Hewlett Packard
Enterprise

HPE Storage Optimizer

Software Version: 5.4

Best Practices Guide

Document Release Date: November 2016

Software Release Date: November 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 software updates, go to https://softwaresupport.hpe.com/patches.

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

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

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

Support

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

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

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

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

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

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

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

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

Contents

Chapter 1: Introduction

Use this guide to assist in installing and configuring your HPE Storage Optimizer environment for the best performance.

Review the prerequisites

Ensure that your environment meets all hardware, software, and third-party component requirements as described in the *HPE Storage Optimizer Installation Guide* or *Support Matrix*.

Physical server recommendation

HPE recommends that you use physical servers for the HPE Storage Optimizer environment, especially for SQL Server.

Antivirus recommendations

For performance reasons, if you are running antivirus software on the HPE Storage Optimizer host machines, you must ensure that it does not monitor the Storage Optimizer directories and any fileshares that have been indexed.

Some advanced antivirus software can scan the network and might block some Storage Optimizer traffic, which can cause errors.

Where possible, exempt the Storage Optimizer and IDOL processes from this kind of network traffic analysis.

Sizing

For assistance on sizing your HPE Storage Optimizer environment, contact HPE Support.

Size and scale considerations for HPE IDOL Content

> **NOTE:**
> For more information on HPE IDOL Content, see the *HPE IDOL Explorer* and the *HPE IDOL Getting Started Guide*.

A major part of the design of an IDOL-based architecture depends on:

- how many documents you need to index into each instance of Content.
- how many Content servers must run on each physical server, or equivalently, how many documents in total you need to index on each physical server.

If you know the total number of documents, the total number of physical servers required follows easily from these two points. However, in many cases you do not know the final number of documents, but instead need to scale the systems over time.

The upper limit on the number of documents for a server is likely to be determined by performance requirements. Multiple smaller servers tend to perform better than a single large one, because they can work in parallel on queries.

The key question is how many documents a server can index while remaining small. This value is tied to the local HPE IDOL configuration, the type of data that you want to index, the type of queries that you will use, and the pattern of indexing and querying.

While you can derive very rough estimate figures by comparing with similar systems on similar hardware, the only reliable way to get useful statistics is to test your proposed system. That is, run an instance of HPE Storage Optimizer and HPE IDOL in the configuration you intend to use, on candidate hardware, and monitor performance while you index realistic data and send realistic queries.

When simulating load and assessing performance, you might also want to consider the following points:

- Do you need to query servers during indexing, or can you set up indexing to occur only during quiet times when query load is low (for example, overnight)?
- How many queries do you expect IDOL to handle simultaneously?

These statistics can give you a sensible maximum size for a single instance of Content, and also its likely system resource usage (footprint on disk, process memory size). After these values are known, you can determine the key values above.

Obviously, total disk usage by all Content servers on a machine cannot exceed the space available. Normal memory usage by all IDOL processes on a server ideally must fit in the machine's physical RAM, with enough free space remaining for the OS to effectively cache file system data.

Chapter 2: SQL Server guidelines

Dedicated HPE Storage Optimizer Database server

To gain the best performance, HPE recommends that you do not install any other HPE Storage Optimizer components on the SQL Server.

If you are deploying the HPE Storage Optimizer databases to a server hosting other HPE Storage Optimizer components, such as connectors, configure SQL Server to limit the resources it consumes

Large-scale HPE Storage Optimizer deployments

For large-scale deployments, use Enterprise Edition of SQL Server, as it enables you to split the tables on different files and optimizes the performance.

> **NOTE:**
> For assistance on sizing your HPE Storage Optimizer environment, contact HPE Support.

Configure Windows power options on the server

In Windows, the default power settings balance power efficiency and performance. For the SQL Server to have consistent, predictable, and high performance, set the power option to High Performance. This additional processing capacity comes with higher power utilization.

> **TIP:**
> The configuration can be used on new and existing installations of HPE Storage Optimizer.

To set the power options

1. Open **Control Panel** > **Power Option**.
2. Click **High Performance** and then click **OK**.

   The server power options are set.

Install and configure SQL Server

To install SQL Server, follow the installation instructions provided by Microsoft.

Install the latest SQL Server service packs

Install the latest SQL Server service packs for your edition of SQL Server.

Minimum SQL permissions

The user account that deploys and installs the ControlPoint databases must have the following permissions configured in SQL Server:

• **Dbcreator, public** — required to create the databases

• **SecurityAdmin** — required to create users in the databases

DBO permissions are the minimum SQL permissions that can be used after the initial deployment.

SQL Server performance guidelines

When you configure the SQL server during installation, use the following guidelines to prepare the server for the HPE Storage Optimizer databases.

Configure SQL Server memory options

Use the two server memory options, **min server memory** and **max server memory**, to reconfigure the amount of memory (in megabytes) that is managed by the SQL Server Memory Manager for a SQL Server process used by an instance of SQL Server.

The default setting for **min server memory** is 0, and the default setting for **max server memory** is 2147483647 MB. By default, SQL Server can change its memory requirements dynamically based on available system resources.

Set the maximum server memory, instead of using the default.

For example, reserve 1 GB of RAM for the OS, 1 GB for each 4 GB of RAM installed from 4 to 16 GB, and then 1 GB for every 8 GB RAM installed over 16 GB RAM.

For more information on setting the min server memory and max server memory, see your SQL Server documentation.

> **TIP:**
> The configuration can be used on new and existing installations of HPE Storage Optimizer.

Configure Max Degree of Parallelism (MAXDOP)

When an instance of SQL Server runs on a computer that has more than one microprocessor or CPU, it detects the best degree of parallelism, that is, the number of processors employed to run a single statement, for each parallel plan execution.

You can use the **max degree of parallelism** option to limit the number of processors to use in parallel plan execution. SQL Server considers parallel execution plans for queries, index data definition language (DDL) operations, and static and keyset-driven cursor population.

To determine the proper value for MAXDOP, use the Best Practice Analyzer (BPA) or following the rules available from Microsoft support, https://support.microsoft.com/en-us/kb/2806535.

Configure cost Threshold of Parallelism in SQL Server

Use the **cost threshold for parallelism** option to specify the threshold at which SQL Server creates and runs parallel plans for queries. SQL Server creates and runs a parallel plan for a query only when the estimated cost to run a serial plan for the same query is higher than the value set in cost threshold for parallelism. The cost refers to an estimated elapsed time in seconds required to run the serial plan on a specific hardware configuration. Only set **cost threshold for parallelism** on symmetric multiprocessors.

Longer queries usually benefit from parallel plans. The performance advantage negates the additional time required to initialize, synchronize, and terminate parallel plans. The cost threshold for parallelism option is actively used when a mix of short and longer queries is run. The short queries run serial plans, whereas the longer queries use parallel plans. The value of cost threshold for parallelism determines which queries are considered short, and they should therefore be run using serial plans.

In certain cases, a parallel plan may be chosen even though the query's cost plan is less than the current cost threshold for parallelism value. This can happen because the decision to use a parallel or serial plan is based on a cost estimate provided before the full optimization is complete.

By default, the **cost threshold for parallelism** is set to 5. For servers with multiple CPUs, this setting should be set to 50. This takes advantage of the CPU by allowing smaller queries to run in parallel and allowing larger queries to have more threads to run in parallel.

For more information, see your SQL Server documentation.

Configure indexes and statistics

The indexes should be reconfigured on a weekly basis. Any index which has 30% or more fragmentation should be rebuilt.

Statistics should be updated daily and have Auto Update statistics enabled. For highly active servers with constant updates and inserts it may be beneficial to update stats every hour.

Chapter 3: Best practices for HPE Storage Optimizer

This section describes the best practices for your HPE Storage Optimizer environment.

Databases

This section describes the best practices for your HPE Storage Optimizer databases.

Back up the ControlPoint and IDOL databases

Before you back up the ControlPoint and IDOL databases, prepare the environment by disabling scheduled tasks and stopping services. This ensures that the ControlPoint and IDOL database backups remain in sync.

To prepare the environment

1. Allow any executing policy phases to complete.
2. In the HPE Storage Optimizer Administration dashboard, disable the Assign Policies and Execute Policies tasks using the Scheduled Tasks, to prevent new policies from being assigned to documents.

   > **NOTE:**
   > Be sure to disable all of the tasks: Normal, Low and High priority.

3. Check the Distributed Connector queue by issuing the command:

   `http://`*`distributedconnectorhost`*`:`*`port`*`/a=queueinfo&queuename=fetch&queueaction=getstatus`

   If the Distributed Connector is working with HTTPS, check the queue by issuing the command:

   `https`*`://distributedconnectorhost`*`:`*`port`*`/a=queueinfo&queuename=fetch&queueaction=getstatus`

   The default port number is 7000.

   All actions should be `Finished`.
4. When all connector actions and executing policy phases have completed, stop the following services:
   a. HPE Storage Optimizer Engines
   b. Distributed Connector
   c. Individual connectors and Connector Framework Services.

   The services are stopped.
5. Back up the ControlPoint databases.
   - ControlPoint
   - ControlPoint Audit
   - ControlPointMetaStore

- ControlPointMetaStore Tags
- ControlPoint Document Tracking
- ReportServer. Available if your environment is configured for reports.
- ReportServerTempDB. Available if your environment is configured for reports

6. Back up the IDOL databases.

Connectors

This section describes the best practices for various HPE Storage Optimizer connector components.

Enforce connector security

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

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

To enforce security

1. Navigate to the following location:

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

2. Locate the `<appSettings>` section.
3. Edit the `"SecureMetaStoreContent"` value from `"false"` to `"true"`.

   Example

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

4. Save the file.

Policies

This section describes the best practices for HPE Storage Optimizer policies.

Policy execution

- Allow any executing policy phases to complete before changing the ControlPoint environment.

  Database changes, restores of the HPE IDOL content databases or software upgrades that occur while policy executions are running may leave the environment in an inconsistent state.
- When manually applying a category to a policy, select the child level category.

Monitor policy conflicts

The number of policy conflicts in HPE Storage Optimizer can affect the performance of the policy executions.

To optimize policy execution performance, HPE recommends that you keep the number of policy conflicts as low as possible.

> **NOTE:**
> For more information on policy conflicts and how to resolve them, see the *HPE Storage Optimizer Administration Guide* or Help Center.

To monitor the policy conflicts

1. In the HPE Storage Optimizer Administration dashboard, navigate to **Administration**, then click **Conflict Management.**

   The Conflict Management page opens, listing any policy conflicts.

   You can manually resolve the conflicts, or configure HPE Storage Optimizer to automatically attempt to resolve them.

Configuration changes and service restarts

For environmental changes to take effect immediately, you must stop and restart the following services.

Changed configuration area	Service Name
Distributed Connector	Storage Optimizer Distributed Connector
Edge Filesystem Connector	Storage Optimizer Edge Filesystem Connector Storage Optimizer Edge Filesystem Connector Framework Storage Optimizer Edge Archive
Edge Archive Service	**Windows:** Storage Optimizer Edge Archive Service **Linux:** `hploggedfs` process
Filesystem Connector	Storage Optimizer Filesystem Connector Storage Optimizer Filesystem Connector Framework
IDOL	Storage Optimizer IDOL

Scheduled tasks and schedules

HPE Storage Optimizer includes a number of scheduled tasks to automatically perform jobs that are required to manage policies, generate statistical information for monitoring purposes, and so on. You can control how often these automated tasks run through schedules.

For more information on configuring scheduled tasks and schedules, see the *HPE Storage Optimizer Administration Guide* or the Administration Help system.

Configure HPE Storage Optimizer schedules for large systems

The following section describes HPE Storage Optimizer schedule configurations to use in large HPE Storage Optimizer systems. Depending on your requirements and hardware, you can combine the solutions in this section as required.

Change the number of scheduler threads

Each HPE Storage Optimizer Scheduler runs a defined number of threads, each processing a batch of items every time it runs. The default number of threads is eight. The optimal number of threads depends on your requirements and the system processor.

To change the number of Scheduler threads

1. Open the **HPE Storage Optimizer Configuration Manager**.
2. Click **Engine**.

   The Engine Setting page opens.
3. Under **Engine Settings**, enter the number of threads in the **Enter the number of threads to use to process items** box.

   HPE recommends one thread per core.
4. Click **Deploy**.

   HPE Storage Optimizer redeploys.

Install multiple HPE Storage Optimizer schedulers

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

Global Settings considerations

You can configure the system by using settings stored centrally in the **ControlPoint.dbo.CPGlobalSettings** table in the ControlPoint database.

Consider the following when changing settings in the Global Settings table in the ControlPoint database:

- In SQL Server, back up the ControlPoint database before attempting to make any changes to the **CPGlobalSettings** table.
- Empty values for items in the Global Setting table do not imply that the setting is zero (0).

  HPE recommends that you retain the default value of a global setting instead of setting a value to zero.
- Care should be taken when changing values in the Global Settings table.

Erroneous values in the Global Settings table can lead to lower performance or blocked execution progress in your HPE Storage Optimizer environment.

Selected Global Settings table values

SettingName	Description
Autonomy.ControlPoint.DataAnalysis.TagBatchSize	Used by the Data Analysis service to perform tagging of documents, for example, tagging documents for duplicates. This value specifies the batch size of the tagging in Data Analysis service. This value must be a non-zero value. **Default:** 10000
Autonomy.ControlPoint.IdolDocumentProcessing BatchSizeWithResults	This setting affects many areas of HPE Storage Optimizer, such as how many documents are discovered in a single discoverer action. It also affects the Storage Optimizer Engine queries IDOL for documents. When this value is set too high, it can cause HPE IDOL to become unresponsive. **Default:** 5000 **NOTE:** Setting this value to a larger value may accelerate policies, but it is not recommended.
Autonomy.ControlPoint.ScheduleLockAgeMinutes	Some policy phases can place a lock on a schedule in the **CPScheduleLock** table. For example, the conflict resolution resolver places a lock on the policy execution schedule when the conflict resolution resolver is busy.

SettingName	Description
	This value specifies the timeout period of the lock. **Default:** 30 **NOTE:** HPE recommends that this value not be set high, as items take a long time to expire.

Appendix A: Export statistics

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.

Before you begin

Install Microsoft Excel to the HPE Storage Optimizer server.

To export statistics

1. Run the Statistics Export Utility, which is available at the following location:

   ```
   Storage Optimizer x64\HPE Storage Optimizer Utilities\Statistics Export
   Utility\ControlPointStatisticsUtility.exe
   ```

   The HPE Storage Optimizer Analysis window opens.
2. Enter the host name in the **Host** box, and then click **OK**.

   The export dialog box appears. The Analysis Tasks section lists all analyzed repositories on the host system.
3. (Optional) To re-analyze a repository, select it, and then click **Re-analyze**.
4. (Optional) **To add a custom analysis task**
   a. Click **New**.

      The New Custom Analysis Task dialog box opens.
   b. Enter a **Task Name**.
   c. Click **OK**.

      The task is added to the list.
5. Select an analysis task.
6. In the **Export Task** section, select a Microsoft Excel template from the list, and then click **Export**.

   The data exports to Excel and appears according to the selected template. Potential Obsolete and Trivial disk space appears in the Obsolete-AllPotential and Trivial-AllPotential charts.

Appendix B: AppSettings in ControlPointTimer.config

The following is a reference for the `<AppSettings>` in `ControlPointTimer.config` file.

Settings	Usage
NumberOfTimerThreads	Number of Threads for the timer engine
ExceptionWaitTime	If 5 exceptions have been thrown in a row wait the amount of time indicated
ClientSettingsProvider.ServiceUri	
SleepSeconds	Thread sleep seconds for ingestion
MaxExecutionFrequencySeconds	Used in Phase execution
CallbackProcessor.MaxInstancesRunning	Used in collect cleanup
SqlDiscoverer.MaxInstancesRunning	Not used anywhere in the code
UndoSqlDiscoverer.MaxInstancesRunning	Not used anywhere in the code
CacheExpirationSettingsCSV	CSV for long expiry seconds, short expiry seconds used during ControlpointFrameworkRegistration andPolicyExecutionRegistration
LoadBalancingSettingsCSV	CSV for maxLatestNoWorkCount, maxPhaseIgnoreSeconds, slidingIgnoreSecondsIncrease

The following parameters are needed to enable secure connections with IDOL and Connectors

SecurePorts	Boolean value, used to determine if the specified metastore port must be added to the metastore port list
MetaStorePort	Port number
LDAPServer	
LDAPBaseObject	
LDAPUseSSL	Boolean to use SSL
LDAPMaxResults	Maximum number of results to retrieve
XMLGroupMembershipFile	Filename containing group information

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 Best Practices Guide (HPE Storage Optimizer 5.4)

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!