Hewlett Packard
Enterprise

HPE Storage Optimizer

Software Version: 5.2

Administration Guide

Document Release Date: March 2016
Software Release Date: March 2016

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

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 updates or to verify that you are using the most recent edition of a document, go to: **https://softwaresupport.hp.com**

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.hp.com/hppcf/createuser.do**

Or click the **Register** link at the top of the HPE Software Support page.

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.hp.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
- 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.hp.com/hppcf/createuser.do

To find more information about access levels, go to:

https://softwaresupport.hp.com/web/softwaresupport/access-levels

HPE Software Solutions Now accesses the HPESW Solution and Integration Portal Web site. This site enables you to explore HPE Product Solutions to meet your business needs, includes a full list of Integrations between HPE Products, as well as a listing of ITIL Processes. The URL for this Web site is **http://h20230.www2.hp.com/sc/solutions/index.jsp**

About this PDF Version of Online Help

This document is a PDF version of the online help. This PDF file is provided so you can easily print multiple topics from the help information or read the online help in PDF format. Because this content was originally created to be viewed as online help in a web browser, some topics may not be formatted properly. Some interactive topics may not be present in this PDF version. Those topics can be successfully printed from within the online help.

Contents

Chapter 1: Introduction

This chapter provides an overview of the following.

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.

Storage Optimizer Architecture

Storage Optimizer has a Web application user interface. Functionality is available through several Dashboards.

Components

Storage Optimizer includes the following components.

- Storage Optimizer Dashboard
- Storage Optimizer Engine

- Storage Optimizer Data Analysis
- Storage Optimizer Connectors

Storage Optimizer Dashboard

The 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 service is included in the Storage Optimizer.

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

Storage Optimizer Engine

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

The following services are included in the Storage Optimizer Engine.

- **Storage Optimizer Engine service** is a Windows service that executes all scheduled tasks
- **CallbackHandler** is an IIS Web application that receives notifications from IDOL connectors

Storage Optimizer Data Analysis Service

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

Storage Optimizer Connectors

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

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

- The **Storage Optimizer StoreAll Connector** service scans and performs actions on items in StoreAll repositories. This connector type has a connector framework deployed alongside.

Related Documentation

The following documents provide more detail on Storage Optimizer.

- *Storage Optimizer Installation Guide*
- *Storage Optimizer Remote Analysis Agent Technical Note*
- *Storage Optimizer Legacy Data Cleanup Administration Guide*

The following documents provide more detail on connectors.

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

Chapter 2: Manage Role-Based Security

This section explains how to add users and how to apply role-based security in Storage Optimizer.

Introduction

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. You can use the Configuration Manager to identify a Storage Optimizer System Administrator and an LDAP server/base Distinguished Name. The System Administrator can 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 Storage Optimizer Security

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

To enable role-based security

1. Open the Storage Optimizer Configuration Manager.
2. On the Security panel, select **Enable Security**.
3. Under System Administrator Account, identify a Storage Optimizer System Administrator by typing a **Domain** and a **Username**.
4. Under Active Directory Settings, identify the LDAP server by typing 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. For example, for Base DN entered as `'file:\\MACHINENAME\folder\userFile.xml'`, the sample file structure can be as follows:

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

```
            <group name="Test1" displayName="Test Group 1" />
            <group name="Test2" displayName="Test Group 2" />
      </user>
      <user name="hpe\$Ciaran.Storage Optimizer" displayName="$Ciaran.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 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 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
  - **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.

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. In the upper right is a drop-down menu that you can use to filter the permissions by type: Administration, Category, Policy, Repository, or the default, All Security.
2. (Optional) **To add a user or user group**
   a. Click **Add**.
   b. Type the name of the user or user group in the text box, and then 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.
3. To edit permissions, click **Edit** by the desired user or group.

   A Permissions dialog box opens.
4. Select one or more roles to assign to the user or group (see "User Roles" on the previous page for more details on roles), and then click **Apply**.

   The selected roles appear 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.
5. When you finish adding users and groups and setting the permissions, click **Save**.

   The security settings apply 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 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

Use the next procedure to set All Category security 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. Type 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 (see "User Roles" on page 14 for details on user roles).

   The selected roles appear 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.

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. Type 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 (see "User Roles" on page 14 for details on user roles).

   The selected roles appear 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 button in the upper-right corner of the repository or policy panel.
2. Click **Security**.

   The Security dialog box opens.
3. (*Optional*) **To add a user**
   a. Click **Add**.
   b. Type 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 (see "User Roles" on page 14 for details on user roles).

   The selected roles appear 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.

Chapter 3: Manage Repositories

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

Repositories

Storage Optimizer manages content that you scan into MetaStore. Repositories provide a view of the data that Storage Optimizer manages. The Register Repositories scheduled task discovers new repositories and registers them in 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: Storage Optimizer is limited to managing documents in defined repositories.

Different repositories support different types of policy actions. The following tables list supported policy phases by repository.

See "Related Documentation" on page 11 for a list of connector documents.

	Dispose
Exchange	x
File System	x
Hadoop	x
Notes	x
SharePoint 2007	x

	Dispose
SharePoint 2010	x
SharePoint 2013	x
SharePoint Remote	x
StoreAll	x

	Secure (Remove)	Secure (Leave)	Secure (Link Shortcut)
Exchange	x	x	
File System	x	x	x
Hadoop	x	x	
Notes	x	x	
SharePoint 2007	x	x	
SharePoint 2010	x	x	
SharePoint 2013	x	x	
SharePoint Remote	x	x	
StoreAll	x	x	

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 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 62.

You can change the repository status manually. See "Change Repository Status" on page 24.

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, which is accessible through 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**. Type the repository name.
   - **Description**. Type 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.

     Note: To use Archive policies, you must select the File System Agent. You can specify the archive policies in the Direct Policy Execution field under the Settings section.
4. The necessary settings are dynamically loaded into the Details section after you select a connector.
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. This Analysis type is available only if you select the File System Agent for the Filesystem repository.
     - **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 85).

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 determine the repository schedule.
   - **Start Time**. Specify a start time, if necessary. **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, if necessary. 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 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 85).

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 either:
  -  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 85).
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 85).

The properties appear on the left of the Repositories dashboard's menu bar.

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  .

   A Filter dialog box opens.

2. Type filter text in the dialog 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 such 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**, type 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.
     i. Click **Add**.

        The Add New Repository dialog box opens.

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

     iii. Click **Save**.

4. (*Optional*) In the **Properties** group, add any required custom properties (see "Custom Properties" on page 85).
   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 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 button.
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 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 Storage Optimizer to apply policies to the repository content automatically (see "Apply Policies Automatically" on page 62).

To change the status of a repository

1. On the **Repositories** dashboard, click the menu button in the upper-right corner of the repository.

   A menu opens. Depending on the current status of the repository, the 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 opens. 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, click the menu button in the upper-right corner of a repository.

   A menu opens.
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.

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 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 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, click the menu button in the upper-right corner of the repository.

   A menu opens.
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).

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 35
   - view data by statistical analysis, by tag, or by any other method
   - combine the resulting file list with filters

   Click Actions > **Create Subset**.

   The Create Subset dialog box opens.
2. Type a **Name** and **Description**, and 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.
3. Click **Save**.
4. 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

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

You can use XML to describe the tree structure to appear when users browse the content. If no XML is provided, users can view the content, but there is no hierarchical navigation, and users have to filter the full list of content in the repository.

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.

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:
   - **Name**. Type the repository name.
   - **Description**. Type a description for the repository.
   - **Type**. Select the repository type as XML.
   - **XML Source**. Specify whether to structure the repository from XML input or using an existing XML file. Note that XML Input takes precedence.
   - **XML File Path** or **XML String**. Type the XML file path if you selected File as the option under XML Source, or provide 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 85).
   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.
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 and the associated Connector Framework Service.

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

Sample XML Repository

The following sample XML shows definition of 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>
```

Chapter 4: Clean Up Legacy Data

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

Introduction

When 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.

  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 Storage Optimizer identifies as ROT data. See "Configure Potential ROT Rule Sets" on page 40.
- 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 " on page 35 for more details.

When you identify data that requires a cleanup action, you can tag it appropriately. See "Clean Up Legacy Data" on page 39.

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.

Duplicated data is rarely necessary to maintain. It is likely that you will dispose of it during cleanup. 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 " on page 35.
3. When you identify the duplicate data you want to clean up, take the appropriate cleanup action. See "Clean Up Legacy Data" on page 39.

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 31
   - **by Size**
   - **by Type** see "View Data by File Type" on the next page
   - **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 " on the next page.
3. When you identify the data to clean up, take the appropriate cleanup action. See"Clean Up Legacy Data" on page 39.

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 39 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. See"Clean Up Legacy Data" on page 39.

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 " below.
3. When you identify the data to clean up, take the appropriate cleanup action. See"Clean Up Legacy Data" on page 39.

Common File List Operations

A number of tasks common to all file lists in 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 (IDOL) repositories.

To search within a file list

1. To the right of the file list, click  .

   The Title filter dialog box opens.
2. Type 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. Type 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 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  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" below.

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 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" on the previous page)
- 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. Type 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 36.

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 36.

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**.

   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 page 37.

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 31), 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 e-mail addresses.
- Send it for **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 42.

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 31.
2. Select the files, and then click  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, type 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

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 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. Type a **Display Name** for the master location, and select a master location from the **Repositories** list.
5. Click **Add**.

   The Add Master Location dialog box closes.
6. 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, type 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.

     For more information on categories, see "Categories" on page 67.
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. Type 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

Storage Optimizer includes four default tags: *Preserve, Protect, Remove*, and *Review* (see "Clean Up Legacy Data" on page 39 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 type 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. Type the reason in the dialog box, and then press **Enter**.

   A text box opens in the Comment section.
6. Type 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 type 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 "Add a Repository" on page 21).

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 5: Manage Target Locations

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

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. For example, 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. The following table lists supported policy phases by target location.

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

Add a Target Location

Use the following procedure to add a target location to your 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.

Name	of the target location
Description	of the target location
Connector Group	to use when sending work to the target location
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**.

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. In the upper-right corner of the target location panel, click the **Menu** button, and then click **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 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**. Type the Target URL.
   - **StoreAll**. Specify the Path and the name of the Connector Config Section.
5. Click **Save**, and then click **OK**.

Map Repository Fields to the Target Location Metadata

If you want 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, type a part of the field name in the **Source Field** text box.
   As you type, a list of matching fields opens. 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. When you finish, click **Save**.

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

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 Storage Optimizer to use when copying, moving, or declaring documents to target locations.

The 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 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. 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, type a part of a field name. As you type, a list of matching fields opens. Select a field from the list.
   - In text boxes, type 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. When you finish, 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 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. Click the menu button in the upper-right corner of the target location panel.
   A confirmation message opens.
3. Click **Delete**.

Chapter 6: Manage Policies

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

Policies

A policy defines the rules and actions to perform on information content. 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 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

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 55.

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.

In addition, 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 "Appendix A - Archiving Command Line Utility" on page 113.

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.

You can perform the following actions using 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.

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 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 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 71.

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*.

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. See "Schedule Plans" on page 73 for more information.

In the first step of execution, 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, 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, type or select the following information.
   - **Name** of the template
   - **Description** is an optional description of the policy template
   - **Phases**. Specify one or more policy phases by clicking **Add**.

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 Storage Optimizer executes the associated Action. You can use the following values. ◦ **System Default** (default) ◦ **Review** ensures that 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 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** group, 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 Storage Optimizer Engine checks the policy for items to execute.
   - **Schedule Plan** selects how frequently 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 73 for more details.
   - **Priority** determines the priority of the policy that Storage Optimizer uses during automatic conflict resolution. See "Automatically Resolve Conflicts" on page 78

5. (*Optional*) In the **Properties** group, 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** group, 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.

As an example, you can create a policy using the Archive policy template.

To create a new 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. This target location is defined in **Administration -> Target Locations**.
4. In the Execution Rules section, specify the criteria that the content must meet for 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 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. Storage Optimizer uses this priority during automatic conflict resolution. See "Automatically Resolve Conflicts" on page 78

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

Deleting Archive Policy

There are two ways of applying Delete Archive Policy. You can create a new repository or you can edit an existing archive repository. The steps to perform both operations are provided here:

1. Creating a new repository:
   - Create a new **Delete Archive Policy**.
   - Create a new repository and select the **Edge File System Connector** from the list of connectors.
   - From the settings, select the newly created Delete Archive Policy.
   - Run the policy.
2. Editing an existing archive repository:
   - Create a new "Delete Archive Policy".
   - Edit the existing archive repository.
   - From the settings, add the newly created Delete Archive Policy along with the existing archive policy.
   - Run the policies.

Create a Policy

You can create a policy without using policy templates for special cases.

To create a policy

1. On the **Policies** dashboard, click **+**.
   The Add Policy dialog box opens.
2. Select **Blank (default)** 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 one or more policy phases by clicking **Add**, and then specify the following information:

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 Storage Optimizer executes the associated Action. This option ensures that 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 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**

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 Storage Optimizer Engine checks the policy for items to execute.

- **Schedule Plan** selects how frequently 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 73 for more details.
- **Priority** determines the priority of the policy, which Storage Optimizer uses during automatic conflict resolution. See "Automatically Resolve Conflicts" on page 78
- **Policy Approver Email Address** selects the email address of the policy approvers to be notified about the review before the policy execution.

5. (*Optional*) In the **Properties** section, click **Add** to associate any properties and values that are appropriate for the policy.
6. (*Optional*) In the **Assign To** section, click **Add** to select categories. The policy will be assigned content associated with the selected categories.
7. When you finish adding phases, click **OK**.

Related Topics

- "Policy Execution Rules" on page 61

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 85.

To sort the policy list

1. Configure one or more custom properties that apply to policies. See "Create a Custom Property" on page 85.
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 85.

   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. Type filter text in the dialog 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, click the menu button in the upper-right corner of the policy.
   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 one or more policy phases by clicking **Add**, or click the Edit icon to change existing phases.

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 Storage Optimizer executes the associated Action. This option ensures that 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 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 Storage Optimizer Engine checks the policy for items to execute.
   - **Priority** determines the priority of the policy that Storage Optimizer uses during automatic conflict resolution. See "Automatically Resolve Conflicts" on page 78
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 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*), 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 Storage Optimizer applied the previous phases to.
Execution Rules	Documents must meet the new rules for Storage Optimizer to apply the action associated with the policy. The new rules do not change existing documents that Storage Optimizer previously assigned to the policy.
Categories	Storage Optimizer applies the policy to documents in the current category list.

Policy Setting	Effect
	The policy association remains for documents in other categories that the policy previously applied to.

Related Topics

- "Create a Policy" on page 56
- "Policy Execution Rules" below

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 opens.
4. Type 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 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.

Related Topics

- "Scheduled Tasks" on page 71
- "Change Repository Status" on page 24

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**.

   Storage Optimizer applies the policies to the documents.

View the Policies on Items

The 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 that are 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.

Status	Description
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.

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

- Double-click a policy from the **Policies** page.

  The Policy Summary page displays:
  - **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

- Click a rectangle at the bottom of the page to change the granularity of the time axis on the policy activity charts: *Hour, Day, Week, Month, Quarter, Year*.

Chapter 7: Categories

Categories identify what content 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 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

You can view the category settings from the Categories dashboard.

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 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 the box 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 opens. To include the category contents, leave the box checked.
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  .
2. Browse to the location of the category XML file, and then click **Open**.
3. Select whether to **Keep** or **Remove** child categories and how to handle encountered duplicates (keep existing, merge or overwrite).
4. Click **Import**.

Chapter 8: Scheduled Tasks

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

Storage Optimizer provides a number of scheduled tasks out of the box. Not all the 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 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.

- **Reevaluate 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 Storage Optimizer to remove policy for these documents. By default, a policy is not removed after it is assigned by a category.

Statistics

- **Calculate Compliance** calculates a measure of how many documents in a repository are being managed through a 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.
- **Metadata Index Tuning** rebuilds fragmented analysis.
- **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.

Default Scheduled Task Configuration

The Storage Optimizer installation installs and configures the following tasks and schedules.

Name	Interval	Enabled
Assign Policies	60 minutes	Yes
Cleanup Policies	60 minutes	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	60 minutes	Yes
Calculate Conflict Statistics	60 minutes	Yes
Metadata Compact	7 days	Yes
Metadata Consistency Check	7 days	Yes
Metadata Index Tuning	24 hours	Yes
Register Repositories	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 (every 10 minutes), Normal (every 4 hours), and Low (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 53 and "Create a Policy" on page 56).

You can edit the Execute Policies tasks to change the default run frequencies as required. See "Edit a Scheduled Task" below.

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 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 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. In the upper-right corner of the task, click the menu button.
3. Click **Edit**.
4. Edit the fields as required. You can edit the following fields.

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 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.

5. Click **Save**.

Remove a Scheduled Task

When you no longer require a scheduled task, you can remove it from Storage Optimizer.

To remove a scheduled task

1. On the Administration dashboard, click **Scheduled Tasks**.
   The Scheduled Tasks page opens.
2. In the upper-right corner of the task, click the menu button, and then click **Delete**.
   A confirmation dialog box opens.
3. Click **Delete**.

Run Scheduled Tasks

Enabled scheduled tasks run whenever a Storage Optimizer Scheduler is active. You start Storage Optimizer Schedulers from the Service Control Manager.

You can install one 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 Storage Optimizer systems, you must deploy multiple Schedulers.

Related Topics

- "Change the Number of Scheduler Threads"
- "Install Multiple Storage Optimizer 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. In the upper-right corner of the task, click the menu button, and then click **Run Now**.

   The Storage Optimizer Scheduler runs the task when it next checks for tasks to run (by default, every 60 seconds).

Configure Storage Optimizer Schedules for Large Systems

The following section describes Storage Optimizer configurations to use in large 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 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 8. 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**.
3. Under **Engine Settings**, change the number of threads to use, and then click **Deploy**.

   Storage Optimizer redeploys.

Install Multiple Storage Optimizer Schedulers

For high processing volumes, you can install multiple Storage Optimizer Schedulers on several machines. You must modify the configuration of each Scheduler to point to the Storage Optimizer SQL Server database.

Chapter 9: Manage Policy Conflicts

The Conflict Management page, which is accessible through the Administration dashboard, displays policy execution conflicts that 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 Conflict

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

Storage Optimizer automatically reports policy conflicts when it encounters them. When 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 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 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 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 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, Storage Optimizer allows it to execute.
- If it has a lower priority than the others, 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**.

   The Settings page opens.
2. On the **General** tab, click **Details**.

   The page opens by default.
3. Under **Details**, change the **Autoresolve Conflicts** option to **Yes**, and then 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.

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 10: Issue Management

This section describes the Issue Management administrative function.

Manage Issues

The Issue Management page, which is accessible through the Administration dashboard, displays events of interest to 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 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 11: Health Check

You can use the Health Check page to verify key configuration settings in your Storage Optimizer deployment.

- "Check Storage Optimizer Health"
- "Run Advanced Health Check Reports"

Check Storage Optimizer Health

The Health Check page allows you to check the status of components and tasks in your Storage Optimizer system.

When you run the health check, the current health of the system appears. You can view any warnings or errors, and correct them in your system. For example, the health check reports whether Storage Optimizer can contact the IDOL connectors.

To check Storage Optimizer health

1. On the Administration dashboard, click **Health Check**.

   The Health Check page opens.
2. Click one of the refresh options.
   - **Refresh all** checks the health of all components
   - **Refresh group** checks the health of the specific group, Connectors or Storage Optimizer
   - Individual components or checks, such as **Storage Optimizer Fields Check**.

   The health check runs, and you can view the results of each test.

Run Advanced Health Check Reports

In addition to the basic Storage Optimizer health check, you can use Storage Optimizer Assist to run advanced health check reports on connector status, execution activity, and policies. 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 Storage Optimizer Assist**.

   The Storage Optimizer Assist page opens.
3. Select the **Report** that you want to run.

Chapter 12: 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

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

   The Add Property dialog box opens.
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.

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" on page 21
- "Edit Repository Settings" on page 24
- "Create a Policy Template" on page 53
- "Create a Policy" on page 56
- "Edit a Policy" on page 59

Chapter 13: Export Statistics to Excel

You can use a 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.

To export statistics

1. Run the **ControlPointStatisticsUtility**.

   The Storage Optimizer Analysis window opens.
2. Specify the host name, and then click **OK**.

   The export window opens. The Analysis Tasks section lists all analyzed repositories on the host system.
3. (*Optional*) To reanalyze 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. Type a **Task Name**.
   c. Click **OK**.

      The Task is added to the list.
5. Select an analysis task.
6. In the Export Task area, 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.

Chapter 14: Storage Optimizer Connectors

This section provides information on the supported 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.

Storage Optimizer Exchange Web Service Connector

Summary

The Exchange WS connector can be used to analyse 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 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 Connector 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.

Adding 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

Storage Optimizer File System Connector

Summary

The file system connector can be used to analyse and execute policy on documents and files held in Windows file shares. Limited capability is available for documents and files held in Linux file systems or Novell Netware file systems.

Supported Capability

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

A file system location can be established as a target location for relevant policy types. The following policy types can utilise 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 Connector 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.

Last Access Dates

Note: Recording updates to last access dates is typically disabled in Windows Server 2008, Vista, Windows 7, and Windows 8 through the Windows registry for performance reasons. This can be changed using the `fsutil` utility.

To ensure last access time updates are recorded:

```
fsutil behavior set disablelastaccess 0
```

To turn off last access time updates:

```
fsutil behavior set disablelastaccess 1
```

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

Note: Refer to Windows documentation for your specific version of Windows before making changes to the last access date behaviour.

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 analysed.

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.

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

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
Secure Shortcut	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 Uri	Enter the root URI of the Hadoop file system to connect to when securing 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 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.

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

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 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.

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

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

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

Secure Leave	Yes
Secure Remove	Yes
Secure Shortcut	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`

Also, see Setting up a Document Store for Declare in Place Policies if you need to use a Declare in Place policy.

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.

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

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

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
Secure Shortcut	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:

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
```

You can verify successful deployment of the web service by accessing the following URL:

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

Post-install steps

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

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.

Also, 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`

Also, see Setting up a Document Store for Declare in Place Policies if you need to use a Declare in Place policy.

Setting	Section	Description
IndexActivityFeeds	TaskName or Default section	Set to false to disable analyzing of activity feeds.

Setting	Section	Description
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.

A sample LUA script is as follows:

```
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
```

Troubleshooting

Troubleshooting the web service is covered in a chapter of the SharePoint 2010 Connector Guide.

Adding 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.

Defining 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.

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.

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 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\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
```

Also, see Setting up a Document Store for Declare in Place Policies if you need to use a Declare in Place policy.

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 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.

Storage Optimizer SharePoint Remote Connector

Summary

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

Supported Capability

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

Dispose	Yes
Secure Leave	Yes
Secure Remove	Yes
Secure Shortcut	No

[1] Requires an additional web service to be installed for SharePoint 2010 or 2013. Not available for SharePoint Online.

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

Secure Leave	Yes[1]
Secure Remove	Yes[1]
Secure Shortcut	Yes[1]

[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
```

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 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.

Storage Optimizer StoreAll Connector

Summary

The StoreAll connector can be used to analyse and execute policy on documents and files held in a StoreAll repository.

Supported Capability

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

Dispose	Yes
Secure Leave	Yes
Secure Remove	Yes
Secure Shortcut	No

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

Secure Leave	Yes
Secure Remove	Yes
Secure Shortcut	Yes

DeployTool Configuration

When selecting a StoreAll 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.
StoreAll Credentials Username	Optional. The user name to be used for basic authentication.
StoreAll Credentials Password	Optional. The password to be used for basic authentication. The password can be encrypted.
StoreAll API URL	The URL of the StoreAll API to be used by default when securing documents to a target location of type StoreAll.

Configure StoreAll Connector

You can find information on configuration settings in the Connector Guide or by accessing the following URL:

```
http://<Connector host>:13100/a=help
```

Adding New Repository of type StoreAll

When adding a new repository of type StoreAll, you must provide the following parameters:

StoreAll URL	The URL of the StoreAll API to be used to access the StoreAll system where the repository is located.
Root Path	The location under which documents are to be registered for subsequent analysis and/or management.

Defining a Target Location of Type StoreAll

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

Connector Config Section	The configuration setting in the connector configuration file containing the details required for securing documents to the StoreAll target location. The default value for the section name is DefaultTargetLocationConfig. This section contains the details of the StoreAll API entered in Storage Optimizer DeployTool.
Path	A target path to be used for documents secured to this target location.

Appendix A - 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 Windows .NET 4.5, and it 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

Synopsis

`stub.exe -create –source <path> -archiveP01 <path> -archiveP02<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 the `-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` (mandatory): 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. For 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, execute:

   ```
   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, execute:

   ```
   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, execute:

   ```
   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, execute:

   ```
   stub.exe -rehydrate -source C:\src
   ```
5. To rehydrate any archived file in the source and its child directories, execute:

   ```
   stub.exe -rehydrate -source C:\src -r
   ```
6. To dump the reparse metadata in the stubbed file's reparse point, execute:

   ```
   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, execute:

   ```
   stub.exe -delete -source C:\Store1\MonthlyAssets.pdf
   ```

Appendix B - Setting up a Document Store for Declare in Place Policies

Declare in Place policies make a copy of an item and insert it into HPE Records Manager.

Next, a Hold is placed on the original item in the source repository, therefore only source repositories that support Holds should be used with this policy. For example, HPE Records Manager, TRIM, and Sharepoint.

For Declare in Place policies to be successfully executed, some additional configuration is required on HPE Records Manager. You need to define a Document Store in HPE Records Manager so as to point to the repository on which to place the Hold.

1. From File, click **New** and select the **Document Store** tab in HPE Records Manager. Then select **CFS Connector** and click **OK**:
2. Then enter a name for the Document Store, the AUTN_GROUP property (For example, SharePoint 2010, if you are using SharePoint 2010) and the configuration information which is the location and port of the Distributed Connector.

After this is setup the Declare in Place policy is executed successfully.

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 (Storage Optimizer 5.2)

Just 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!


Hewlett Packard
Enterprise


Java
COMPATIBLE