter by:

    - **Title**
    - **Age**, by date of creation, last access, or last modification. You can combine multiple selections to identify date ranges as required.
    - **File Size**
    - **File Type**
    - **Users**
    - **Group** allows you to filter by Active Directory group to see what documents are available to

different groups

- **Tags** (Analyzed repositories) Select a tag, and then click **+** to include documents with the selected tag in the list, or click **-** to exclude documents with the selected tag from the list.
- **Custom Property** is only available if a HPE Storage Optimizer administrator configured any
- **Policies**
- **Potential** (Analyzed repositories) One or more of Redundant, Obsolete, and Trivial. You can select subsets of the Obsolete and Trivial criteria. For Redundant information, you can show all duplicates or only duplicates of specific repositories.

2. Click **Filter**.

   The filter applies and the file list refreshes. You can apply multiple filters to a file list as required.

   To clear all filters, click **X** above the filter icons.

Sample lists

If the file list is very long, you can take a sample percentage or a number of the total, which may make your analysis easier although, of course, some desired information may be excluded from the sample.

To sample a file list

1. Click the sample icon ( ) above the file list.

   The Sample dialog box opens.

2. Select the number or percentage to sample from the file list.

   The file list refreshes and shows the desired sample of the total.

View files and file properties

File lists only display the file names. To view more details, you can view the file contents in a browser or view the file properties.

To view files

1. Click the file name.

2. Click one of the options at the bottom of the Properties area.

   - **View** the file in a new browser window
   - **Download** the file to open it locally

To view basic file properties

- Click the file name.

   The area beneath the file expands to display basic file properties, such as name and location. You can configure which properties are included in the basic properties list. For more information, see Configure item properties, on the next page.

To view advanced file properties

1. Select the box next to the file name.

2. Above the file list, click **Actions > Advanced Properties**.

The Advanced Properties dialog box opens. It displays all basic file properties, as well as a variety of fields, which can be useful to select when building rules.

Configure last accessed date

By default, the last accessed date displayed in file properties is the value captured when the document was last scanned or re-scanned. If source documents are accessed after the scan, the last accessed date does not update in HPE Storage Optimizer. As a result, policies associated with last accessed dates may run prematurely if the files in question are still accessed by users regularly. This is less of a problem for dormant data.

You can configure the FileSystem connector to re-scan documents whenever the last accessed date changes in the source file.

Windows systems disable the last-access time stamps for performance reasons. When working with documents that users continue to access, ensure that you enable the last-access time stamps in Windows before you build policies linked to the last accessed date.

To enable re-scan of documents when the last-accessed date changes

1. Open the FileSystem Connector configuration file in an editor.

2. Uncomment the following line in the `[Default]` section.

   `IngestIfLastAccessChanged=true`

3. Save the file.

4. Restart the FileSystem Connector.

Configure item properties

You can configure which item fields appear in the file properties list. The default properties are:

- Name
- Location
- Created Date
- Last Modified Date
- Last Accessed Date (see Configure last accessed date, above)
- Creating User
- File Type
- File Size

To access the property configuration page

1. On the **Administration** page, click **Settings**.

2. On the **General** tab, click **Fields**.

3. Expand the **Item Properties** section.

To add an item property

1. Click **Add**.

   The Add Property dialog box opens.

2. Enter a **Display Name**.

3. Select a property **Type**: either **Date** or **String**.

4. Click the **Fields** box, and then select a property from the list.

   You can add multiple fields to a property.

5. When you finish, click **Add**.

6. Click **Save**.

   The new property appears when you view file properties. See View files and file properties , on page 102.

To remove an item property

1. Click **X** at the right of the property row.

2. Click **Save**.

   The property no longer appears when you view file properties. See View files and file properties , on page 102.

Display document summaries

You can toggle the display of document summaries in file lists. By default, summaries appear in search result lists and are hidden in all others. Summaries are available for only those documents that at the Content level.

To toggle document summaries

- In the menu bar, click .

Export item data

You can export item properties from file lists to .csv files that you can open in a spreadsheet program, such as Microsoft Excel. For example, you may need to export item data to send to a superior for approval before taking action on the items.

To export item data from file lists

1. Select the items from any file list.

2. Click **Actions > Export**.

   HPE Storage Optimizer generates a .csv file that you can save and open. The file lists the selected item properties. You can change the type of information exported to the file by changing the displayed file properties. See Configure item properties, on the previous page.

Clean up legacy data

Cleaning up legacy data is generally a two-stage process. The first stage consists of identifying and tagging content for removal, preservation, protection, or review. After you tag the content, an information manager creates policies that match the tagged content and applies the appropriate actions.

When you identify legacy data for cleanup by exploring analyzed data (see View repository data , on page 97), you can tag it.

Four tags are available by default.

- **Remove** it from the repository. The items appropriate for removal include:
  - redundant information, such as duplicates, convenience copies, or decommissioned documents
  - obsolete information, such as very old files, irrelevant or unused files, or files that have not been accessed or updated in a long time
  - trivial information, such as personal files, media files, or system files
- **Preserve** it in a records repository. Items appropriate for preservation can include compliance records, business records, master copies, or other items of business importance.
- **Protect** it in a secure archive. Items appropriate for protection include confidential information, security risks, or documents that contain personally identifiable information such as Social Security numbers, IP addresses, or email addresses.
- **Review**. You can tag data for different reviews depending on content, including HR review, legal review, business review, or records review.

You can create custom tags or modify the default tags as required. See Create and modify tags , on page 108.

Ideally, by the time you finish cleaning up your legacy data, only active data should remain in the repository.

Tag files

When you identify data for cleanup, you must first tag it for removal, preservation, protection, or review.

To tag files

1. Identify the files to tag in a file list. See View repository data , on page 97.

2. Select the files, and then click the tag icon (🏷) above the file list.

3. In the Tags dialog box, select one of the tag names.

   - **Remove**
   - **Preserve**
   - **Protect**
   - **Review**

4. Select a **Reason** for applying the tag.

   The available reasons vary depending on which tag you select.

   When you select certain Reason values, a Comment list opens, where you can select an extra comment about the data.

   For example, if you select the **Remove** tag, and then select **Redundant** as the reason, you can select one of the following comments.

   - **Duplicate**
   - **Convenience copy**
   - **Superseded**

- **Decommissioned**
- **No value**

5. Click **Apply**.

   The file list refreshes and all tagged files display a tag icon in the Tags column.

Collaborate on data analysis through comments

You can use comments to collaborate with colleagues on repository analysis in real time. You can also use comments to make notes for reference.

To add a comment to a repository

1. On any repository detail page, click **Comments** in the menu bar.

   The Comments dialog box opens.

2. Click **Add Comment**, and in the text box, enter a comment.

3. Click **Save Comment**.

   The comment appears in the Comments dialog box. The comment icon on the menu bar displays the total number of comments made for the current repository.

Configure potential ROT rule sets

HPE Storage Optimizer identifies potential ROT information according to rule sets. You select a rule set to use when analyzing a repository or repository subset. One rule set is available by default. According to the default rule set, a file is considered to be:

- **Redundant** if it is a duplicate of another file in the repository or in any master location
- **Obsolete** if it was last accessed or modified five or more years ago
- **Trivial** if is an image, audio, video, or system file

You can create multiple rule sets as required to address particular use cases. You can add master locations that HPE Storage Optimizer uses to identify redundant information. You can also specify which file types to identify as trivial, or the age of files to identify as obsolete.

To add master locations for duplicate detection

1. On the Administration dashboard, click **Settings**.

   The Settings page opens.

2. On the Analysis tab, click **Master Locations**.

   The Master Locations page opens.

3. Under Details, click **Add**.

   The Add Master Location dialog box opens.

4. Enter a **Display Name** for the master location.

5. Select a master location from the **Repositories** list.

6. Specify whether the Master Location should be included in the default Potential Set by clicking **Yes** or **No**.

> **NOTE:**
> Duplicates will be visible by set when you specify a master location at the root node of the repository. If multiple master locations are present, duplicates are visible in the **Duplicates by Location** filter.

7. Click **Add**.

8. Click **Save**.

   The new location is added to the Master Locations list.

To add a potential set

1. On the Administration page, click **Settings**, on the Analysis tab, click **Potential Sets**, and then under Details, click **Add**

   The Add Potential Set dialog box opens.

2. Under Details, enter a **Name** and **Description** of the set.

3. Select or change any of the following settings as required. Under:

   - **Redundant:** select a master location to use to identify duplicate items.
   - **Obsolete:** select the criteria to use to identify obsolete information.
   - **Trivial:** select the criteria to use to identify trivial information.

4. Click **Add**.

   The Add Potential Set dialog box closes.

5. Click **Save**.

   You can use the new potential rule set when analyzing or reanalyzing repositories, repository groups, and subsets.

Configure a file group

You can configure the file group names and the file type extensions that comprise a file type group in Administration > Settings > Analysis > File Groups.

File type groups are used on the "Item Types" metric of an analyzed repository.

To configure a file group

1. Select **Administration > Settings > Analysis > File Groups**.

2. (*Optional*) To add a new File Group, click **Add**.

3. Enter the File Group **name** and the **file extensions**.

   Separate the extensions with a comma (no spaces).

4. Click **Add**, and then click **Save**.

5. Re-analyze all repositories to reflect the file group changes.

   Failure to re-analyze repositories can have unexpected results.

Re-analyze a repository

If the information in an analyzed repository, repository group, or repository subset has recently changed, you can re-analyze it to ensure that the statistical data is up to date. Any tags that you applied to content are maintained.

To re-analyze a repository or repository group

1. On any detail page, click the menu button by the repository name, and then click **Re-analyze Repository**.

   The Re-analyze Repository dialog box opens.

2. Select the **Potential Set** to use to identify ROT information, and then click **Re-analyze**.

   The repository or group is reanalyzed.

To re-analyze a subset

1. On any detail page, click the menu button by the repository name.

2. Click **Re-analyze all Subsets**.

   The Re-analyze Repository dialog box opens.

3. In the confirmation dialog box, click **Yes**.

   All subsets are re-analyzed.

Create and modify tags

HPE Storage Optimizer includes four default tags: **Preserve, Protect, Remove**, and **Review** (see Clean up legacy data, on page 104 for the descriptions), each of which includes several predefined **Reason** and **Comment** options that you can select to explain why you applied certain tags to certain documents.

The default tags and their corresponding Reason and Comment options may not be sufficient, depending on your business requirements, so in such cases, you can create your own tags and define your own Reason and Comment options. You can also add custom Reason and Comment options to the default tags.

To create and modify tags

1. On the **Administration** dashboard, click **Settings**.

   The Settings page opens.

2. On the **Analysis** tab, click **Tags**.

   The tag creation page opens. It contains three collapsible lists: Name, Reason, and Comment.

3. In the Name section, do either of the following actions.

   - To create a tag, click **Add**, and then enter a tag name.

   - To modify a tag, click the tag.

   The Reason section opens.

4. Click **Add**.

A text box is added to the Reason list.

5. Enter the reason in the dialog box, and then press **Enter**.

   A text box opens in the Comment section.

6. Enter a comment in the dialog box, and then press **Enter**.

7. (*Optional*) Continue to add as many tags, reasons, and comments as required.

   The new or modified tags are available for use in data clean up.

8. When you finish, click **Save**.

Modify analysis details

You can change analysis configuration details as required if the default settings are not appropriate for the data in your repositories. You can:

- change the **maximum number of segments** that can appear in area charts, such as the Redundant, Obsolete, and Trivial charts. When the number of segments exceeds the maximum, the data is presented in a bar chart.

- add **custom fields** to use during data analysis if the default fields (size, date, and so on) do not meet your needs. Before you can add custom fields, an administrator must configure them.

- add **category hierarchies** to use in the analysis. Before you can add categories, an administrator must configure them.

   To modify analysis details

   1. On the Administration dashboard, click **Settings**.

      The **Settings** page opens.

   2. Click the **Analysis** tab, and in the Details section, change any of the following settings as required.

      ○ **Max Segments**. Change the maximum number of segments that can appear in an area chart. If the number of segments in the data exceeds the maximum, the data is presented in a bar chart. The default value is **5**.

      ○ **Custom Fields**. Click **Add** to add custom fields that an administrator configured. In the dialog box, select a custom field from the **Fields** list, and then enter a **Display Name**.

      ○ **Categories**. Click **Add** to add categories that an administrator configured. In the dialog box, select the category to add, and then click **Save**.

         After you add one or more categories, you must select a **Summary Category** to appear on the repository Summary page.

   3. Click **Save**.

Select a connector for manual scan

You can select or change the connectors used by default when adding repositories of different types (see).

To select a connector for manual scanning

1. On the **Administration** dashboard, click **Settings**.

   The Settings page opens.

2. On the **Connectors** tab, click **Locations**.

   The connector location page opens.

3. In the **Details** section, select the desired connector from the appropriate list. and then click **Save**.

Chapter 10: Scheduled tasks

HPE Storage Optimizer includes a number of scheduled tasks to automatically perform jobs that are required to manage policies, generate statistical information for monitoring purposes, and so on. You can control how often these automated tasks run through schedules.

You can configure tasks to run on a scheduled basis or you can configure tasks to run only once.

- Default scheduled tasks
- Add a scheduled task
- Edit a scheduled task
- Remove a scheduled task
- Run scheduled tasks
- Configure HPE Storage Optimizer schedules for large systems

Default scheduled tasks

HPE Storage Optimizer provides a number of scheduled tasks out of the box. Only certain tasks run by default. You can enable or disable them to suit your requirements.

You can also add tasks that run only once, or you can configure additional scheduled tasks, for example, to execute a specific policy frequently.

Default scheduled task types

This section describes the scheduled tasks that are available by default.

Policies

- **Assign Policies** automatically assigns policies to documents, depending on their categorization. After you set up category associations for policies, this task automatically assigns the policies to documents that match the category.
- **Cleanup Policies** removes policy actions from documents after you remove a policy in HPE Storage Optimizer. When you remove a policy through the User Interface, the policy action remains on the document until the Cleanup Policies task runs.
- **Execute Policies (High)** applies the policy action to documents that have policies applied. It executes the action only when the document meets the policy rules. This Execute Policies task checks for items ready to execute every 10 minutes.
- **Execute Policies (Normal)** applies the policy action to documents that have policies applied. It executes the action only when the document meets the policy rules. This Execute Policies task checks for items ready to execute every 4 hours.
- **Execute Policies (Low)** applies the policy action to documents that have policies applied. It executes the action only when the document meets the policy rules. This Execute Policies task checks for items ready to execute every 24 hours.
- **Process Issues** processes the Abort and Retry actions from the Issue Management administration page.

- **Remove Policy Assignments** determines when to remove a policy from documents that no longer match the categories used when a category is retrained. You can configure HPE Storage Optimizer to remove policy for these documents.

  By default, a policy is not removed after it is assigned by a category.

- **Notify Policy Approvers** emails policy approvers for Policies configured for Review before execution.

Statistics

- **Calculate Compliance** calculates a measure of how many documents in a repository are being managed through a HPE Storage Optimizer policy assignment.

- **Calculate Conflict Statistics** updates metrics related to policy conflicts.

System

- **Metadata Compact** removes repositories that have been marked for permanent deletion.

- **Metadata Consistency Check** ensures that all stored metadata is consistent.

- **Register Repositories** automatically finds and registers all repositories . Data sources that contain documents and that are not one of the recognized types (Filesystem, SharePoint 2007, SharePoint 2010, SharePoint 2013, Exchange) are defined as XML type.

- **Workflow Batch** creates batched items and initiates workflows.

Default scheduled task configuration

The HPE Storage Optimizer installation installs and configures the following tasks and schedules.

Name	Interval	Enabled
Assign Policies	30 minutes	Yes
Cleanup Policies	24 hours	Yes
Execute Policies (High)	10 minutes	Yes
Execute Policies (Normal)	4 hours	Yes
Execute Policies (Low)	24 hours	Yes
Process Issues	60 minutes	Yes
Reevaluate Policy Assignments	60 minutes	No
Calculate Compliance	24 hours	Yes
Calculate Conflict Statistics	60 minutes	Yes
Metadata Compact	7 days	Yes
Metadata Consistency Check	7 days	Yes

Name	Interval	Enabled
Register Repositories	24 hours	Yes
Workflow Batch	60 minutes	No
Notify Policy Approvers	24 hours	Yes

Schedule plans

Schedule plans determine how frequently the tasks check for items that are ready to execute. You can use schedule plans to ensure that critical policies run more frequently than others.

Three Execute Policies tasks run on different schedule plans:

- **High** runs every 10 minutes
- **Normal** runs every 4 hours
- **Low** runs every 24 hours

When you define a policy or a policy template, you must select a schedule plan. See Create a policy template, on page 67 and Create a policy, on page 71.

You can edit the Execute Policies tasks to change the default run frequencies as required. See Edit a scheduled task, on the next page.

Add a scheduled task

You can create new scheduled tasks to control when specific operations execute.

You can create scheduled tasks to run specific tasks immediately, or to run on separate schedules.

To create a scheduled task

1. On the Administration dashboard, click **Scheduled Tasks**.

   The Scheduled Tasks page opens.
2. On the menu bar, click **+**.

   The Add New Scheduled Task page opens.

3. Specify the following information.

Name	The name of the task.
Description	The description of the task.
Schedule Type	The type of scheduled task to create.
Start At	The time and date that the schedule starts.
Run Once	Whether the task runs only once or on a schedule.
Frequency	The frequency that the scheduled task runs. Specify the frequency in hours and minutes.
Enable Scheduling	Whether to enable the task. Enabling the task means it runs (either once or according to the schedule) whenever the HPE Storage Optimizer scheduler is running.

4. Click **Save**.

Edit a scheduled task

You can alter scheduled task settings from the Administration dashboard.

To edit a scheduled task

1. On the Administration dashboard, click **Scheduled Tasks**.

   The Scheduled Tasks page opens.

2. Select a scheduled task panel, click the menu icon (≡) and select **Edit**.

   The Scheduled Tasks page opens.

3. Edit the fields as required.

Name	The name of the task.
Description	The description of the task.
Schedule Type	The type of action the task performs.
Start At	The time and date that the schedule starts.
Run Once	Whether the task runs only once or on a schedule.
Frequency	The frequency that the scheduled task runs. Specify the frequency in terms of hours and minutes.
Enable Scheduling	Whether to enable the task. Enabling the task means it runs (either once or according to the schedule) whenever the HPE Storage Optimizer scheduler is running. You can disable a task, for example, during system maintenance. The task does not run until you enable it again.

4. Click **Save**.

Run scheduled tasks

Enabled scheduled tasks run whenever a HPE Storage Optimizer Scheduler is active. You start HPE Storage Optimizer Schedulers from the Service Control Manager.

You can install one HPE Storage Optimizer Scheduler for each server. The number of threads you configure for the Scheduler determines the overall rate at which it processes items. On large HPE Storage Optimizer systems, you must deploy multiple Schedulers.

Run tasks immediately

Scheduled tasks run according to a defined frequency. You can advance the start date and time of the next schedule cycle to force it to run immediately.

To run a scheduled task immediately

1. On the Administration dashboard, click **Scheduled Tasks**.

   The Scheduled Tasks page opens.

2. Select a scheduled task panel, click the menu icon (≡) and select **Run Now**.

   The HPE Storage Optimizer Scheduler runs the task when it next checks for tasks to run (by default, every 60 seconds).

Configure HPE Storage Optimizer schedules for large systems

The following section describes HPE Storage Optimizer configurations to use in large HPE Storage Optimizer systems. Depending on your requirements and hardware, you can combine the solutions in this section as required.

Change the number of scheduler threads

Each HPE Storage Optimizer Scheduler runs a defined number of threads, each processing a batch of items every time it runs. The default number of threads is eight. The optimal number of threads depends on your requirements and the system processor.

To change the number of Scheduler threads

1. Open the **HPE Storage Optimizer Configuration Manager**.
2. Click **Engine**.

   The Engine Setting page opens.
3. Under **Engine Settings**, enter the number of threads in the **Enter the number of threads to use to process items** box.
4. Click **Deploy**.

   HPE Storage Optimizer redeploys.

Install multiple HPE Storage Optimizer schedulers

For high processing volumes, you can install multiple HPE Storage Optimizer Schedulers on several machines. You must modify the configuration of each Scheduler to point to the HPE Storage Optimizer SQL Server database.

Remove a scheduled task

When you no longer require a scheduled task, you can remove it from HPE Storage Optimizer.

To remove a scheduled task

1. On the Administration dashboard, click **Scheduled Tasks**.

   The Scheduled Tasks page opens.
2. Select a scheduled task panel, click the menu icon (☰) and select **Delete**.

   A confirmation dialog box appears.
3. Click **Delete**.

Chapter 11: Issue management

This section describes the Issue Management administrative function.

- Manage issues
- Resubmit failed items
- Abort failed items

Manage issues

The Issue Management page, which is accessible through the Administration dashboard, displays events of interest to HPE Storage Optimizer Administrators. Typically these events require manual intervention to resolve.

For example, the Issue Management list may report when:

- a connector or the distributed connector stops
- a dispose action fails due to a lack of permission on the target document
- a copy action cannot access the target location
- the configured `Temporary Location` cannot be accessed

Multiple occurrences of individual events can be filtered and then processed using a single retry or abort instruction. This bulk handling mechanism makes it easy to resolve environmental issues and to replay the underlying HPE Storage Optimizer actions.

The issues described in this section are typically the result of problems in the system environment such as: incorrect permissions, access problems, and so on.

Resubmit failed items

After you resolve the problem that caused the issue, you can resubmit items.

To resubmit items

1. On the Administration dashboard, click the **Issue Management** panel or tab.

   The Issue Management page opens.

2. Select all items that you want to retry.

   You can filter on any of the columns, and then click **Select All** to select multiple common items.

3. In the Actions menu, click **Retry**.

Abort failed items

You can abort items that failed in processing. Aborting a policy execution removes the policy tag from all the selected items.

> **NOTE:**
> The policy can be reapplied to some or all of the aborted items whenever the Apply Policies from Category task runs again.

To abort failed items

1. On the Administration dashboard, click the **Issue Management** panel.

   The Issue Management page opens.

2. Select all items that you want to abort.

   You can filter on any of the columns, and then click **Select All** to select multiple common items.

3. Click **Abort**.

Chapter 12: Health Checks

You can use the Health Checks page to verify key configuration settings in your HPE Storage Optimizer deployment.

- Check HPE Storage Optimizer health
- Run advanced health check reports
- Usage details, on the next page

Check HPE Storage Optimizer health

The Health Check page allows you to check the status of components and tasks in your HPE Storage Optimizer system.

When you run the health check, the page displays the current health of the system. You can view any warnings or errors, and correct them in your system.

For example, the health check reports whether HPE Storage Optimizer can contact the connectors.

To check HPE Storage Optimizer health

1. On the Administration dashboard, click **Health Checks**.

   The Health Checks page opens. The following groups summarize aspects of your configuration settings.

   - **IDOL and Connectors.** Summarizes the status of IDOL and connector components.
     - **IDOL port status.** Verifies the IDOL action and index ports.
     - **Distributed connector port status.** Verifies the Distributed Connector action port.
     - **Connector callback.** Verifies the accessibility of the callback site.
     - **List connectors.** Verifies which connectors have registered with the Distributed Connector.
   - **Storage Optimizer**. Verifies access to various Storage Optimizer components.
     - **Insert temporary location.** Verifies access to the temporary location used by connectors when performing inserts.
     - **Policy parent category ID.** Verifies the validity of the parent category ID for new policy categories.
     - **Target location insert configuration.** Verifies read access to the target location insert configuration shared directory.
   - **Edge Connector Status.** Verifies access to the Edge Filesystem connector server.

     The status displays the following possible states:

     - Green checkmark — indicates that all Edge Filesystem connector services are running and stable.
     - Yellow exclamation — indicates that at least one Edge Filesystem connector web service is down.
     - Blue question — indicates that the Edge status service is down, and the connector web services cannot be queried.

       If this status occurs, you must troubleshoot your Edge Filesystem connector services.

To refresh the health check

1. Click one of the refresh options ( ):

   - **Refresh all**. Refreshes the status of all components.
   - **Refresh group**. Refreshes the status of a specific group.
   - **Individual component.** Refreshes the status of a specific component.

     The health check runs, and you can view the results of each test.

Run advanced health check reports

In addition to the basic HPE Storage Optimizer health check, you can use HPE Storage Optimizer Assist to run advanced health check reports on connector status, execution activity, and policies.

> **NOTE:**
> To run the policy reports, you must have configured at least one policy.

To run advanced health check reports

1. On the Administration dashboard, click **Health Check**.

   The Health Check page opens.

2. Click **View the advanced HPE Storage Optimizer Assist**.

   The HPE Storage Optimizer Assist page opens.

3. Select the **Report** to run.

Usage details

The Usage Details page allows you to check the amount of data managed by your HPE Storage Optimizer system.

To check HPE Storage Optimizer usage

1. On the Administration dashboard, click **Usage details**.

   The Usage Details page opens.

2. The **Usage Details** section displays the current and cumulative managed data details and numbers of documents ingested by HPE Storage Optimizer.

   - The **Breakdown by Index Type** section displays a chart of the following current managed data by index type:
     - **Repository Metadata and Metadata only.**
       - **Repository metadata** contains metadata retrieved from the repository.
       - **Metadata** contains a small analysis that contains all metadata.
     - **Full index content** Contains all metadata and item content.
   - The **Repository Type (Active)** charts display active repositories by connector type, with current and cumulative managed data.
   - The **Repository Type (Deleted)** charts display deleted repositories by connector type, with managed data.

- The **Repository Group (Top 10)** charts display the top ten repository groups with current and cumulative managed data.

Chapter 13: Audit Reports

> **NOTE:**
> The HPE Storage Optimizer audit reports feature is available if you configured SQL Server Reporting Services as part of the configuration of your HPE Storage Optimizer environment.
>
> The HPE Storage Optimizer data source in SQL Server Reporting Services (SSRS) must be configured with a set of credentials for access to the audit reports.
>
> For more information, see the prerequisites in the *HPE Storage Optimizer Installation Guide*.

To view audit reports

1. On the Administration dashboard, click **Audit Reports**.

   The SQL Server Reporting Services page for HPE Storage Optimizer Reports opens.

2. Click **HPE Storage Optimizer Reports.**

3. Click **English**.

   A list of reports opens.

   Click **Details View** for additional details on the reports.

Name	Description
Category Activity.rdl	Reports on Storage Optimizer Policy applied by Category
Category Report.rdl	Reports on training categories
Category Training Activity.rdl	Reports on training categories
Classification Activity.rdl	Reports on Storage Optimizer classifications and alerts.
Document Activity.rdl	Reports on lifecycle changes to documents under the management of specified Storage Optimizer Policies.
Policy History.rdl	Reports on authorized changes to Storage Optimizer Policies.
Repository History.rdl	Reports on authorized changes to Storage Optimizer source repositories.
Security Activity.rdl	Reports on Storage Optimizer Security changes.
Tag Activity.rdl	Reports on activity related to a specific tag name.
Tag History.rdl	Reports on authorized changes to Storage Optimizer tags
User Activity.rdl	Reports on lifecycle changes to documents under the management of Storage Optimizer Policy by a specified user.
Subreports	
Policy History Phases.rdl	Reports on authorised changes to ControlPoint Policy phase settings.

Name	Description
Policy History Training and Assignment.rdl	Reports on authorised changes to ControlPoint Policy training and assignment settings.
Security Changes.rdl	Reports on ControlPoint Security changes.

4. Click a report name. The report page opens.

5. Specify any necessary parameters, such as dates, and click **View Report.**

   The selected report opens.

Chapter 14: Custom properties

This section describes how to create custom properties, which you can use to sort and filter repositories and policies.

- Create a custom property
- Update the internal configuration of custom columns
- Add property values to repositories and policies

Create a custom property

HPE Storage Optimizer administrators can create custom properties to apply to repositories and policies. These properties allow users to sort and filter large repository and policy lists on the respective dashboards.

For example, you can create a *Region* property with three values, *Americas*, *Europe*, and *Asia*, and then apply the property values to your repositories. On the Repositories dashboard, you can then sort or filter the list by region.

You can filter by multiple properties to further refine your repository or policy list. For example, you can filter by *Region*, and then by a second property, such as *Department*, to identify all IT repositories in the Americas region.

To create a custom property

1. On the **Administration** dashboard, click the **Settings** panel.

    The Settings page opens.

2. On the **General** tab, click **Properties**.

    The Properties page opens.

3. In the **Details**section, click **Add**.

    The Add Property dialog box appears.

4. Specify the following information.

    - **Name** is the property name.
    - **Values**. Click **Add** to add as many property values as required.
    - **Availability**. Select whether to enable the property for policies, repositories, or both.
    - **Filtering**. Enable or disable the property for repository and policy list filtering.

5. Click **Save**.

    The property appears in the **Details** list.

Update the internal configuration of custom columns

Whenever you change the attributes of a custom column, for example, the length of the column, you must perform the following steps.

1. Re-run the following script to update the internal configuration of the custom column.

```
USE ControlPointMetaStore
EXEC MetaStore.ConfigureAddDocument
EXEC MetaStore.ConfigureUpdateDocument
EXEC ControlPointMetadata.ConfigureImportDocument
GO
```

2. Restart the Storage Optimizer MetaStore service.

Add property values to repositories and policies

After you create custom properties, you can apply property values when you add or edit repositories or policies. For more information, see

- Add a repository
- Edit repository settings
- Create a policy template
- Create a policy
- Edit a policy

Chapter 15: Customize HPE Storage Optimizer

This section is a reference for customizing your HPE Storage Optimizer environment.

Change Sample sizes when browsing a repository

If a repository file list is very long, you can take a sample percentage or a number of the total, which may make your analysis easier although, of course, some desired information may be excluded from the sample.

By default, HPE Storage Optimizer has the following sample sizes:

- **Number of documents:** 5, 50, 100, 500, 1000, 5000.

For more information on sample lists, see Sample lists , on page 102

If you require more flexibility in defining sample sizes, you can configure the system by using settings stored centrally in the Global Settings table in the ControlPoint database.

To customize samples sizes

1. In SQL Server, back up the ControlPoint database before attempting to make any changes to the CPGlobalSettings table.
2. Run the following SQL statement:

   ```
   update ControlPoint.dbo.CPGlobalSettings set SettingValue='<numberDocuments>' where
   SettingName='Autonomy.ControlPoint.Views.SampleSizesCSV'
   ```

   where

   - *<numberDocuments>* is the new integer value for the number of documents to sample.

   ### Examples

   To add 1, 6 and 20 as number of documents to sample, run the following SQL statement:

   ```
   update ControlPoint.dbo.CPGlobalSettings set SettingValue='1, 6, 20' where
   SettingName='Autonomy.ControlPoint.Views.SampleSizesCSV'
   ```

Limitations

The customization of sample sizes has the following limitations:

- Integer values for numbers of documents must be greater than zero (0) and less than half of the total number of documents.

  ### Example

  For a set of 100 documents, the number of documents should be set between 1 and 49.

Insert Configuration

Insert Configurations are predefined settings used to map custom fields a specified target location connector.

> **NOTE:**
> Only Administrators can create Insert Configurations.

Before you begin

If you need to use insert configurations, ensure that the `InsertConfigEnabled` parameter in the `<AppSettings>` in `ControlPointTimer.config` is set to true to enable insert configurations.

For more information, see InsertConfigEnabled, on the next page or the *HPE Storage Optimizer Best Practices Guide.*

Create an insert configuration

To define an insert configuration

1. On the Administration dashboard, click **Insert Configuration**.

   The Insert Configuration page opens.

2. Select the **Connector Group**.

   The available options depend on the connectors that are active.

3. Select the **Insert Configuration** file to customize.

   You can modify the default file, duplicate a configuration file, or create a new one.

4. In the Field Mapping section, enter a part of the field name in the **Source Field** box.

   As you type, a list of matching fields is displayed. Select the field from the list.

5. (Optional) Click f to select a **Preprocessing** option.

   Any preprocessing options modify the final output.

   Example

   > You can map the date field and, by selecting one of the date options, change the date format from **Epoch** time to **ConvertEpochToLocal** time format.

   > The resulting date has a date format with the local time zone offset.

   > 2017-04-11T09.07.05-04:00

   When you select a preprocessing option, the f icon darkens.

6. In the **Target Name** box, enter the target metadata field name.

7. (Optional) Repeat steps 4 to 6 to add as many custom mappings as required.

8. In the Name Mapping section, define the format for inserted item names.

   - Select **Add Field** to add a new field mapping.
   - Select **Add Text** to add a new text mapping.

9. Click **Save**.

   > **NOTE:**
   > If you add custom field mappings in Insert Configuration, you must restart the HPE Storage Optimizer Engine so that HPE Storage Optimizer picks up the new custom fields.

AppSettings in ControlPointTimer.config

The following is a reference for the `<AppSettings>` in `ControlPointTimer.config` file.

Settings	Usage
NumberOfTimerThreads	Number of Threads for the timer engine
ExceptionWaitTime	If 5 exceptions have been thrown in a row wait the amount of time indicated
ClientSettingsProvider.ServiceUri	
SleepSeconds	Thread sleep seconds for ingestion
MaxExecutionFrequencySeconds	Used in Phase execution
CallbackProcessor.MaxInstancesRunning	Used in collect cleanup
CacheExpirationSettingsCSV	CSV for long expiry seconds, short expiry seconds used during ControlpointFrameworkRegistration andPolicyExecutionRegistration
LoadBalancingSettingsCSV	CSV for maxLatestNoWorkCount, maxPhaseIgnoreSeconds, slidingIgnoreSecondsIncrease
ClearLocksAtStartUp	Boolean value. Set this value to true to clear locks in the ExecutionLog table during Storage Optimizer Scheduler startup. **NOTE:** Set this to true only when Storage Optimizer has only one Scheduler instance deployed. Engine crashes or unexpected restarts can leave locks on the execution items. Enabling this option clears the locks upon Scheduler start and therefore avoids putting policy executions on hold for a long period of time.
InsertConfigEnabled	Boolean value. Default value is 'false'. Set this value to true to enable querying MetaStore for insert configurations. Setting this value to false allows the engine to skip querying MetaStore for the insert configuration values, thus improving the execution performance for insert actions to target locations.

Settings	Usage
	If you need to use custom insert configurations, enable this option.
	If you do not use insert configurations, setting this option to false will improve performance for policy executions.

The following parameters are needed to enable secure connections with HPE IDOL and Connectors

Settings	Usage
SecurePorts	Boolean value, used to determine if the specified metastore port must be added to the MetaStore port list
MetaStorePort	Port number
LDAPServer	
LDAPBaseObject	
LDAPUseSSL	Boolean to use SSL
LDAPMaxResults	Maximum number of results to retrieve
XMLGroupMembershipFile	Filename containing group information

Chapter 16: Configure HPE Storage Optimizer MetaStore for metadata ingestion

This section provides an overview of the steps necessary for configuring HPE Storage Optimizer MetaStore to capture additional data during document ingestion. A set of examples will be used to show where and how this data can be captured.

- Data Mapping
- Additional data capture
- Examples
  - Example 1 – single value for the same document
  - Example 2 – single value hash for the same document
  - Example 3 – multiple values for the same document
  - Example 4 – multiple values hashed for the same document
- Existing data and re-ingestion
- Field text and advanced properties

Data Mapping

Document metadata is captured by a list of instructions dynamically generated based on information held in the **MetaStore.MapTable** and **MetaStore.MapColumn** tables.

A stored procedure named **MetaStore.MapField** handles the complexity of these mapping tables. Run this stored procedure to register data mappings for any additional document metadata to be captured into HPE Storage Optimizer MetaStore.

MetaStore.MapColumn

Field	Description
GroupNumber	Used when a source field is mapped to multiple times the same target table.
	For example, use GroupNumber for a complex field such as "ADDRESS" with a value {CITY="BFS", NUMBER=10, STREET="Queens"}. The inclusion of the same GroupNumber for the separate address parts keeps the information together within the one row in the target table.
	Default: 1
SourceName	The field to be extracted from the source document.
ExtractPath	The value of this field is typically null, except when a value is to be parsed from the source field.

Field	Description
TargetColumn	The name of the column where the captured value is to be stored.
TargetTransform	The type of transformation to be used before storing the captured value.
TargetTransformParams	When a transformation requires additional configuration, the configuration can be placed in the TargetTransformParams field. The value of this field is typically null.
SupportingTable	The name of the target hash table, if any. This field should be populated when the extracted data is to be hashed into a separate hash table.
CanUpdate	Indicates whether the information captured to the target column can be modified after creation.
Inherit	Indicates whether the information captured to the target column, when modified, should be captured to child documents. Examples of such inheritance would be security.
AlternativeFieldSource	The alternate field to be extracted from the source document when SourceName cannot be extracted.
AlternativeFieldSourceTransform	The alternate transform to be used when AlternativeFieldSource is specified.

MetaStore.MapTable

Field	Description
GroupNumber	See GroupNumber
SourceName	See SourceName
TargetType	The TargetType values are as follows: • "MVF" if the table can capture multiple values for the same document. For example, more than one row can exist for a given document. • "SVF" if the table can capture single values for the same document. For example, a maximum of one row can exist per document.
TargetTable	The name of the table to populate.
TargetMVPSuffix	Supports the extraction of a suffix from the source field name to further populate a column in the target table. For example, assuming data exists in the source document like:

Field	Description
	CPPATH1=\\c\ CPPATH2=\\c\test\ CPPATH3=\\c\test\folder\ Then it is possible to map CPPATH* as the SourceName and indicate that the value extract from * should be placed in the field configured by TargetMVPSuffix, for example "Level".
TargetMVPSuffixTransform	Specifies the transform to use when extracting a suffix. See TargetMVPSuffix.

MetaStore.MapField

The stored procedure **MetaStore.MapField** handles the complexity of the mapping tables by defaulting a number of optional parameters to typical values.

Parameter Name	Required	Default Value
@GroupNumber	No	(1), defaults to a single field mapping
@SourceName	Yes	
@TargetType	No	('SVF') , defaulting Single-valued Field(SVF)
@TargetTable	Yes	
@TargetMVPSuffix	No	(NULL), defaults to not specified
@TargetMVPSuffixTransform	No	(NULL), defaults to not specified
@ExtractPath	No	(NULL), defaults to not specified
@TargetColumn	Yes	
@TargetTransform	Yes	
@TargetTransformParams	No	(NULL), defaults to not specified
@SupportingTable	No	(NULL), defaults to not specified
@CanUpdate	No	(1) , defaults to TRUE
@Inherit	No	(0), defaults to FALSE
@AlternativeFieldSource	No	(NULL), defaults to not specified
@AlternativeFieldSourceTransform	No	(NULL) , defaulting to not specified

Additional data capture

HPE Storage Optimizer MetaStore includes the database schemas, **Metadata** and **ControlPointMetadata**.

Metadata and the corresponding tables (for example, **Metadata.Document**) are used for the default set of captured properties only. Extensions to this default set must be captured into the **ControlPointMetadata** schema instead.

- If the additional data to be captured is a single value field (SVF), then it must be captured in the **ControlPointMetadata.Additional table**.

- If the additional data to be captured is a multivalue field (MVF) instead, then a new table must be created within the **ControlPointMetadata** schema to accommodate the multiple values for each document.

All multivalue tables should also include a repository identifier and a MD5 hash of the document DREREFERENCE. **ControlPointMetadata** also comprise of hash table types. These tables are utilized to reduce the storage footprint for information that is readily repeated. Each hash table has the same basic format comprising a repository identifier, a raw value and a MD5 hash of the raw value.

Examples

This section documents the steps required to capture additional metadata into HPE Storage Optimizer MetaStore. It uses a number of examples to do so and includes corresponding SQL statements that need to be loaded and executed.

The examples make use of metadata fields `AU_DOCUMENT_EDITOR_STRING` and `AU_DOCUMENT_AUTHOR_STRING` to illustrate the differences between SVF and MVF table setup.

For any new field that is added to metadata, it needs to be added to the appropriate field type in `FieldTypeInfo`.

> **NOTE:**
> `AU_DOCUMENT_AUTHOR_STRING` is already captured in HPE Storage Optimizer MetaStore by default.

Example 1 – single value for the same document

Documents comprise a single `AU_DOCUMENT_EDITOR_STRING` value.

This will be recorded in the **ControlPointMetadata.Additional** table in a new field named **LastEditedBy**. Data mappings must be configured to instruct the MetaStore service on how to capture and record this field value during document ingestion.

To map data

1. In SQL Server, add a new column to the **ControlPointMetadata.Additional** table to support the capture of the `AU_DOCUMENT_EDITOR_STRING` string value:

```
USE ControlPointMetaStore
GO
ALTER TABLE ControlPointMetadata.Additional
ADD LastEditedBy NVARCHAR(255) NULL
GO
```

2. Configure `AU_DOCUMENT_EDITOR_STRING` data mapping using the `MetaStore.MapField` stored procedure:

```
USE ControlPointMetaStore
GO
EXEC MetaStore.MapField
@SourceName            = 'AU_DOCUMENT_EDITOR_STRING',
@TargetTable           = 'ControlPointMetadata.Additional',
@TargetColumn          = 'LastEditedBy',
@TargetTransform       = 'ToString'
GO
```

3. Refresh document ingest, import and update sequences to support the newly captured `AU_DOCUMENT_EDITOR_STRING` field in MetaStore.

```
USE ControlPointMetaStore
GO
EXEC MetaStore.ConfigureAddDocument
EXEC MetaStore.ConfigureUpdateDocument
EXEC ControlPointMetadata.ConfigureImportDocument
GO
```

4. Restart the ControlPoint MetaStore service to utilize the refreshed sequences.

5. If you add custom fields in Insert Configuration, you must restart the HPE Storage Optimizer Engine.

Example 2 – single value hash for the same document

Documents comprise a single `AU_DOCUMENT_EDITOR_STRING` value. This example assumes that this string value is readily repeated throughout.

A new hash table, **ControlPointMetadata.EditorHash**, will be created to help reduce storage footprint.

A MD5 hash of `AU_DOCUMENT_EDITOR_STRING` will be recorded in the **ControlPointMetadata.Additional** table in a new field named **LastEditedByHash**. Data mappings must be configured to instruct the MetaStore service on how to capture and record this field value during document ingestion

To map data

1. Create a new hash table, **ControlPointMetadata.EditorHash**, to support the `AU_DOCUMENT_EDITOR_STRING` string value and MD5 hash value mappings.

```
USE ControlPointMetaStore
GO
IF OBJECT_ID(N'ControlPointMetadata.EditorHash', N'U') IS NULL
BEGIN
```

```
CREATE TABLE ControlPointMetadata.EditorHash
(
    RepositoryId    INTEGER         NOT NULL,
    HashKey     BINARY(8)           NOT NULL,
    Value           NVARCHAR(255)       NOT NULL,
    CONSTRAINT  ControlPointMetadata_EditorHash_PK
    PRIMARY KEY NONCLUSTERED(RepositoryId, HashKey) WITH FILLFACTOR = 80
)
END
GO
```

2. Add a new column to the **ControlPointMetadata.Additional** table to support the MD5 hash of the AU_DOCUMENT_EDITOR_STRING string value.

```
USE ControlPointMetaStore
GO
ALTER TABLE ControlPointMetadata.Additional
ADD LastEditedByHash BINARY(8) NULL
GO
```

3. Create a foreign key relationship from the source table to the corresponding hash table.

```
USE ControlPointMetaStore
GO
ALTER TABLE ControlPointMetadata.Additional
ADD CONSTRAINT  ControlPointMetadata_Additional_FK_LastEditedByHash
FOREIGN KEY (RepositoryId, LastEditedByHash)
REFERENCES ControlPointMetadata.EditorHash(RepositoryId, HashKey)
GO
```

4. Configure AU_DOCUMENT_EDITOR_STRING data mapping using the MetaStore.MapField stored procedure.

```
USE ControlPointMetaStore
GO
EXEC MetaStore.MapField
@SourceName         = 'AU_DOCUMENT_EDITOR_STRING',
@TargetTable        = 'ControlPointMetadata.Additional',
@TargetType         = 'SVF',
@TargetColumn       = 'LastEditedByHash',
@TargetTransform    = 'HashValue',
@SupportingTable    = 'ControlPointMetadata.EditorHash'
GO
```

5. Refresh document ingest, import and update sequences to support the newly captured AU_DOCUMENT_EDITOR_STRING field in HPE Storage Optimizer MetaStore.

```
USE ControlPointMetaStore
GO
EXEC MetaStore.ConfigureAddDocument
EXEC MetaStore.ConfigureUpdateDocument
```

```
EXEC ControlPointMetadata.ConfigureImportDocument
GO
```

6. Restart the HPE Storage Optimizer MetaStore service to utilize the refreshed sequences.

7. If you add custom fields in Insert Configuration, you must restart the HPE Storage Optimizer Engine.

Example 3 – multiple values for the same document

Documents can comprise multiple AU_DOCUMENT_AUTHOR_STRING values. These will be recorded in the **ControlPointMetadata.Author** table. Data mappings must be configured to instruct the MetaStore service on how to capture and record these field values during document ingestion.

To map data

1. Create a table, **ControlPointMetadata.Author** to record all AU_DOCUMENT_AUTHOR_STRING values for each document.

```
USE ControlPointMetaStore
GO
IF OBJECT_ID(N'ControlPointMetadata.Author', N'U') IS NULL
BEGIN
CREATE TABLE ControlPointMetadata.Author
(
        RepositoryId            INTEGER          NOT NULL,
        DocKey                  BINARY(8)                NOT NULL,
        Author                  NVARCHAR(255)            NOT NULL
        CONSTRAINT  ControlPointMetadata_Author_PK
        PRIMARY KEY CLUSTERED(RepositoryId, DocKey, Author)
        WITH FILLFACTOR = 80
)
END
GO
```

2. Configure AU_DOCUMENT_AUTHOR_STRING data mapping using the MetaStore.MapField stored procedure.

```
USE ControlPointMetaStore
GO
EXEC MetaStore.MapField
        @SourceName                     = 'AU_DOCUMENT_AUTHOR_STRING',
        @TargetTable                    = 'ControlPointMetadata.Author',
        @TargetType                     = 'MVF',
        @TargetColumn                   = 'Author',
        @TargetTransform                = 'ToString'
GO
```

3. Refresh document ingest, import and update sequences to support the newly captured AU_DOCUMENT_AUTHOR_STRING field in MetaStore.

```
USE ControlPointMetaStore
GO
```

```
EXEC MetaStore.ConfigureAddDocument
EXEC MetaStore.ConfigureUpdateDocument
EXEC ControlPointMetadata.ConfigureImportDocument
GO
```

4. Restart the HPE Storage Optimizer MetaStore service to utilize the refreshed sequences.

5. If you add custom fields in Insert Configuration, you must restart the HPE Storage Optimizer Engine.

Example 4 – multiple values hashed for the same document

Documents can comprise multiple AU_DOCUMENT_AUTHOR_STRING values. This example assumes that these string values are readily repeated throughout.

A new hash table, **ControlPointMetadata.AuthorHash**, will be created to help reduce storage footprint. Hashed AU_DOCUMENT_AUTHOR_STRING values for each document will be stored in **ControlPointMetadata.Author**. Data mappings need configured to instruct the MetaStore service on how to capture and record these field values during document ingestion.

To map data

1. Create a new hash table, **ControlPointMetadata.AuthorHash**, to support the AU_DOCUMENT_AUTHOR_STRING string value and MD5 hash value mappings.

```
USE ControlPointMetaStore
GO
IF OBJECT_ID(N'ControlPointMetadata.AuthorHash', N'U') IS NULL
BEGIN
CREATE TABLE ControlPointMetadata.AuthorHash
(
        RepositoryId    INTEGER         NOT NULL,
        HashKey BINARY(8)               NOT NULL,
        Value           NVARCHAR(255)           NOT NULL,
        CONSTRAINT  ControlPointMetadata_AuthorHash_PK
        PRIMARY KEY NONCLUSTERED(RepositoryId, HashKey) WITH FILLFACTOR = 80
)
END
GO
```

2. Create a table, **ControlPointMetadata.Author** to record all MD5 hashes for AU_DOCUMENT_AUTHOR_STRING values for each document.

```
USE ControlPointMetaStore
GO
IF OBJECT_ID(N'ControlPointMetadata.Author', N'U') IS NULL
BEGIN
CREATE TABLE ControlPointMetadata.Author
(
        RepositoryId            INTEGER         NOT NULL,
        DocKey                  BINARY(8)               NOT NULL,
```

```
        AuthorHash              BINARY(8)              NOT NULL
        CONSTRAINT  ControlPointMetadata_Author_PK
        PRIMARY KEY CLUSTERED(RepositoryId, DocKey, AuthorHash)
        WITH FILLFACTOR = 80,
        CONSTRAINT  ControlPointMetadata_Author_FK_AuthorHash
        FOREIGN KEY (RepositoryId, AuthorHash)
        REFERENCES ControlPointMetadata.AuthorHash(RepositoryId, HashKey)
)
END
GO
```

3. Configure `AU_DOCUMENT_AUTHOR_STRING` data mapping using the `MetaStore.MapField` stored procedure.

```
USE ControlPointMetaStore
GO
EXEC MetaStore.MapField
        @SourceName                 = 'AU_DOCUMENT_AUTHOR_STRING',
        @TargetTable                = 'ControlPointMetadata.Author',
        @TargetType                 = 'MVF',
        @TargetColumn               = 'AuthorHash',
        @TargetTransform            = 'HashValue',
        @SupportingTable            = 'ControlPointMetadata.AuthorHash'
GO
```

4. Refresh document ingest, import and update sequences to support the newly captured `AU_DOCUMENT_AUTHOR_STRING` field in MetaStore.

5. Restart the HPE Storage Optimizer MetaStore service to utilize the refreshed sequences.

6. If you add custom fields in Insert Configuration, you must restart the HPE Storage Optimizer Engine.

Existing data and re-ingestion

The steps outlined in the examples ensure that the new field, `AU_DOCUMENT_EDITOR_STRING`, is captured for new document files being ingested.

Existing data will need to be re-ingested in order to capture values for this new metadata field.

> **NOTE:**
> If you add custom fields in Insert Configuration, you must restart the HPE Storage Optimizer Engine so that HPE Storage Optimizer picks up the new custom fields.

To re-ingest data

- select **Re-Index Repository** on the Repositories dashboard.

- remove the connector database file from the connector installation directory, followed by a connector service restart.

Field text and advanced properties

The new metadata has been captured into HPE Storage Optimizer MetaStore through document ingestion. In order to make use of this new data for field text purposes and to return as part of the Properties/Advanced Properties within the HPE Storage Optimizer Dashboard, a number of further changes are required.

Field Text

In order to make the new field available within the category field text builder, a new Rule Builder Fields mapping must be configured within the HPE Storage Optimizer Administration Dashboard.

To support this, a database view modification must be made to ensure the new field is available from the list of rule builder available fields in the HPE Storage Optimizer UI.

To add a new field within the category field text builder

1. Open SQL Management Studio and expand **Databases > ControlPointMetaStore > Views**.

   a. Select **MetaStorePro.FieldTypeInfo**, right click and click **Script View as > Alter To > New Query Editor Window**.

      > **NOTE:**
      > For any new field that is added, it needs to be added to the appropriate field type in `FieldTypeInfo`.

      Examples:

      - A new field, `AU_DOCUMENT_EDITOR_STRING`, must be appended to both 'Match' and 'RulesBuilderInc' FieldType list of supported fields and then executed.

      - A new date field must be appended to both the 'NumericDate' and 'RulesBuilderInc' FieldType list of supported fields and then executed.

2. On the HPE Storage Optimizer Administration dashboard, click **Settings**.

   The Settings page opens.

   a. On the General tab, select **Fields**. In the Rule Builder section, add a new field by clicking **Add** (+).

      The Add New Field page opens.

   b. Enter a name for the new field in the **Display Name** box.

   c. Select the new metadata field from the **Fields** list.

   d. Click **Add**.

      After the new field mapping is added, the new metadata captured into MetaStore can be used for category training purposes.

Properties and Advanced Properties

The new field is available within the HPE Storage Optimizer UI in the Advanced Properties list after you restart Internet Information Service (IIS).

To configure a new property mapping

1. On the HPE Storage Optimizer Administration dashboard, click **Settings.**

   The Settings page opens.

2. On the General tab, select **Fields**. In the Item Properties section, add a new item property by clicking **Add** (+).

   The Add Property page opens.

3. Enter a name for the new property in the **Display Name** box.

4. Select the type from the **Type** list.

5. Select the new metadata field from the **Fields** list.

6. Click **Add**.

Appendix A: Statistics Export Utility

You can use the ControlPoint Statistics Export Utility to export data to Microsoft Excel. The type of data exported depends on the state of the repository.

- Statistics can be exported from any analyzed repository.
- Metrics can be requested from any unanalyzed repository.

Sample Microsoft Excel templates are provided with the utility.

Before you begin

Install Microsoft Excel to the HPE Storage Optimizer server.

To export statistics

1. Run the Statistics Export Utility, which is available at the following location:

   ```
   Storage Optimizer x64\HPE Storage Optimizer Utilities\Statistics Export
   Utility\ControlPointStatisticsUtility.exe
   ```

   The HPE Storage Optimizer Analysis window opens.

2. Enter the host name in the **Host** box, and then click **OK**.

   The export dialog box appears. The Analysis Tasks section lists all analyzed repositories on the host system.

3. (Optional) To re-analyze a repository, select it, and then click **Re-analyze**.

4. (Optional) **To add a custom analysis task**

   a. Click **New**.

      The New Custom Analysis Task dialog box opens.

   b. Enter a **Task Name**.

   c. Click **OK**.

      The Task is added to the list.

5. Select an analysis task.

6. In the **Export Task** section, select a Microsoft Excel template from the list, and then click **Export**.

   The data exports to Excel and appears according to the selected template. Potential Obsolete and Trivial disk space appears in the Obsolete-AllPotential and Trivial-AllPotential charts.

Statistics Export Utility command line interface

You can use the Statistics Utility command line interface to export results.

Location

```
Storage Optimizer x64\HPE Storage Optimizer Utilities\Statistics Export
Utility\ControlPointStatisticsUtility.exe
```

Synopsis

```
ControlPointStatisticsUtility.exe -dahost <hostname> -enablehttps 0|1
-sqlhost <hostname> -authtype 0|1 -dataset <repo> -action 0|2|3
-templatepath <path> -exportpath <path>
```

Options

Parameter	Required	Description
-dahost	Required	Specify the host name of the Data Analysis service machine.
-sqlhost	Required	Specify the host name of SQL Server machine.
-authtype	Required	Specify the SQL Server authentication type: 0 is Windows user authentication 1 is SQL Server user authentication
-enablehttps	Required	Specify whether the enable HTTPS. 0 is no 1 is yes **NOTE:** Only set to 1 when Storage Optimizer environment is enabled with HTTPS.
-dataset *<repo>*	Optional	Specify the data set to take action on. Required for export. **NOTE:** The string for *<repo>* is case sensitive. For example: `-dataset fileType`
-action	Required	Specify the type of action to perform: • 0 is export. • 2 is re-analyze. • 3 is delete.
-sqluser	Optional	Specify the user name of a SQL Server user. **NOTE:** Required when `-authtype` is set to 1.
-password	Optional	Specify the password of the SQL Server user.

Parameter	Required	Description
		NOTE: Required when `-authtype` is set to 1.
-templatepath	Optional	Absolute path of the template file. **NOTE:** Required when the `-action` is set to export (0).
-exportpath	Optional	Absolute path of the export file. **NOTE:** Required when the `-action` is set to export (0).
-taskname	Optional	Name of task to be re-analyzed or deleted. **NOTE:** Required when the `-action` is set to re-analyze (2) or delete (3).

Examples

To export data

```
ControlPointStatisticsUtility.exe -dahost cpserver -enablehttps 0
-sqlhost cpserver -authtype 0 -dataset repo -action 0
-templatepath C:\test\Templates\Blank.xltx -exportpath C:\test\export\repo.xlsx
```

To re-analyze a repository

```
ControlPointStatisticsUtility.exe -dahost cpserver -enablehttps 0
-sqlhost cpserver -authtype 0 -action 2 -taskname myTask
```

To delete a task

```
ControlPointStatisticsUtility.exe -dahost cpserver -enablehttps 0
-sqlhost cpserver -authtype 0 -action 3 -taskname myTask
```

Appendix B: Archiving command line utility

The Archiving command line utility archives and stubs a file. In addition, this command recreates a file or directory stub, rehydrates a stubbed file or directory, and dumps the reparse data contents of a stubbed file.

Currently, this utility works on Microsoft .NET Framework 4.5 and supports the following features:

- Archives and stubs a file
- Recreates a file stub
- Recreates all file stubs
- Dumps the contents of the reparse data of a stubbed file
- Rehydrates a stubbed file
- Rehydrates all stubbed files
- Deletes the stubbed source file

Location

```
\Program Files\Hewlett Packard Enterprise\Storage Optimizer\Edge\Archive
Service\Stub\stub.exe
```

Synopsis

```
stub.exe -create -source <path> -archiveP01 <path> -archiveP2 <path>

stub.exe -recreate [-source <path>] -archive <path> -r

stub.exe -rehydrate -source <path> -r

stub.exe -dump -source <path>

stub.exe -delete -source <path>

stub.exe -|-help
```

Options

`-create`

Archives the data from the file specified in thne `-source` parameter and replaces the original file with the reparse point file stub, which contains the information required by the filter drive to recognize the file for the archive redirection.

This option includes the following parameters:

- `-source`: Specify the location of the source file. It can be a file name with the full path or a directory.
- `-archiveP01` (required): Specify the location for archiving the file. It should be a directory.
- `-archiveP02` (optional): Specify an additional location for archiving the file. This too should be a directory.

After the copy operation, the archive location is modified. The source system name and the source path are appended to the archive location.

Example

If the source system name is `SRC_SYS`, the source file is `D:\Logfiles\Monday\Log1.txt`, and the original archive location is `Z:\archive`, then the modified archive location is `Z:\archive\SRC_SYS\D\Logfiles\Monday`, which now includes two files: `Log1.txt` and `Log1.stb`.

> **NOTE:**
> All parameters of this option are mandatory. Also, the `readonly` and `nodelete` parameters are specific to the stub file through the actions handled by the filter driver and are always ON.

`-recreate`

Recreates stub files if they are deleted or damaged due to user actions or file system issues. You can recreate a stub by using the two files at the archive location. You should copy the file from the archive location back to the source location where the stub was deleted.

As a next step, the corresponding `.stb` file that exists along side the archive version of the file is used to replace the reparse information back on the file, and then the file is once again made a sparse file. The copying operation of the archived file to the source skips the main data stream and only copies attributes and alternate data streams.

This option includes the following parameters:

- `-source`: Specify the file name (with the complete path) or directory where the stub file is to be created. This parameter is optional. If it is not specified, the original source location is chosen.

- `-archive`: Specify the file location where the file is archived.

- `-r`: Specify this parameter for the recursive operation. You need to explicit provide this parameter, as it is not recursive by default.

`-rehydrate`

Restores the archived file from the archive, replacing the stubbed file.

This option includes the following parameters:

- `-source`: Specify the file name and the complete path of the stub file.

- `-r`: Specify this parameter for the recursive operation. You need to explicit provide this parameter, as it is not recursive by default.

`-dump`

Displays to the console and logs the reparse metadata stored in the reparse point of the stubbed file, which is specified using the `-source` parameter.

This option includes the following parameter:

- `-source`: Specify the file name and the complete path of the stub file.

`-delete`

Deletes the stubbed file specified using the `-source` parameter. The Windows delete commands are prevented from deleting the stub file when the `nodelete` flag is set. This is enforced by the filter driver.

This option includes the following parameter:

- `-source`: Specify the file name and the complete path of the stub file.

```
-help
```

Displays the usage synopsis for this command line utility.

> **NOTE:**
> If no option is provided with the command, it lists all the available options, with their parameters.

Examples

1. To archive a file to a shared location in your network, run:

   ```
   stub.exe -create -source C:\src\test.txt -archiveP01 \\dest_sys\share1
   ```
   where

   \\dest_sys\share1 is the shared location in your network where the file is archived

2. To recreate a file stub, run:

   ```
   stub.exe -recreate -source C:\src\test.txt -archive \\dest_sys\share1\SRC_
   SYS\c\src\test.txt
   ```
   where

   SRC_SYS is the source system name

3. To rehydrate a stubbed file in the source location, run:

   ```
   stub.exe -rehydrate -source C:\src\test.txt
   ```

   If this command is successful, the file is no longer a stubbed, offline file. It represents the complete file prior to it being archived.

4. To rehydrate any archived file in the source location, run:

   ```
   stub.exe -rehydrate -source C:\src
   ```

5. To rehydrate any archived file in the source and its child directories, run:

   ```
   stub.exe -rehydrate -source C:\src -r
   ```

6. To dump the reparse metadata in the stubbed file's reparse point, run:

   ```
   stub.exe -dump -source C:\Store1\MonthlyAssets.pdf
   ```
   Sample output:
   ```
   Dump: File: c:\Store1\MonthlyAssets.pdf
   Dump: readonly: 1
   Dump: nodelete: 1
   Dump: assetversion: 1
   Dump: source c:\Store1\MonthlyAssets.pdf
   Dump: target: \\SHARE1\D\Store1\MonthlyAssets.pdf
   Dump: target: \\SYSTEMA\D\Store1\MonthlyAssets.pdf
   Dump: AssetId: StubLocalFile
   ```

7. To delete a stubbed file, run:

   ```
   stub.exe -delete -source C:\Store1\MonthlyAssets.pdf
   ```

Appendix C: Support Utility

The ControlPoint Support utility captures system information and configuration file information from your Storage Optimizer environment.

The utility supports the following modes:

- User interface — captures the information and generates a ZIP archive of the results and the report file.
- Command line — see Synopsis, below for command line options and examples.

> **NOTE:**
> Command line enhancements are supported for Storage Optimizer 5.4 and later.
>
> For versions 5.3 or earlier, run the utility with the user interface.

Location

```
\Program Files\Hewlett Packard Enterprise\Storage
Optimizer\Engine\Scheduler\ControlPointSupportUtility.exe
```

Synopsis

```
ControlPointSupportUtility.exe

ControlPointSupportUtility.exe -c
```

Options

`No option`

Generates a ZIP archive of the results and the xml/xslt browser report file.

`-c`

Moves the data to the `\<user>\AppData\Local\Temp` directory for comparison. Does not generate a ZIP archive of the results or the report file.

To generate a report that contains comparison results, you must run the utility with the `-c` option twice.

Example

> **NOTE:**
> The following example applies to ControlPointStorage Optimizer versions 5.4 and later. If you are running version 5.3 or earlier, this example does not apply.

Run the utility as a preparatory step when changing the Storage Optimizer environment.

1. Run the Support utility from the command line as the Administrator.

   ```
   ControlPointSupportUtility.exe -c
   ```

   The utility gathers and copies all of the system information and configuration file information and label it as `Pre` capture data.

2. Perform the changes to the environment.

3. Run the Support utility to gather the data and label it as `Post` data.

   ```
   ControlPointSupportUtility.exe -c
   ```

   The utility runs a comparison feature, which generates a report named `diffReport.txt`. The ControlPoint Support Utility creates the report in the same directory as the utility.

   The report lists any differences between the two `SystemInfo.xml` files, including changes, additions and deletions. In addition, it lists any differences between all configuration files located in the Storage Optimizer installation directory.

Results

When the utility is run with the `-c` option, the locations of the `Pre` and `Post` data files are as follows:

```
<systemroot>\Users\<user>\AppData\Local\Temp\PreLogFiles
```

```
<systemroot>\Users\<user>\AppData\Local\Temp\PostLogFiles
```

```
<systemroot>\Users\<user>\AppData\Local\Temp\PreSystemInfo.xml
```

```
<systemroot>\Users\<user>\AppData\Local\Temp\PostSystemInfo.xml
```

Appendix D: Repository command-line utility

This utility allows you to create new repositories from the command line interface.

Supported repository types

- File System
- SharePoint 2010
- SharePoint 2013
- SharePoint 2016
- SharePoint Remote
- Exchange

Location

`Storage Optimizer x64\HPE Storage Optimizer Utilities\CommandLine Utility`

Synopsis

`StorageOptimizerCommandLine.exe -action repo_create -config_path C:\<configPath>\<fileName>.xml -report_path C:\<reportPath> -enablehttps 0|1`

Options

The options include the following parameters:

Parameter	Required	Description
`-action repo_create`	Required	Creates a repository.
`-config_path <configPath>\<fileName>.xml`	Required	Specify the absolute path to the XML file of repository configuration parameters. See Examples, on page 151.
`-report_path <reportPath>`	Required	Specify the absolute path to a summary report of the utility run.
`-enablehttps`	Required	Specify whether the enable HTTPS. 0 is no

Parameter	Requi red	Description

		1 is yes
		
		NOTE:
		Only set to 1 when Storage Optimizer environment is enabled with HTTPS.
		
		If want to use https, update `ThumbprintValue` and `ControlPointHostDefault` value in configuratuion file under <appSettings>.
		
		`<appSettings>`
		
		`<add key="ControlPointHostDefault" value="localhost" />`
		`<add key="ThumbprintValue" value="PUT_THUMBPRINT_ VALUE_HERE" />`
		`</appSettings>`
		
		For example:
		
		`<appSettings>`
		
		`<add key="ControlPointHostDefault" value="myControlntPointHost.myDomain.com" />`
		`<add key="ThumbprintValue" value="4a 46 9f b8 42 01 25`
		`26 77 53 e2 e5 31 6f 6d 65 f5 b5 4a 10" />`
		`</appSettings>`

Notes

- If you specify Content, restart the IDOL service after the repositories are created.
- To generate properties for the `<Properties>` section in the XML, run the following query:

```
SELECT cd.[Id]
      ,[InternalName]
      ,[FilteringEnabled]
      ,[LookupType]
      ,cm.Id as ValueId
      ,cm.DisplayValue
  FROM [ControlPoint].[dbo].[CPMetadataDefinition] cd,
ControlPoint.dbo.CPMetadataLookup
  cm WHERE cm.MetadataDefinitionId = cd.Id AND cd.InternalName = 'abc2'
```

Examples

File System connector XML

```xml
<?xml version="1.0" encoding="utf-8" ?>
<Repository>
  <details>
    <name>FileSys_Repo_100</name>
    <description>Test FileSystem repository created by utility</description>
    <repo_type>Filesystem</repo_type>
    <connector>CPWIN12R2</connector>
    <aci_port>7200</aci_port>
    <service_port>7202</service_port>
  </details>
  <settings>
    <path>\\CPWIN12R2\share1</path>
    <include_type>txt,doc</include_type>
        <!-- e.g., txt,doc -->
    <exclude_type>exe,pdf</exclude_type>
        <!-- e.g., exe,pdf -->
  </settings>
  <analysis>
    <analysis_type>Metadata_Only</analysis_type>
        <!-- also support: Repository_Metadata_Only; Content -->
    <permissions_and_ownership>Yes</permissions_and_ownership>
    <analyze_subitems>Yes</analyze_subitems>
    <default_tag>Category1</default_tag>
        <!-- this is for the categoryName -->

  </analysis>
  <Properties>
     <Property>
         <id>27</id>
         <InternalName>abc2</InternalName>
         <FilteringEnabled>Yes</FilteringEnabled>
         <ValueId>3</ValueId>
         <DisplayValue>2</DisplayValue>
     </Property>
  </Properties>
  <schedule>
    <start_time>11:45</start_time>
        <!-- also can be time, e.g, 21:35 -->
    <cycle>Run_Once</cycle>
        <!-- can also be Run_Forever -->
    <recurrence_number>2</recurrence_number>
    <recurrence_unit>Hours</recurrence_unit>
        <!-- can also be: Minutes, Days, Weeks -->
```

```
      </schedule>
  </Repository>
```

SharePoint connectors XML

```
  <?xml version="1.0" encoding="utf-8" ?>
  <Repository>
    <details>
      <name>SP_Remote_Repo_1</name>
      <description>Test1 SP Remote repository created by utility</description>
      <repo_type>SharePointRemote</repo_type>
      <connector>YourConnectorComputerName</connector>
      <aci_port>7800</aci_port>
      <service_port>7802</service_port>
    </details>
    <settings>
                <url>http://v-qa-moss</url>
                <!-- URL to your remote sharepoint server -->
    </settings>
    <analysis>
      <analysis_type>Repository_Metadata_Only</analysis_type>
        <!-- also support: Repository_Metadata_Only; Content -->
      <permissions_and_ownership>No</permissions_and_ownership>
      <analyze_subitems>Yes</analyze_subitems>
      <default_tag>CategoryName</default_tag>
        <!-- this is for the categoryName -->

    </analysis>
    <Properties>
    </Properties>
    <schedule>
      <start_time>Now</start_time>
        <!-- also can be time, e.g, 21:35 -->
      <cycle>Run_Once</cycle>
        <!-- can also be Run_Forever -->
      <recurrence_number>2</recurrence_number>
      <recurrence_unit>Days</recurrence_unit>
        <!-- can also be: Minutes, Days, Weeks -->
      </schedule>
  </Repository>
```

Exchange connector XML

```
  <?xml version="1.0" encoding="utf-8" ?>
  <Repository>
    <details>
      <name>Exch_Repo_1</name>
      <description>Test1 Exch repository created by utility</description>
      <repo_type>Exchange</repo_type>
      <connector>YourConnectorComputerName</connector>
```

```
    <aci_port>7600</aci_port>
    <service_port>7602</service_port>
  </details>
  <settings>
    <wsurl>https://mail.exch2010msg1.hpswlabs.hp.com/ews/exchange.asmx</wsurl>

<ldappath>LDAP://cmbge10m1dc1.exch2010msg1.hpswlabs.hp.com:389/DC=exch2010msg1,D
C=hpswlabs,DC=hp,DC=com</ldappath>
    <!-- Your exchange server WS URL and LDAP path -->
   <!-- Optional
    <domain></domain>
    <username></username>
    <password></password>
    Optional -->
  </settings>
  <analysis>
    <analysis_type>Metadata_Only</analysis_type>
       <!-- also support: Repository_Metadata_Only; Content -->
    <permissions_and_ownership>No</permissions_and_ownership>
    <analyze_subitems>Yes</analyze_subitems>
    <default_tag>CategoryName</default_tag>
       <!-- this is for the categoryName -->

  </analysis>
  <Properties>
  </Properties>
  <schedule>
    <start_time>Now</start_time>
       <!-- also can be time, e.g, 21:35 -->
    <cycle>Run_Once</cycle>
       <!-- can also be Run_Forever -->
    <recurrence_number>2</recurrence_number>
    <recurrence_unit>Hours</recurrence_unit>
       <!-- can also be: Minutes, Days, Weeks -->
  </schedule>
</Repository>
```

Send documentation feedback

If you have comments about this document, you can contact the documentation team by email. If an email client is configured on this system, click the link above and an email window opens with the following information in the subject line:

Feedback on Administration Guide (HPE Storage Optimizer 5.5)

Add your feedback to the email and click **Send**.

If no email client is available, copy the information above to a new message in a web mail client, and send your feedback to AutonomyTPFeedback@hpe.com.

We appreciate your feedback!