Hewlett Packard Enterprise

Technical white paper

HPE Data Protector Integration with Autonomy IDOL Server

Introducing e-discovery for HPE Data Protector environments

Table of contents

Summary

This technical white paper describes the integration of HPE Data Protector (**Data Protector**) with Autonomy IDOL Server (**IDOL Server**), which introduces e-discovery for enterprise environments where Data Protector is used as the data protection application. The white paper includes information which, together with information in the Data Protector end-user documentation and the IDOL Server end-user documentation, guides you through all usage aspects of the Data Protector IDOL Server integration (**integration**): installation and configuration, indexing Data Protector backup images, restoring on a basis of full content search, and troubleshooting. Where needed, cross-references point to Data Protector and IDOL Server documentation items. The white paper ends with a glossary, which explains some of the frequently used terms, and references to relevant resources on the web.

Introduction

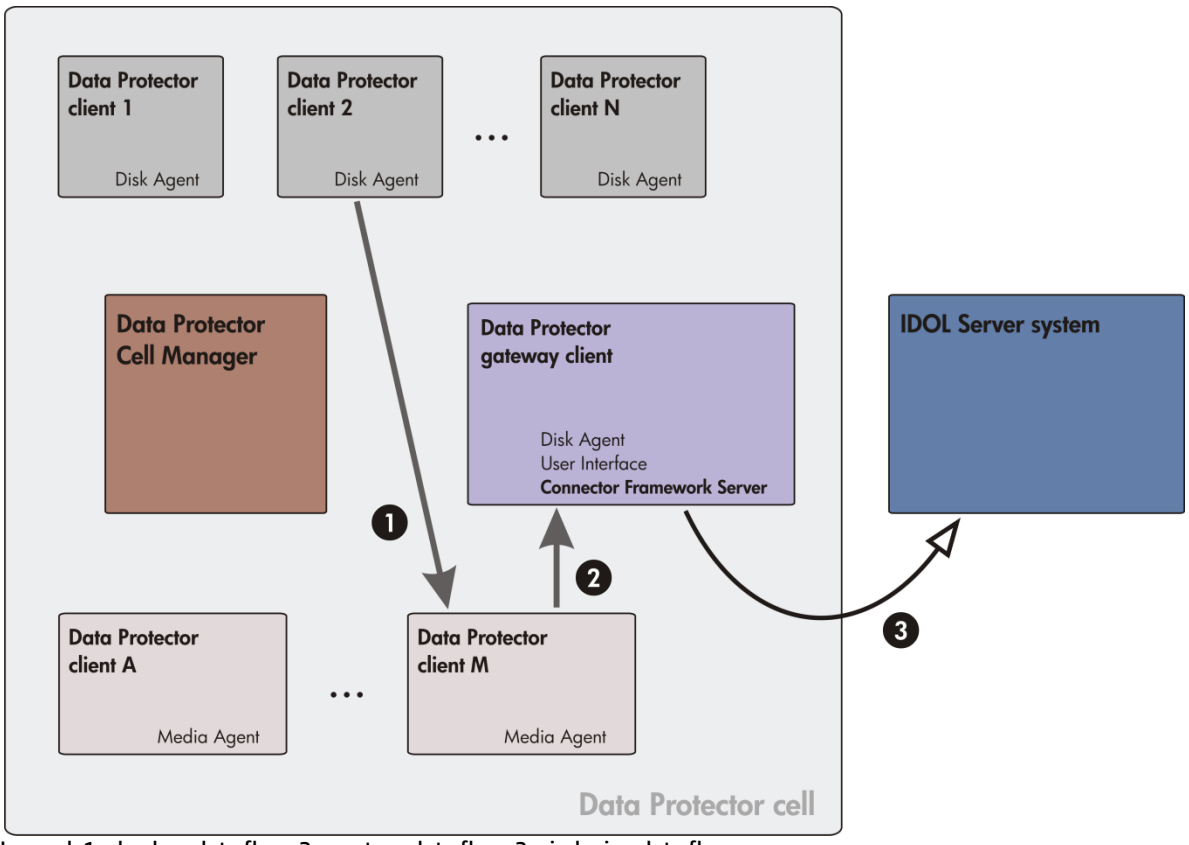
The selective restore capabilities of Data Protector have been limited to the restore-by-query functionality available in its graphical user interface. The Data Protector IDOL Server integration together with the IDOL Server overcomes this limitation by introducing advanced functionality for performing restore based on full content search. Interactively selected Data Protector backup data is exposed to the automatic indexing process driven by the IDOL Server (**on-demand indexing**). Acquired metadata tags stored in the IDOL Server can later be used in the legal hold use cases and for e-discovery and other purposes where full-content keyword search is required. To avoid potential exposures of sensitive data, the integration enables customers complete control over what is indexed.

Integration concepts

The indexing process is started from the Data Protector GUI as a post-restore operation, driven by special pre-exec and post-exec scripts bundled with Data Protector. Selected data from the selected Data Protector backup image is restored to a **temporary restore location** on the Data Protector **gateway client**, where the Autonomy Connector Framework Server processes it and sends content extracts and metadata of the files to the IDOL Server where they are indexed. The IDOL Server can be hosted on a system in the Data Protector cell, including the gateway client itself, or a system outside the cell.

When retrieval is required, the IDOL Server is used to find the files matching the keyword terms. Once the Data Protector backup session IDs are determined from their metadata, the desired files can be restored using ordinary Data Protector restore functionality.

Figure 1: Example of the infrastructure hosting the Data Protector IDOL Server integration



Legend: 1 – backup data flow, 2 – restore data flow, 3 – indexing data flow.

Limitations and considerations

Limitations

The integration has the following limitations:

- The integration functionality cannot be used from the Data Protector command line interface.
- The only supported Data Protector backup object type is *filesystem*.
  Only backup data created in Data Protector filesystem backup sessions can be indexed.
- The only supported backup type is *full backup*.
- Cross-platform restore is not supported.
  The selected backup object version must be restored to a system which uses the same operating system as the system from which the backup data originates.
- The backup object version selected for restore applies to all filesystem objects which compose the backup image. Any version selections for individual filesystem objects in the backup image are ignored by the integration.
- When indexing is enabled for the restore session, custom pre-exec and post-exec scripts cannot be used in the same session. The Pre-exec and Post-exec restore options are disabled in the Data Protector GUI.
- Simultaneous indexing of data that belong to multiple backup object versions is not supported.
- The integration does not provide functionality for mapping file system security attributes of the indexed objects to the security settings of the indexing metadata inside the IDOL Server.

Considerations

Consider the following when planning your integration-related activities:

- You can choose the desired backup object version in the Restore context of the Data Protector GUI in two ways:
  - By selecting a restore object.
    In this case, its *latest* backup object version will be restored and used for indexing.
  - By selecting a restore session

In this case, the backup object version *corresponding to the selected session* will be used.

For more information, see the *HPE Data Protector Help*.

- Decide which data can be indexed and which ones not. At the restore time, you can manually select which files will be indexed, regardless of which data compose the selected backup object version. You can make this selection in the Data Protector restore wizard.

Installing the integration

Operating system requirements

The integration supports the following operating systems on the gateway client:

- Windows Server 2008 R2 (64-bit) (only Full installation option)
- Windows Server 2003 – x86 (32-bit)

For operating system requirements for systems with other roles in the Data Protector cell, see the *HPE Data Protector Platform and Integration Support Matrix*.

Required applications

The following application software versions are required by the integration:

- HPE Data Protector 8.00
- HPE Data Protector 8.10
- Autonomy IDOL Server 7.6 and Autonomy Connector Framework Server 7.6

Prerequisites for installation

The following are prerequisites for installation of the integration:

- A Data Protector cell is correctly set up for backup and restore purposes, with the Data Protector Cell Manager, the Installation Server, and the Disk Agents installed on the appropriate systems.

  For more information, see the HPE Data Protector Concepts Guide, the *HPE Data Protector Installation and Licensing Guide*, and the *HPE Data Protector Help*.
- On the system which is to become the gateway client, the following is installed and configured appropriately:
  - Autonomy Connector Framework Server

    Connectivity between the Connector Framework Server and the IDOL Server must be established and verified.

  For installation and configuration instructions, see the *IDOL Getting Started Guide* and other IDOL documentation items.
- Optionally, on an arbitrary system in the Data Protector cell, the following is installed:
  - Data Protector User Interface

    This is required only in case you want to trigger the indexing-enabled restore sessions from a system which is not the gateway client.

Installation procedure

On the system which is to become the gateway client, install the following Data Protector components locally or remotely:

- Disk Agent

  This component provides the integration scripts.
- User Interface

  This component provides Data Protector commands that are invoked by the integration scripts.

For installation instructions, see the HPE Data Protector Installation and Licensing Guide and the *HPE Data Protector Help*.

Configuring the integration

Before you can start indexing your backup data, you need to configure the integration. Perform the steps below.

Step 1: Inspect the predefined parameters in the `omniidol.bat` script, and adjust them if necessary. The script resides in the directory *`Data_Protector_home`*`\bin` on the gateway client. The following parameters are available; their values are defined at the beginning of the script:

- `RESTORE_DIR_PATH`: Defines the path to the temporary restore location.

  Default: `C:\tmp\DataProtectorIndexing`
- `LOG_LEVEL`: Defines the logging verbosity level for status and event messages reported by the integration. A higher level "includes" all levels. The following levels are available:

  `1` – logging of errors (ERROR)

  `2` – logging of warnings (WARN)

  `3` – logging of informational messages (INFO)

  `4` – logging of debugging messages (DEBUG)

  `5` – logging of tracing messages (TRACE)

  Default: `3`
- `CFS_ADDRESS`: Defines the network interface address used by the Autonomy Connector Framework Server.

  Default: `127.0.0.1`
- `CFS_PORT`: Defines the port number used by the Autonomy Connector Framework Server.

  Default: `12800`
- `CLEANUP_FLAG`: If enabled (uncommented), this parameter causes all data at the temporary restore location to be removed after the indexing is complete. To retain indexed data, leave it disabled.

  Default: disabled
- `IDOL_DATABASE`: This parameter defines the name of the IDOL Server database where index of the backup object version will be stored. For specific Connector Framework Server configurations, which automatically add the IDOL Server database name to the indexed files, this parameter must be left empty. For more information, see *the IDOL Server Administration Guide*.

  Default: empty

Step 2: Ensure that the directory corresponding to the temporary restore location exists and is empty.

IMPORTANT: Having a non-empty temporary restore location when the restore session starts causes redundant data to be indexed later on, and eventually results in misleading information about which data are actually present in the backup image.

Step 3: Provision enough free space at the temporary restore location. The minimum required amount depends on the size of the chosen backup object version.

Step 4: Ensure that a corresponding Data Protector user exists for the user account which will be starting indexing-enabled restore sessions. This user must be a member of a Data Protector user group having the following Data Protector user rights: Monitor, Abort, Mount request, Start restore, Restore to other clients, See private objects. For more information on Data Protector user and user group management, see the *HPE Data Protector Help*.

Indexing Data Protector backup data

You can include data in the indexing process by invoking a restore session for the chosen backup object version, after selecting the restore option which enables indexing. Before starting indexing, consider topics discussed in Limitations and considerations.

Assumption and a prerequisite at this point are that the backup object version with the data to be indexed already exists, and that indexing preferences that suit your needs are already set in the IDOL Server.

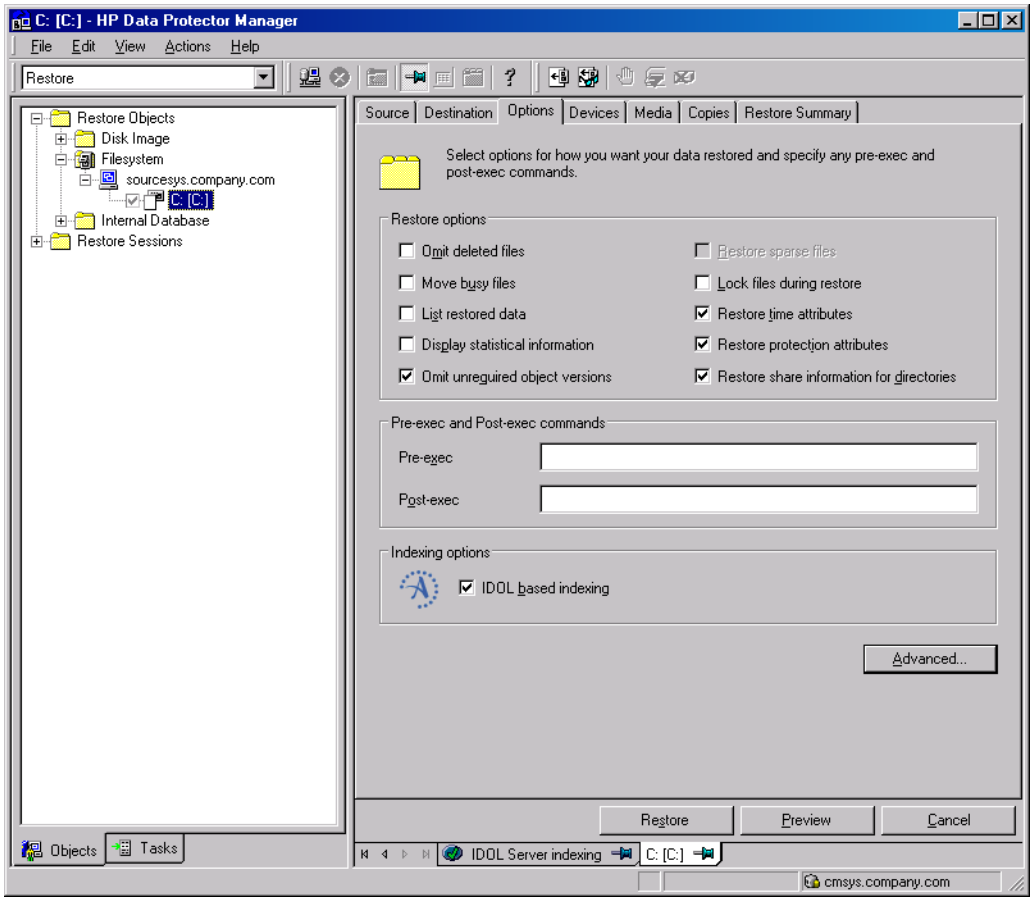
Follow the steps:

1. Launch the Data Protector graphical user interface.
2. In the Context List, click **Restore**.
3. In the Scoping Pane, select the desired backup object version in either way:
   - Expand **Restore Objects**, **Filesystem**, and finally the fully qualified domain name (FQDN) of the Data Protector client from which the data was backed up. The latest backup object version is used in this case.

- Expand Restore Sessions, expand the chosen backup object version, and finally expand the FQDN of the Data Protector client from which the data was backed up.

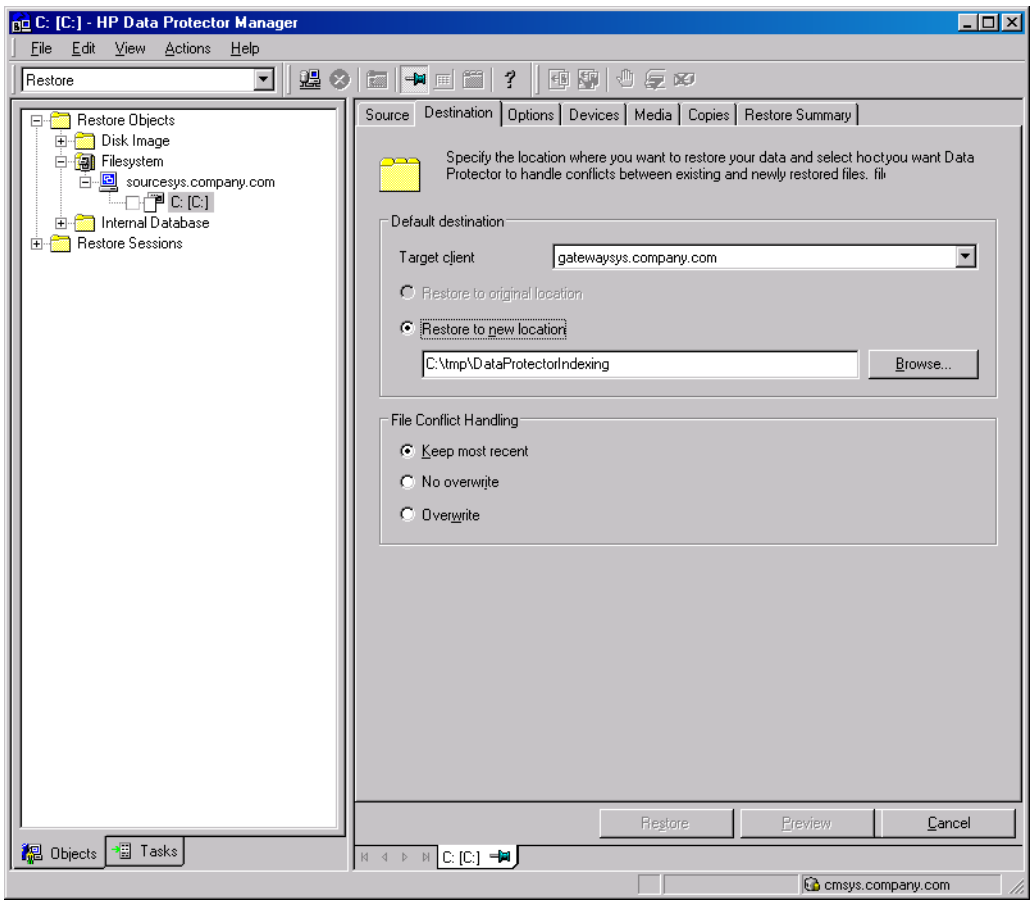
4. Click the mount point that corresponds to the chosen backup object.
5. In the Results Area, in the Source pane, select folders and files to define what should be restored and indexed.
6. Click the **Options** tab.

Figure 2: Selecting the option that enables indexing



7. In the Options pane, under Indexing options, select **IDOL based indexing**.
8. Click the **Destination** tab.

Figure 3: Specifying the gateway client and the temporary restore location



9. In the Destination pane, from the Target client drop-down menu, select the FQDN of the gateway client.
10. In the Restore to new location text box, update the default path to the temporary restore location so that it matches the actual path you configured as part of the integration configuration.
11. Click the **Restore Summary** tab. Review and optionally change the list of objects selected for restore and indexing.
12. Optionally, to preview the restore process, click **Preview**.

    If the default session preview behavior is not reconfigured with the related Data Protector global options, the restore preview will not trigger the integration scripts and will skip indexing.
13. Click **Restore**. The Start Restore Session dialog box appears.
14. In the Start Restore Session dialog box, click **Finish** to start the restore session and the subsequent indexing process.

Restoring indexed data

To successfully retrieve indexed files from a specific backup object version, the corresponding Data Protector backup image must still be available for restore. The retrieval process itself consists of three phases.

Perform them in the following sequence:

Phase 1: In the IDOL Server user interface, invoke a keyword search, and choose the items of interest (files) from the search results.

Phase 2: Using the IDOL Server user interface, determine the associated Data Protector backup session IDs for the chosen files from their indexing metadata. The metatada tags are created by the indexing process.

Phase 3: Using either Data Protector user interface (GUI or CLI), restore data from the corresponding backup object versions.

TIP: If restore of only a few files or even a single file is required, you can use the restore-by-query functionality of the Data Protector GUI to reduce the restore scope and speed up the process.

For instructions on how to perform a Data Protector restore, see the following:

- *HPE Data Protector Help*, for restoring using the GUI
- *HPE Data Protector Command Line Interface Reference*, for restoring using the CLI

Troubleshooting

This section provides problem-solving information when using the Data Protector IDOL Server integration.

For general Data Protector troubleshooting information, such as log and event reporting, warnings, and diagnostics, see the *HPE Data Protector Troubleshooting Guide* or the *HPE Data Protector Help.*

For Connector Framework Server troubleshooting information, see the *IDOL Server Administration Guide.*

Before you begin

Before you start determining the root cause of your problem:

- Ensure that the latest official patch bundles or patches for your Data Protector version are installed. On how to verify this, see the *HPE Data Protector Help* index: "patches".
- Get familiar with general Data Protector limitations as well as recognized issues and workarounds. For more information, see the *HPE Data Protector Product Announcements, Software Notes, and References.*

Problems and workarounds

Problem

An error about the non-existing temporary restore location is reported in the Data Protector session output
There are two possible causes for this problem: either the temporary restore location does not exist, or a wrong directory is specified for the Data Protector restore option Restore to new location.

Action

Perform the following:

1. Check if the specified directory exists, and create it if needed.
2. Ensure that the path of this directory is used as both the value of the `RESTORE_DIR_PATH` parameter in the `omniidol.bat` script (see Configuring the integration) and the value of the restore option Restore to new location in the Data Protector GUI.
3. Restart the restore session.

Problem

An inter-process communication error related to the Connector Framework Server is reported in the Data Protector session output
The root cause of this problem is the inability of Data Protector to establish a connection with the Connector Framework Server on the gateway client.

Action

Perform the following:

1. Check if the values of the parameters `CFS_ADDRESS` and `CFS_PORT` configured in the `omniidol.bat` script match your effective Connector Framework Server configuration, and adjust them if necessary. See Configuring the integration.
2. Verify that the Connector Framework Server is running and that it is listening on the port specified in the `omniidol.bat` script.
3. Restart the restore session.

Glossary

The following table explains some of the terms and acronyms used in this document.

Term or acronym	Description
backup object version	A Data Protector backup image created during a Data Protector backup session. The same backup specification is used to produce several different backup object versions at different times.
Data_Protector_home	A reference to the directory containing Data Protector program files. Its default path is *%ProgramFiles%*\OmniBack, but the path can be changed in the Data Protector Setup Wizard at installation time.
gateway client	A system in the Data Protector cell on which the temporary restore location resides. The Data Protector Disk Agent and the Connector Framework Server are installed on it.
IDOL	Intelligent Data Operating Layer
indexing	A process in which the IDOL Server indexes the chosen Data Protector backup data.
indexing-enabled restore session	A Data Protector restore session for which the post-restore indexing process is enabled interactively.
integration scripts	The scripts that enable indexing of the backed up and subsequently restored data.
temporary restore location	A directory where the Data Protector Disk Agent restores backup data to and from where the Connector Framework Server acquires data to be indexed by the IDOL Server. It resides on the gateway client.

For more information

Visit the following HPE Data Protector online resources to get more information:

http://www.hpe.com/software/dataprotector

www.hp.com/go/imhub

www.hp.com/go/software

Call to action

To read more about HPE Data Protector, visit http://www.hpe.com/software/dataprotector.

To read more about Autonomy IDOL Server and Autonomy IDOL, visit www.autonomy.com/content/Products/products-idol-server/index.en.html and idol.autonomy.com.

Get connected

hp.com/go/getconnected

Current HPE driver, support, and security alerts
delivered directly to your desktop

4AA4-0284ENA, Created June 2014

Hewlett Packard Enterprise