Informatica® Data Archive
6.5.1

Administrator Guide

Informatica Data Archive Administrator Guide
6.5.1
November 2021

Publication Date: 2022-02-10

Table of Contents

Preface

Read the *Data Archive Administrator Guide* to learn how to configure and set up Data Archive. Learn about system configuration, database users and privileges, and how to set up source and target connections.

Informatica Resources

Informatica provides you with a range of product resources through the Informatica Network and other online portals. Use the resources to get the most from your Informatica products and solutions and to learn from other Informatica users and subject matter experts.

Informatica Network

The Informatica Network is the gateway to many resources, including the Informatica Knowledge Base and Informatica Global Customer Support. To enter the Informatica Network, visit https://network.informatica.com.

As an Informatica Network member, you have the following options:

- Search the Knowledge Base for product resources.
- View product availability information.
- Create and review your support cases.
- Find your local Informatica User Group Network and collaborate with your peers.

Informatica Knowledge Base

Use the Informatica Knowledge Base to find product resources such as how-to articles, best practices, video tutorials, and answers to frequently asked questions.

To search the Knowledge Base, visit https://search.informatica.com. If you have questions, comments, or ideas about the Knowledge Base, contact the Informatica Knowledge Base team at KB_Feedback@informatica.com.

Informatica Documentation

Use the Informatica Documentation Portal to explore an extensive library of documentation for current and recent product releases. To explore the Documentation Portal, visit https://docs.informatica.com.

If you have questions, comments, or ideas about the product documentation, contact the Informatica Documentation team at infa_documentation@informatica.com.

Informatica Product Availability Matrices

Product Availability Matrices (PAMs) indicate the versions of the operating systems, databases, and types of data sources and targets that a product release supports. You can browse the Informatica PAMs at https://network.informatica.com/community/informatica-network/product-availability-matrices.

Informatica Velocity

Informatica Velocity is a collection of tips and best practices developed by Informatica Professional Services and based on real-world experiences from hundreds of data management projects. Informatica Velocity represents the collective knowledge of Informatica consultants who work with organizations around the world to plan, develop, deploy, and maintain successful data management solutions.

You can find Informatica Velocity resources at http://velocity.informatica.com. If you have questions, comments, or ideas about Informatica Velocity, contact Informatica Professional Services at ips@informatica.com.

Informatica Marketplace

The Informatica Marketplace is a forum where you can find solutions that extend and enhance your Informatica implementations. Leverage any of the hundreds of solutions from Informatica developers and partners on the Marketplace to improve your productivity and speed up time to implementation on your projects. You can find the Informatica Marketplace at https://marketplace.informatica.com.

Informatica Global Customer Support

You can contact a Global Support Center by telephone or through the Informatica Network.

To find your local Informatica Global Customer Support telephone number, visit the Informatica website at the following link:
https://www.informatica.com/services-and-training/customer-success-services/contact-us.html.

To find online support resources on the Informatica Network, visit https://network.informatica.com and select the eSupport option.

CHAPTER 1

Starting Data Archive

This chapter includes the following topics:

Starting the ILM Application Server

The web application deploy directory contains scripts to start and stop the ILM application server. The script names depend on the operating system.

▶ To start and stop the ILM application server, run one of the following files or scripts:

- Windows. Run batch files startApplimation.bat and stopApplimation.bat.
- UNIX or Linux. Run shell scripts startApplimation.sh and stopApplimation.sh.

You can start or stop the ILM application server in the background. You can also abort the ILM application server if it stops responding. To check the status of the ILM process, run the following command:

```
ps -ef | grep informia
```

Configure Proxy Server Settings

If the ILM application server is running under a network behind a proxy server, you must configure the `startApplimation.bat` (on Microsoft Windows) or `startApplimation.sh` (on Linux or UNIX) file with the proxy server details. If the Enterprise Data Manager client is running behind a proxy server, configure the proxy details on the local machine.

If the ILM application server is installed under a network behind a proxy server, add the following Java Virtual Machine parameters to the `startApplimation.bat` or `startApplimation.sh` file:

- -Dhttp.proxyHost=*<proxyServerHostName/IP>*
- -Dhttp.proxyPort=*<ProxyServerPort>*

If the Enterprise Data Manager client is running behind a proxy server, configure the proxy details on the local machine.

Logging In to Data Archive

After you start the ILM application server, you can access Data Archive.

Use the following URL to access the user interface:

```
http://hostname:portnumber/WebApplicationName
```

CHAPTER 2

System Configuration

This chapter includes the following topics:

System Configuration Overview

Before you create a Data Archive project, you must configure properties in both the configuration properties file and the system profile. The properties that you configure depend on the system behavior and the features that you want to implement.

The configuration properties file is a text file you create after you install Data Archive. You can access the system profile through the Data Archive user interface.

If you use the Data Vault Service, you must also configure environment variables to set character encoding.

Conf.properties File

The configuration properties file is a text file that includes a list of parameters and associated values. Data Archive reads the configuration file to determine settings related to archive and partitioning activities, system-wide properties, and Data Archive functional settings.

The Data Archive installation directory contains a template of the configuration properties file called `conf.properties.default`. Copy this template, rename it `conf.properties`, and configure it based on your environment.

After you update `conf.properties`, restart the ILM application server for the changes to take effect.

General Properties

You must configure general properties for Data Archive before you run an archive project.

The `conf.properties` file includes the following general properties:

informia.api.authentication

You can enable API authentication to check if a password given in the API URL is valid. To enable or disable API authentication, configure this property.

Valid inputs are Y and N. Default is N.

Password authentication is available only for API's built with the authentication mechanism.

informia.appendHostAndPortInLog

Displays the host and the port number on every line of the `applimation.log` file. Enable this property only at the request of Informatica Global Customer Support.

informia.applicationHostname

IP address of the ILM application server.

The embedded web server binds to one IP address. If the ILM application server needs to be accessible to multiple IP addresses, enable the informia.bindTomcatToMultipleIPAddress property.

informia.applicationPort

Port on which the ILM application server starts. If you change the default port, verify that the port number is not used by another application.

Default is 8080.

informia.applicationSessionTimeout

Browser session timeout period in minutes.

Default is 20. If blank, default is 10.

informia.appViewGetAllowed

Enables you to use a URL to view the Data Discovery portal application view for entities.

Default is disabled.

informia.bcp.columnSeparator

The column separator to designate distinct columns in an SQL query when you extract table data to flat files. The column separator must be unique. To use the Data Vault Service, comment out this property.

For example, `%@%` acts as a column separator.

Use `^#^` as the column separator when you extract table data for SAP application retirement.

informia.bcp.maxLOBFilesInADir

The maximum number of large files, such as BLOB or CLOB files, to store in a folder in the BCP directory. For example, if you set the value to 10,000 and you have 30,000 BLOB files, Data Archive creates three folders. Each folder will contain 10,000 files. The size of each file is determined by the `informia.LoaderDatFileMaxSizeLimit` property.

Default is 100,000.

The BCP directory path is `<bcp_location>/<job_id>/<table_id>/<column_id>/A<number>` where:

- `table_id` is the `meta_table_id` in `am_meta_tables`

- `column_id` is the `meta_column_id` in `am_meta_columns`

informia.bcp.rowSeparator

Uses the specified value, which must be unique, as a row separator. A row separator designates distinct rows in an SQL query when you extract table data to flat files. To archive to the Data Vault, comment out this property.

For example, `%#%` is a row separator.

Use `@#@#` as the row separator when you extract table data for SAP application retirement.

informia.bindTomcatToMultipleIPAddress

Enables the embedded web server to bind to multiple IP addresses. Enter Y if the ILM application server needs to be accessible to multiple IP addresses. Enter N if you want Data Archive to use the first IP address from the network configuration.

Default is N.

informia.bulkExtractionMode

Increases the performance of data extraction to flat files. Disable if a query contains data that cannot be extracted, such as images. Disable only at the request of Informatica Global Customer Support.

informia.dataDiscoveryAppendJobIdPredicate

Appends the job ID to Data Discovery queries, so that query results do not include a table more than once. For example, if you run a retirement cycle twice, Data Discovery displays the tables from that cycle one time.

informia.dataDiscoveryMetadataQuery

The metadata connection for the Data Discovery portal.

Default is AMHOME.

informia.defaultArchiveAreaRetention

The default retention period for an archive area within the Data Vault Service.

Use one of the following values:

- -1. Indefinite retention.
- 0. Immediate expiration.

Default is -1.

informia.disableAddingExtractJobIdPredicate

Determines whether job IDs are considered when creating a query to display data on the **Browse Data** page.

Use one of the following values:

- N. Consider job IDs when creating a query to display data on the **Browse Data** page.
- Y. Ignore job IDs when creating a query to display data on the **Browse Data** page.

Default is Y.

informia.disableTrendGraph

Determines if Data Growth Analyzer calculates the trend graph. When you set to Y, the trend graph is disabled. Set to Y for extended date ranges, such as 2,000 years. If you do not set this value to Y for date ranges with lengthy time spans, then the trend graph appears as an empty image.

When you set to N, Data Growth Analyzer computes the trend graph.

Default is N.

informia.dropInterimTablesOnTerminate

Drops interim tables when you terminate an archive or retirement job. If you are in a development or debugging phase, you may want to select false to examine interim tables in the source staging schema. If you are in a production phase, you might want to select true to eliminate unnecessary tables in the source staging schema.

Default is False.

informia.enableDDMSecurity

Encrypts user information and adds the encrypted text to each Data Vault query sent to Dynamic Data Masking. If you want to mask sensitive information in Data Discovery search results or in data visualization reports, enable this property. You must have Informatica Dynamic Data Masking installed and configured to enable masking. When a user queries data in Data Vault, masking rules that you set up in Dynamic Data Masking apply based on the Data Vault access role assigned to the user. Dynamic Data Masking replaces sensitive information such as a credit card number with a randomly generated number or Xs in the query results.

Default is No.

informia.enable.enhanced.import

Enables the enhanced import option for importing metadata accelerators and entities in the Enterprise Data Manager. Valid inputs are Y and N.

Default is Y.

informia.enable.mining

Enables the "import metadata from database" option in the Enterprise Data Manager. Valid inputs are Y and N.

Default is Y.

informia.enable.traditional.import

Enables the traditional method of importing entities and custom product family versions. Valid inputs are Y and N.

Default is N.

informia.encryptionkey.command

If you use a third-party encryption key generator to encrypt Data Vault files, provide the command to run the encryption key generator. The command provided should return only the encryption key.

informia.encryption.AdminTimeout

Defines the admin timeout in seconds.

informia.encryption.DeleteOldDataFiles

Specifies whether or not to delete the original unencrypted data files during the encryption job. Valid property values are Y and N. Default value is N. For more information on this property, see the chapter "Scheduling Jobs" in the *Data Archive User Guide*.

informia.encryption.IDVDeBugMode

Specifies whether or not Data Vault runs the encryption job in debug mode. Valid property values are Y and N. Default value is N.

informia.encryption.RegistrationThreadCount

Defines the number of registration threads used to encrypt the Data Vault data. Default value is 5.

informia.encryption.SCTRegistrationType

Defines whether the SQL or admin registers the encrypted .SCT files in Data Vault. Valid property values are SQL and admin. Default is SQL.

informia.encryption.WorkerThreadCount

Defines the number of worker threads used to encrypt the Data Vault data. Default value is 10.

informia.faThreadCount

Determines the number of threads used in the Data Vault Loader job.

By default, the ILM engine calculates the value by dividing the number of CPU cores by two. For example, if the server has eight CPU cores, then the default value is 4.

informia.fasLoadAdditionalAttributes

Loads tables with row length greater than 100,000,000 bytes.

informia.fasLoadProcessingThreads

Specifies command line options for the Data Vault Loader job.

informia.fasMigrationErrorThresholdCount

Defines the maximum number of errors that the Data Vault Service migration process can generate before the migration process stops loading data into the Data Vault. Default is 4.

informia.idv.ssl.pemCertificatePath

Defines the path to the PEM certificate file that enables SSL communication with Data Vault.

informia.isKeyStorePassEncrypted

Determines if the password for the keystore file is encrypted. This property applies if you enable the ILM application server to run on HTTPS and must create a keystore file.

Default is False.

informia.isSapOnWindows

Determines if an additional row separator is required for BCP file generation when the retirement job uses an SAP function module to generate BCP files for data in transparent HR and STXL tables, ADK files, and attachments. Required only for SAP application retirement if the SAP application is hosted on Windows.

Use one of the following values:

- Y. Operating system that hosts the SAP application is Windows. The retirement job appends `\r` as a suffix to the BCP file row separator that is configured in the informia.bcp.rowSeparator property. The `\r` is required to confirm the end of rows on Windows.
- N. Operating system that hosts the SAP application is not Windows. The retirement job uses the value configured in the informia.bcp.rowSeparator property as the BCP file row separator.

Default is Y.

informia.julianLimiterYear

The year by which to limit Julian dates that do not include a century during Julian to Gregorian date conversion for Data Discovery Portal searches.

For example: informia.julianLimiterYear=1975

The 5-digit Julian date "15001" could be either 1915 or 2015. If you set the limiting year to 1975, then the dates "05001," "15001," and "65001" are converted as 2005, 2015, and 2065. The Julian dates "75001," "85001," and "95001" are converted as 1975, 1985, and 1995.

informia.keywordSearchIndexDir

Specifies the location of the search indexes used in Search Data Vault. You must enter the file path of the folder containing the search indexes. The folder must be in the same root directory as the Data Archive installation. The ILM Engine must have read and write privileges to the folder.

You must set this property to run the jobs that create and delete the search indexes. Both jobs use the file path entered in this property to find the search indexes.

Based on your operating system, use one of the following formats for the file path:

- On a Windows operating system, use a double back slash between folder names. For example,

  ```
  C:\\KeywordSearchIndex
  ```

- On all other operating systems, use a single forward slash between folder names. For example,

  ```
  /Data/KeywordSearchIndex
  ```

There is no default value.

informia.loadDataVisualization

Loads the Data Visualization option so that users can create reports and dashboards. Set this property to Y or N.

If you install Data Archive on multiple machines, you must set this property to Y on at least one machine to handle Data Visualization requests. Set this property to N on any machine that you do not want to handle the Data Visualization process.

Default is Y.

informia.LoaderDatFileMaxSizeLimit

Maximum size of each large object file such as BLOB or CLOB, created and stored in the BCP directory by the `informia.bcp.maxLOBFilesInADir` property.

Default is 1024 MB.

informia.maxActiveAMThread

Maximum active Java threads to run at one time. A common configuration is two times the number of cores available on the ILM application server. Increase the number of threads dependent on machine capability and the number of tables you want to extract data from. The more active threads you can run at once, the faster the archive process will complete.

Default is 10.

informia.maxActiveIndexThreads

Maximum number of Java threads to run at one time. Each thread creates a search index. You create search indexes to be able to search for records across applications in Data Vault.

Optional. Minimum is 1.

Default is one less than the number of cores on the machine hosting the ILM Engine.

informia.maxCachedStatements

Maximum number of cached statements stored in the ILM repository. When the Data Archive web application accesses the ILM repository, it caches the SQL statements to improve future query performance.

Default is 100.

informia.maxNoResultsIteration

Maximum number of milliseconds that the retirement job waits for a response from the SAP system. For every package of rows that the job processes, the job waits for a response from the SAP system.

If the SAP system does not provide a response within this time, the job fails.

SAP application retirement only.

Default is 5,000,000.

informia.miningCommitCount

Number of tables that the Import Metadata from Database job commits to the database at each commit interval. If the Enterprise Data Manager returns a memory error, reduce the property value.

Default is 10,000.

informia.MonitorQuickTxnRestoreQueue

Enables or disables monitoring requests submitted through the quick transaction restore API. Controls the behavior of the monitor thread used to poll the status of the restore jobs triggered by the external API.

Set to Y to enable monitoring requests. Set to N to disable monitoring requests.

Default is Y.

informia.NULLRepresentationInBcp

Stores the NULL character used by the Data Vault Service to load null data into table columns in the Data Vault. This property contains the value `\u007F`, which represents the null character in the Data Vault Service. If this property is not enabled and the source column contains null data, the Data Vault Service sets the column in the Data Vault to an empty string. To disable the property, use the pound character (#) character to comment out the property. Default is enabled.

informia.oracleClientPath

Web server location for the Oracle client.

informia.privacyEnableCardinalityHint

Enables or disables Oracle Cardinality Hint. Oracle Cardinality Hints can improve query performance.

informia.retainTrailingSpaces

For Data Vault Service installations on AIX, determines if you want to retain trailing spaces in archived data for VARCHAR fields when you query data from the Data Vault. This property applies only to archive areas that you create after you configure this property. This property does not apply to existing archive areas or archived data. If you do not retain trailing spaces, the Data Vault Service might return validation errors.

Use one of the following values:

- N. Data Vault Service drivers truncate the trailing spaces.
- Y. Data Vault Service drivers retain the trailing spaces.

Default is N.

informia.showQuartzLog

Displays quartz log messages inside the quartz log. The quartz log is a log of all the jobs you schedule in Data Archive. If set to Y, the log might be very large. Enable only at the request of Informatica Global Customer Support.

Use one of the following values:

- N. Log messages do not appear.
- Y. Log messages appear.

Default is N.

informia.sqlServerNVarcharMaxAndVarcharMaxToVarchar

Determines whether you want to convert the Microsoft SQL Server data types `NVARCHAR(MAX)` and `VARCHAR(MAX)` to the `VARCHAR(32000)` data type in the Data Vault. The default value N, converts these source data types to the `CLOB` data type.

Use one of the following values:

- N. Converts the Microsoft SQL Server data types `NVARCHAR(MAX)` and `VARCHAR(MAX)` to the `CLOB` data type.
- Y. Converts the Microsoft SQL Server data types `NVARCHAR(MAX)` and `VARCHAR(MAX)` to the `VARCHAR(32000)` data type.

If you plan to use the Application Retirement for Healthcare accelerator reports and your source database is Microsoft SQL Server, set this property to Y before you retire the healthcare data to the Data Vault.

Default is N.

informia.sqlServerVarBinaryAsVarchar

Converts the Microsoft SQL Server data type `VARBINARY` to the `VARCHAR` or `CLOB` data type in Data Vault.

In some SAP tables, the column with data type `VARBINARY` is a primary key. In this case, set this property to Y before running the archive job.

Default is N.

informia.sslEnabled

Enables the HTTP application connection to use SSL authentication.

Use one of the following values:

- N. Disables SSL authentication.
- Y. Enables SSL authentication.

Default is N.

informia.sslKeystoreFile

Location of the keystore file that is used to configure HTTPS. Required if you enable HTTPS support.

informia.sslKeystorePass

Plain text password for the keystore file. Required if you enable HTTPS support.

informia.techViewMaxDisplayRowCount

Maximum number of rows displayed for each child table in the Search Data Vault technical view. Each row in the child table is a record that references the record in the parent table. To view all the records in the child table, use a value of -1.

Default is 100.

Note: If you enter -1 or a high value and the child table has a large number of records, the technical view page might take a long time to display.

informia.TempDirectory

Temporary folder for export and import used with the above attribute.

informia.updateSCTCommitInterval

For the online mode of SCT file registration in Data Vault, this property defines the number of queries that are run before the File Archive Loader job commits the queries to bring the SCT files online. The job executes one query for every 999 SCT files.

If a table has less than 999 SCT files, the File Archive Loader job commits once for all of the files in the table.

Default is 10,000.

informia.useAppendHintForRestore

Determines whether restore cycles use an APPEND hint in the copy to destination insert as select statements. Setting it to true requires an exclusive lock on the ERP table. The true value might also cause an Oracle timeout error to occur.

Default is False.

informia.useColumnDescriptionOnBrowseDataPage

Determines whether to display the following column descriptions on the **Browse Data** page:

- Column descriptions for special and transparent tables.
- Column descriptions for an alias name in a materialized view query.

Use one of the following values:

- N. Do not display column descriptions for special and transparent tables. Do not display column descriptions for an alias name in a materialized view query.
- Y. Display column descriptions for special and transparent tables. Display column descriptions for an alias name in a materialized view query.

Default is N.

informia.useDbaViewsInSeamlessAccess

Enables you to use DBA_* or ALL_* views when you create the seamless access layer. Enter ALL_* if the source administrator user does not have DBA_* views access.

Default uses DBA_* views.

informia.useTimestampServerForJARSigning

Specifies whether or not to sign JAR files that use the time stamp server. Use one of the following values:

- N. Enter N if you do not want to sign JAR files that use the time stamp server.
- Y. Enter Y if you want to sign JAR files that use the time stamp server.

Default value is Y.

informia.validateKeystore

Validates if the keystore has a certificate issued to the host before starting the ILM application server.

Default is false.

informia.validateRowCountBeforeDelete

Determines if archive and restore jobs validate the row count before the job deletes data from the source. You can validate row count for archive jobs with the archive and purge cycle and for restore jobs with the database archive restore cycle.

Use one of the following values:

- N. Disables row count validation. The job deletes from the source without row count validation.
- Y. Enables row count validation. Before the job deletes data from the source, the job compares the delete count with the insert count of the copy to destination step. If the row count is different, the job errors out.

Default is N.

informia.welcomePage

Determines the Data Archive page that appears when the user logs in. To set the welcome page, enter the page-specific portion of the URL. The following list provides page-specific URLs to Data Archive pages:

- For the **Keyword Search** page, enter:

  ```
  /quickSearch.htm?action=getQuickSearchHome
  ```
- For the **Search Data Vault** page, enter:

  ```
  /fileArchiveSearch.htm?action=displayFileArchiveAreas
  ```
- For the **Data Visualization Reports and Dashboards** page, enter:

  ```
  /dataVisualization.htm?action=getReportsandDashboards
  ```

By default, the **Monitor Jobs** page is the welcome page.

validHosts

Determines the machines that have access to run JavaServer Page-based APIs for standalone jobs that you can run through external APIs. Enter the IP address or the host name of the machines that can access the JavaServer Pages from the external application. Use a comma to separate multiple values.

The following text is an example of the property:

```
validHosts=host1, 10.11.22.33, hyw172967.abc.com, dev.*
```

By default, no machines have access.

informia.XMLTYPE_Storage_Option

Specifies the storage type of the XMLTYPE columns that the ILM Engine creates in staging tables during the archive process.

For Oracle versions 11.2.0.3 and later, enter the value "CLOB."

informia.SQlSafeValidation

Enables or disables SQL safe validation. SQL safe validation prevents SQL injection on user input parameters. If enabled, blocks user input that can be prone to SQL attacks and logs the user out. The default value is Y for enabled.

IBM DB2 Native Utilities Properties

Before you use IBM DB2 native utilities for data movement, configure properties for Data Archive to use IBM DB2 native utilities. Optionally, you can configure additional command parameters to use in the import and export commands. Uncomment the properties to enable them.

General Properties

The `conf.properties` file includes the following properties for IBM DB2 native utilities:

informia.db2ClientPath

Location of the IBM DB2 Connect client root directory. Required to use any of the IBM DB2 native utilities for data movement. The IBM DB2 Connect installation must be on the same machine that hosts the ILM application server.

If you do not configure this property, Data Archive uses JDBC for data movement.

informia.db2HPUPath

Location of the IBM DB2 HPU utility root directory. Enter the full path to the main directory of the IBM DB2 HPU utility installation. Required to use the IBM DB2 HPU utility to export data. The IBM DB2 HPU utility installation must be on the same machine that hosts the ILM application server.

If you do not configure this property, Data Archive uses the export utility to export data.

informia.exportLobLocation

Location on the IBM DB2 database that temporarily stores LOB data. Required for the load client utility to process tables that have LOB datatypes. The location must be on the same machine that hosts the ILM application server.

If you do not configure this property, the archive job uses the import utility to process tables with LOB datatypes. The import performance might be impacted if the archive job uses the import utility.

informia.exportUtilName

IBM DB2 utility that Data Archive uses to export data. Required if you want to use the IBM DB2 HPU utility to export data.

If enabled, default value is `HPU`. Data Archive uses the IBM DB2 HPU utility to export data.

If disabled, Data Archive uses the export utility to export data.

Default is disabled. By default, if you configure the informia.db2ClientPath property, Data Archive uses the export utility unless you specify another utility in this property.

informia.hpu.columnSeparator

Optional. Character that identifies the column separator for the IBM DB2 HPU utility to export data.

If enabled, the archive job passes the value to the control file that the IBM DB2 HPU utility uses to export data. The job uses this value to create export files. Configure to one byte length. The IBM DB2 HPU utility does not support multibyte length separators.

If disabled, the IBM DB2 HPU utility default separator is used to export data.

informia.hpu.rowSeparator

Optional. Character that identifies the row separator for the IBM DB2 HPU utility to export data.

If enabled, the archive job passes the value to the control file that the IBM DB2 HPU utility uses to export data. The job uses this value to export files. Configure to one byte length. The IBM DB2 HPU utility does not support multi-byte length separators.

If disabled, the IBM DB2 HPU utility default separator is used to export data.

informia.importUtilName

IBM DB2 native utility that Data Archive uses to import data.

If enabled, default is IMPORT. Data Archive uses the import utility. For IBM DB2 on AS/400 and z/OS, you must enable the property to use the import utility.

If disabled, default is LOAD_CLIENT. Data Archive uses the load client utility. The load client utility is available for IBM DB2 on AIX, Linux, UNIX, or Windows.

By default, if you configure the informia.db2ClientPath property, Data Archive uses the load client utility unless you specify another utility in this property.

informia.TempDirectory

Temporary directory that stores the import and export files for the IBM DB2 native utilities. Each table has a separate import and export file. Data Archive deletes the corresponding import and export files after it processes the table. The location must be on the same machine that hosts the ILM application server.

At a minimum, the size of the temporary directory should equal the size of the tables that you run in parallel. For example, you need at least 10 GB temporary storage if your average table size is 1 GB and you use 10 threads for parallel processing.

For the IBM DB2 HPU utility, location that temporarily stores the control file that Data Archive generates when the job copies data to the destination. The IBM DB2 HPU utility uses parameters in the file for the export process. After the job completes, the job deletes the control file.

Import and Export Command Parameter Properties

You can add IBM DB2 command parameters to the script that the export or import utility uses to import or export data. By default, no additional command parameters are required.

Add command parameters to the property that corresponds to the location in which you want to insert the statement. You can use any command parameters that the import and export commands support in that location. Enter the property values in the supported IBM DB2 syntax for the import and export commands.

You can add the following optional properties to the `conf.properties` file:

informia.db2ImpStart

Statement that is inserted after the load file name of the import command. If you configure separators for the HPU utility, you must add a statement for the import utility to use the same separators.

The following value is an example of a statement that creates a dump file:

```
MODIFIED BY chardel@ coldel% delpriorityчar DUMPFILE=/u01/db2/Dump.dmp FASTPARSE
ANYORDER warningcount 10000
```

informia.db2ImpMiddle

Statement that is inserted before the insert statement of the import command.

informia.db2ImpLast

Statement that is inserted after the insert statement of the import command.

informia.db2ExpStart

Statement that is inserted after the unload file name of the export command.

informia.db2ExpLast

Statement that is inserted after the select statement of the export command.

LDAP User Authentication Properties

LDAP authenticates users in an LDAP server. Use LDAP authentication to enable a single sign-on feature across multiple applications. You must configure properties for Data Archive to use LDAP user authentication.

The `conf.properties` file includes the following LDAP properties:

authenticationMethod

Determines the type of user authentication.

If commented, default is native user authentication. You maintain users in Data Archive.

If uncommented, default is `LDAP`. You maintain users in the LDAP directory service and synchronize users to Data Archive.

ldap.attribute.email

Configures the LDAP email address, AM_USERS.EMAIL_ADDRESS.

Default is `mail` for Active Directory and Sun LDAP.

ldap.attribute.fullName

Configures the ILM full user name, AM_USERS.FULL_NAME.

Default is `displayName` for Active Directory.

Default is `uid` for Sun LDAP.

ldap.attribute.groupclassname

Name of the LDAP directory service object class that you use to group members or security groups.

Default is `group` for Active Directory.

Default is `groupOfUniqueNames` for Sun LDAP.

ldap.attribute.ismemberof

Name of the LDAP directory service attribute that indicates a user is a member of a group.

Default is `memberOf` for Active Directory.

Default is `isMemberOf` for Sun LDAP.

ldap.attribute.member

Configures the LDAP member name.

Default is `member` for Active Directory.

Default is `uniqueMember` for Sun LDAP.

ldap.attribute.organizationName

Configures the LDAP organization name, AM_USERS.ORGANIZATION_NAME. Optional.

No default value. If you do not set this property, the organization name of the user is set to `LDAP User`.

ldap.attribute.userName

Configures the ILM user name, AM_USERS.USER_NAME.

Default is `sAMAccountName` for Active Directory.

Default is `uid` for Sun LDAP.

ldap.pageSize

Restricts the number of entries returned in a single page.

- 0. Paging is disabled.
- Greater than 0. Paging is enabled. Value indicates the number of entries to return in a single page.

ldap.syncRoles

Determines the location where you maintain role assignments for users.

If commented, you maintain role assignments in the users account in Data Archive.

If uncommented, enables role assignment synchronization from the LDAP directory service to user accounts in Data Archive. You maintain role assignments for users in the LDAP directory service. When users log in to Data Archive, Data Archive synchronizes the role assignments from the LDAP directory service and updates the role assignments in the user account in Data Archive.

If you uncomment the property, enter one of the following values:

- False. Disables role assignment synchronization.
- True. Enables role assignment synchronization.

Default is false.

ldap.useSSL

Determines if you access the LDAP directory service through SSL.

Use one of the following values:

- False. Disables SSL authentication.
- True. Enables SSL authentication.

Default is false.

ldap.roleNamePrefix

Adds a user-defined custom prefix to the Data Archive-specific LDAP role names. If enabled, only roles with the prefix are considered for LDAP role synchronization and role search.

For example, to add the prefix "GLOBAL_US_ILM-" to the LDAP role name, configure the property as follows: `ldap.roleNamePrefix=GLOBAL_US_ILM-`

When you create a custom prefix for LDAP roles, you must create Data Archive-specific LDAP roles using the prefix. For example, "GLOBAL_US_ILM-Administrator" for the Administrator role or "GLOBAL_US_ILM-Legal_Hold_User" for the Legal Hold User role. Data Archive removes the prefix when it compares LDAP roles to Data Archive roles. For example, the LDAP user "GLOBAL_US_ILM-Legal_Hold_User" is assigned the Legal Hold User role in Data Archive.

You can specify multiple custom prefixes. Separate each prefix by a comma.

For example: `ldap.roleNamePrefix=GLOBAL_US_ILM-,GLOBAL_US_DSG-`

ldap.useRoleNamePrefixForNestedGroups

Determines if nested group searches consider only groups that begin with the role name prefix.

Use one of the following values:

- True. Nested group searches consider only groups that begin with the role name prefix.
- False. Nested group searches consider all role names.

Default is true.

ldap.skipReferrals

Restricts LDAP searches of roles assigned to a user to the user's domain component.

Use one of the following values:

- True. Restricts LDAP searches of roles assigned to a user to the user's domain component.
- False. Disables the restriction.

Default is false.

ldap.ignoreRefErrors

If set to true, Data Archive ignores referral errors in LDAP role searches.

Use one of the following values:

- True.
- False.

Default is true.

Single Sign-On Properties

To enable single sign-on support through an identity provider that uses the SAML standard, configure the single sign-on properties.

The `conf.properties` file includes the following single sign-on properties:

informia.sso.enable

Enables single sign-on. Enter Y to enable single sign-on. Enter N to disable it.

Default is N.

informia.idp.metedata.file

Path to the identity provider's metadata file.

Example: informia.idp.metedata.file = c:\\metadata

informia.key.path

Path to the KeyStore that you created during single sign-on configuration.

Example: c:\\generatedKeys

informia.key.alias.name

Alias name for the KeyStore that you created during single sign-on configuration.

Example: informia.key.alias.name = testkey01

informia.key.password

Password for the KeyStore. Before you update the property, this password must be encrypted using the encrypt password utility provided in Data Archive.

Run the command below to encrypt the password from the ILM directory:

-"encryptPassword.bat testkey01" for Microsoft Windows

-"encryptPassword.sh testkey01" for Unix

Example: informia.key.password = D1YgPl914QpCtSgoHWbsCg==

informia.idp.home.url

The identity provider's home URL. This parameter is not required for all identity providers (Okta, Onelogin).

Configuring the Conf.properties File

The installation includes a default configuration properties template file called `conf.properties.default`. Use the file as a template. Copy the default file and configure the properties in the copied file. If you configure the default file, your changes might be overwritten when you upgrade.

1. Navigate to the Web container folder of the installation directory.
2. Copy the `conf.properties.default` file and rename the file to `conf.properties`.
3. Save the file in the same directory.
4. Open the `conf.properties` file.
5. Define the properties.
6. Save the file.
7. Restart the ILM application server.

Sample Conf.properties File

You might configure the conf.properties file similar to the sample file below.

```
#Use this to update the port on which application starts (if unset, default is 8080)
informia.applicationPort=8080
#Use this to enable https for ILM application
#informia.sslEnabled=Y
#Property that specifies whether the keystore password is encrypted
#informia.isKeyStorePassEncrypted=true
#informia.sslKeystoreFile=path to keystore
#informia.sslKeystorePass=password for keystore
#informia.validateKeystore=false
#Set the session timeout value (Time in mins) (if unset, default is
10)informia.applicationSessionTimeout=20
#Set the maxActive java thread to run at a time (if unset, default is
10)informia.maxActiveAMThread=10
#Set the maxActive java thread to run at a time for a smart partitioning job (if unset,
default is 6)informia.maxActiveSPThread=6
#If you want the quartz log to get printed set the value to Y (if unset, default is N)
#informia.showQuartzLog=Y
#maximum numbers of am statements that are cached (if unset, default is 100)
#informia.maxCachedStatements=100#ORACLE EXPORT IMPORT ATTRIBUTES
#If you have oracle client installed on webserver, point it to the location
#informia.oracleClientPath=
#DB2 EXPORT IMPORT ATTRIBUTES
#Uncomment this parameter to use Export/Import Utilities for DB2. The value should be
pointing to the location of DB2 Connect Client,
#ex: \home\user\sqllib
#informia.db2ClientPath
#Extra Attribute for DB2 EXPORT IMPORT ATTRIBUTES for AS400 and Mainframe only
#informia.importUtilName=IMPORT
#Temp folder for export/import used with above attribute
#informia.TempDirectory
#Temp Folder on DB2 Server to be used by Export / Import Utilities for storing LOB values
#informia.exportLobLocation
#DATA DISCOVERY RELATED ATTRIBUTES
#if you want the xsl debugging enabled, set it to Y (if unset, default is N
#informia.xslDebugMode=Y
#If you want the storage lib to load (make sure the relevent storage files are in path)
(if unset, default is N)
#informia.loadStorageLib=Y
```

```
# Set this to Y and the server will not attampt to display main Applimation web site
page on the login screeninformia.intranetMode=N
# Enable this and set to Y  to enable row count validation before delete
#informia.validateRowCountBeforeDelete=N
#Archive to External Storage Email msg Attributes, default from_addr and to_addr will be
picked up if no entity or pfv specific entries are available.
#entity_-8_from_addr=From_Entity_-8
#entity_-8_to_addr=To_Entity_-8
#entity_-12_from_addr=From_Entity_-12
#entity_-12_to_addr=To_Entity_-12
#pfv_100_from_addr=From_Pfv_100
#pfv_100_to_addr=To_Pfv_100from_addr=ILMAdministratorto_addr=ILMAdministrator
# Data Discovery attributes
#Use this for delimiters in BCP files - should have different Characters
#informia.bcp.columnSeparator=%@%
#informia.bcp.rowSeparator=%#%#
#Use this for Archive to External Storage in Serial Mode
#informia.discovery.serialMode=Y
#Use this for Reconciler Standalone Program
#informia.discovery.issueDelete=Ninformia.dataDiscoveryMetadataQuery=AMHOME
#informia.dataDiscoveryIgnoreJobIdQuery=Y
#informia.maxDiscoveryRows=100
#informia.dataDiscoveryAppendJobIdPredicate=false
#informia.dataDiscoverySimpleQueryConstruct=false
#Data Archive Quick restore Attributes
#informia.MonitorQuickTxnRestoreQueue: Used to control the behavior of the monitor
thread used to Poll the Status of the Restore Jobs triggered by the external API.
#Values:
#Y: Monitor Process would poll the progress and the data can be accessed from jsp/
TxnRestoreApiUtil.jsp
#N: Monitor process would not poll the process
#Blank defaults to Y
#informia.MonitorQuickTxnRestoreQueue=N
# IP address for tomcat binding
# informia.applicationHostname=<IP address>
# where <IP address> is the IP address of the ILM server.
# To bind Tomcat to multiple IP Addresses of the Host
#informia.bindTomcatToMultipleIPAddress=N
# LDAP related properties
## Enable LDAP authentication
## authenticationMethod=LDAP
## Enable LDAP SSL authentication
## ldap.useSSL=true
# Enable LDAP role synchronization
# ldap.syncRoles=true
# Use LDAP Role name prefix
# Multiple comma separated prefixes are supported.
# If set, LDAP security groups must start with this prefix to be considered an ILM role.
# For example:
# ldap.roleNamePrefix=ILM-
# The LDAP group ILM-Legal_Hold_User would be treated as ILM's Legal_Hold_User role
# # ldap.roleNamePrefix=
# Only consider groups that start with the role name prefix in nested group search,
enabled by default
# ldap.useRoleNamePrefixForNestedGroups=true
# Ignore referral errors such as:
# javax.naming.PartialResultException: [LDAP: error code 10 - 0000202B: RefErr:
DSID-0310063C, ...
# ldap.ignoreRefErrors=true
# Restrict LDAP searches of roles assigned to a user to the user's domain component (DC)
# For example: The DC of CN=jdoe,OU=Resources,DC=example,DC=com
# would be DC=example,DC=com and LDAP groups other than DC=example,DC=com will be skipped
# ldap.skipReferrals=false# Nested role/group search, enabled by default
# ldap.enableNestedRoles=true
# The isMemberOf attribute of certain LDAP systems (e.g. OUD) shows both direct and
indirect group memberships from nested groups.
# This can lead to incorrect role assignments if role synchronization is enabled.
# Set the following attribute to true to filter out indirect group memberships.
# ldap.directGroupMembershipCheck=false
## LDAP attribute mapping properties
## Attribute that maps to the ILM user name (AM_USERS.USER_NAME)
```

```
# ldap.attribute.userName=uid
# The default is "uid" for Sun LDAP and "sAMAccountName" for Active Directory
## Attribute that maps to the ILM full user name (AM_USERS.FULL_NAME)
# ldap.attribute.fullName=cn
# The default is "uid" for Sun LDAP and "displayName" for Active Directory
## Attribute that maps to the email address (AM_USERS.EMAIL_ADDRESS)
# ldap.attribute.email=mail
# The default is "mail" for bot Sun LDAP and Active Directory
## Attribute that maps to the email address (AM_USERS.ORGANIZATION_NAME)
# ldap.attribute.organizationName
# No default value. If this property is not set then the user's Organization Name will
be set to "LDAP User".
## Group member attribute name
# ldap.attribute.member=member
# The default is "uniqueMember" for Sun LDAP and "member" for Active Directory
# The name of the attribute that represents the groups the user is member of.
# ldap.attribute.ismemberof=isMemberOf
# The default is "isMemberOf" for Sun LDAP and "memberOf" for Active Directory.
# The name of object class representing a group.
# ldap.attribute.groupclassname
# The default is "groupOfUniqueNames" for Sun LDAP and "group" for Active Directory.
# LDAP Paging support: This property controls the number of entries returned in a single
page. A value of 0 disables paging.
# ldap.pageSize=500
# This property controls whether to drop interim tables when terminating a job (default
= false)
# informia.dropInterimTablesOnTerminate=false
# Data Discovery properties
# Append applimation_job_id predicate to child table query where clauses
# defaults to false if metadata source = AMHOME, otherwise defaults to true
# informia.dataDiscoveryAppendJobIdPredicate=true
# Property to use "row level" security in amhome mode
(informia.dataDiscoveryMetadataQuery=AMHOME)
# informia.dataDiscoverySearchUsingNPASecurity=true
# The useAppendHintForRestore property controls whether restore cycles will use an
APPEND hint in CTD insert as select statements.
# The default setting is false and restore cycles will not use the hint.
# Note that using APPEND will require an exclusive lock on the ERP table and ORA-02049:
timeout: distributed transaction waiting for lock errors might occur.
# informia.useAppendHintForRestore=false
# Use DBA_* or ALL_* views when creating the seamless access layer (default is to use
DBA_* views)
# If set to true then access to the following DBA views is required:
# Seamless Access User: DBA_DEPENDENCIES, DBA_SYNONYMS, DBA_OBJECTS, DBA_POLICIES
# Source User: DBA_VIEWS, DBA_TAB_COLUMNS
#informia.useDbaViewsInSeamlessAccess=true
# Use this parameter to specify the default retention period for the archive folder.
# 0 Immediate expiry allowed.
# -1 Infinite retention applies.
# nd|D|m|M|y|Y  Positive integer with d, or m, or y suffix (case insensitive).
#For example, 500d means 500 days; 24m means 24 months; and 10y means 10 years.
#informia.defaultArchiveAreaRetention=-1
# Use this parameter in case of Legacy Platforms to query database catalog from Database
Drivers Provider's implementation of the API# Y  uses the Driver providers
implementation of JDBC API to query DB object information
# N  uses the universal approach of querying the DB object information (default)
# informia.useDriverProvidedMetadata = Y
#Use this parameter to control renaming of the Staging Directory of ArchiveTarget
#informia.attachmentRenameDir=Y
#Use this paramater to enable Orcale Cardinality Hint. This can be useful to improve the
performace.
#informia.privacyEnableCardinalityHint=Y
#Use this parameter to enable security in JSP based API to manage Jobs.
#you can provide comma separated host entries
#validHosts=host1, 192.168.168.97, dev.*
#Use this parameter to print the host and port in the logs
#informia.appendHostAndPortInLog=N
#Use this parameter to enable retaining of trailing spaces in FAS
#informia.retainTrailingSpaces=N
#Uncomment this property to turn off bulk extraction select statements from the engine
#informia.bulkExtractionMode=Ninformia.virtualView.colSeparatorConstant=&COL_SEPARATOR
```

```
#Uncomment this property to add addtional attributes to FAS load properties
#informia.fasLoadAdditionalAttributes=<entry key="maxSourceLineLength">512000000</entry>
#Use this parameter to limint the concatenated Column length in Bulk Extract mode.
Default is 2000 bytes
#informia.bulkModeMaxConcatenatedColumnLength=2000
#Uncomment this parameter to get the App View over URL using GET method
#informia.appViewGetAllowed=Y
#MIS Integration Properties
#DomainName=<Domain_Name>
#DomainHost=<PCHostName>
#DomainPort=<DISPort>
#nodeName=<PCNode_Name>
#UserName=<PowerCenter_User>
#Password=<PowerCenter_User_Password>
#version=9.1
#UserNameSpace=<PC_Authentication_Method>
#mrsServiceName=<Model_Repository_Name>
#DisName=<DataIntegrationServiceName>
##Set to UK if you want to use uniqueness runs and PK to do Primary Key Profiling runs
instead.
#informia.pkInferenceType=UK
#PC Integration
#PCRSName=<DataDomainServiceName>
#PCISName=<Power Center Integration Service Name>
#ide.thread.sleeptime=3000
#informia.pc.codepage=MS1252
#informia.pc.repdbtype=Oracle
#informia.pc.foldername=ArchiveInt
#informia.isPCOnWindows=Y
#informia.pc.sourceDatabaseType=Oracle
#informia.exposePCForExecution=N
# Legal hold query options
# Prevent legal hold child table queries to run in parallel (default is false)
# legalhold.useWithClauseForChildTableQuery=true
# Number of npa_unique_record_index values in the IN list of legal hold child table
queries (default is 100)
# legalhold.numUniqueIdsInWhereClause=100
# Use this parameter for meta data mining, insert commit interval
# informia.miningCommitCount=10000
# DB2 HPU ExportUtil Configuration
#informia.db2HPUPath=/opt/IBM/DB2TOOLS/HighPerformanceUnload42
#informia.exportUtilName=HPU
#informia.db2Extension=DEL
#These are added after the load file name is mentioned in import command.
#informia.db2ImpStart=
#These are  added before the insert statement mentioned in import command.
#informia.db2ImpMiddle=
#These are added after the Insert statement is mentioned in import command.
#informia.db2ImpLast=
#These are added after the unload file name is mentioned in export command.
#informia.db2ExpStart=
#These are added after the select statement is mentioned in export command.
#informia.db2ExpLast=
#If enabled then it is used only when we use HPU. If disabled then we do not use it in
any utility. Please note that HPU allows only one byte long separator.
#informia.hpu.columnSeparator=
#informia.hpu.rowSeparator=
# Create threading tables as Index Organized Tables (defaults to true):
# informia.createDFSThreadingTableAsIOT=true
# Create index on APPLIMATION_ROWNUM column if tables are not created as IOTs (defaults
to true):
# informia.createIndexOnDFSThreadingTable=true
#Use this parameter to enable DFS Pause & Resume Functionality
#informia.DFS.PauseResume=Y
# Set informia.dataDiscovery.deleteTransactionXML to false in order to keep the
transaction xml (App View) for debugging purposes
# informia.dataDiscovery.deleteTransactionXML=true
# CR 277873 : Setting this property to Y, disables the trend graph calculation for large
datasets in the DGA Administrator Dashboard
# informia.disableTrendGraph=Y
# Data Partition Column Name being added to Partitioned
```

```
tablesinformia.partition.ilm.column=ILM_PART_COLUMN
#if the partitionDbms is N then use native partition(no redefinition)
#informia.partitionDbmsRedef=N
# default value is - " -k 4 -j 1"
#informia.fasLoadProcessingThreads= -a 0x00 -k 4 -j 1 -x
#[in MB's] default value is 1024
#informia.maxBcpGroupSize=1024
# used in the case of mount points
#informia.sctPremissionCommand=chmod 664
#property defines task timout of FAS server(if unset, default is 3000(ms))
#informia.fasServerTaskTimeout=120000
# default value for informia.isSapOnWindows is 'N'informia.isSapOnWindows=Y
# defaule value for maximum number of large object files in a directory
# one megalob file gets generated for per multiple rows per large object column in
table, size of megalob is determined by informia.LoaderDatFileMaxSizeLimit
propertyinformia.bcp.maxLOBFilesInADir = 100000
# list of special characters in table or column Nameinformia.bcp.specialChars = / \\ $
( ) = & ^ @ # % ! *
# set below property to Y to enable Role assignment to partition
#informia.enableAccessLayer=N
# set the below property to Y to treat varbinary with length less than 2000 in SQL
server to be treated as varcharinformia.sqlServerVarBinaryAsVarchar=N
# set the below property to execute run procedure in archive from histroy
#informia.proceduresToExecute.inArchiveFromHistory=java://
com.applimation.archive.dao.impl.MoveFileDAOImpl
#property defines maximum allowed error count before aborting File Archive Loader Job.
(if unset, default is 4
)#informia.fasLoaderErrorThresholdCount=4
#property defines size limit of megalob file in MB. (if unset, default is 20480 MB.)
#informia.LoaderDatFileMaxSizeLimit=20480
# set the below property to N to stop using the InsertAsSelect in CTD and GC step.
Default behavior is Yes.
#informia.insertAsSelectFlag=Y
# Run DGA Data Collection as admin user or application user (default true)
# informia.dga.useAdminUser=true
#property specifies encoding technique used by tomcat
#informia.httpCharacterEncoding=UTF-8
#property defines NULL Representation in bcpinformia.NULLRepresentationInBcp=\u007F
#Property to turn off loading visualization on server startup. Value Y is the default
behaviour.
#informia.loadDataVisualization=N
#Property overrides the default RMI port used to connect to visualization server. the
default value is 1129
#informia.visualizationPort=1129
#Property overrides the default RMI host used to connect to visualization server. the
default value is localhost
#informia.visualizationHost=localhost
#property to enable DDM support queryinformia.enableDDMSecurity=N
#Property to define the lucene index directory, this directory should be outside your
tomcat install, to avoid any impacts on upgrade.
#Provide the absolute path of the directory. Use double slash for Windows path, eg : C:\
\ILM\\lucene
#informia.keywordSearchIndexDir=/u01/611/lucene/
#property to specify number of threads used in data vault Loader job.By Default, the ILM
engine calculates the value by dividing the number of CPU cores by 2
#informia.faThreadCount=
#Data Archive page that appears when user logs in. enter the url of the page you want
the user to first see, based on the user's system-defined role.
#informia.welcomePage=
#Remote storage metadata loading fails if multiple threads are used, to use single
thread, enable this property by replacing 'N' with 'Y'
#informia.metaDataLoaderThreadCount=N
#Property to define the number of threads created by indexing
job.#informia.maxActiveIndexThreads=10
# Storage option for XMLTYPE columns in Staging and History tables
# If set, should be either CLOB or BINARY XML
# informia.XMLTYPE_Storage_Option=CLOB
#Property to define the NVARCHAR(MAX) and VARCHAR(MAX) to
VARCHAR(32000)#informia.sqlServerNVarcharMaxAndVarcharMaxToVarchar=N#set the following
properties to extract the data in non-ascii format
#informia.extractSqlServerBinaryasHexaData=N
```

```
#informia.extractSqlServerVarBinaryasHexaData=N
#informia.extractSybaseVarBinaryasHexaData=N
#set the following properties to extract bit as byte
#informia.extractSybaseBitAsByte=N
# Property to let know the start year of 5-digit julian data
#informia.julianLimiterYear=1940
# For Windows network driver as staging. Wait time in milli seconds. Default value is 0
seconds.
#For a value of 5000. File Archive Loader job waits for 5 seconds.
#informia.regfileWaitTimeForNas=5000
#informia.loadWaitTimeForNas=5000
#Commit intervel for number of update querires when making scts online, if this propery
is not enabled default values is 10000#informia.updateSCTCommitInterval=10000
#Property to be set if default encoding defined by file.encoding and operating system
locale are different #if inactive the default value is taken from property file.encoding
defined in startApplimation
#informia.loadExternalAttachmentsFileEncoding=#Properties to be provided for enabling
Single sign on.
#informia.sso.enable = N
#Path of meta-data file generated after configuring the application to
IDP.#informia.idp.metedata.file = idp.xml#Path of the key which will be used to encrypt
the authRequest send to IDP.
#informia.key.path = key
#Alias name used to generate the key mentioned above.
#informia.key.alias.name = alias
#Password used to generate the key mentioned above in encrypted
format.#informia.key.password = password
# Property hold Certificate file path to enabled SSL from DA side to communicate with
IDV which running on SSL mode#informia.idv.ssl.certificatePath=
# Property to enable the authentication for the API requests, default value is
N.#informia.api.authentication=N#property defines size limit of export XML files in MB.
(if unset, default is 1024 MB.)
#informia.MetaDataExportXMLFileMaxSize=1024
#property to enable import metadata from database option in EDM, default value is Y.
#informia.enable.mining=Y
#property to enable enhanced import option in EDM, default value is Y.
#informia.enable.enhanced.import=Y
#Property to disabled jobids for the browse data
queriesinformia.disableAddingExtractJobIdPredicate=Y
#property to enable traditional import metadata option in EDM, default value is N.
#informia.enable.traditional.import=N
#Command to generate encryption key using third party utility
#informia.encryptionkey.command=java -classpath cryptokey.jar
com.crypto.GenerateKey#property defines the number of admin threads to be used in
encrypting the idv data. default value is 5.
#informia.encryption.RegistrationThreadCount=10
#property defines the number of worker threads to be used in encrypting the idv data.
default value is 10.
#informia.encryption.WorkerThreadCount=50
#property that defines SCT file regisration type in IDV. property values can be sql/
admin , default is sql.
#informia.encryption.SCTRegistrationType=sql
#property defines the admin timeout in seconds
#informia.encryption.AdminTimeout=120
#property defines delete data files or not. default values is N ( no)
#informia.encryption.DeleteOldDataFiles=N
#encrypt idv data job property defines if idv should execute the encrypt commnad in
debug mode or not, default value is N ( no)
#informia.encryption.IDVDeBugMode=N
# Property to disable the signing of JAR files using timestamp server. Default value is
Y.
#informia.useTimestampServerForJARSigning=N
#Set the following property to Y to disable the daylight saving time
#informia.disableDaylightSavingTime=Y
#Set this property to add wait time for the RuntimeUtil used in ssadrv. For a value 10 -
we are adding wait time of 10 seconds[10 iterraions of 1 second each]. Default value is
10 seconds.
#informia.waitTimeForRuntimeStreamRead=10
#Set the following property to Y to display column descriptions for the columns on
Browse Data pageinformia.useColumnDescriptionOnBrowseDataPage=Y
#Set the following property to Y to apply the Business Rules during the Archive from the
```

```
History jobinformia.applyBusinessRulesDuringArchiveFromHistory=N
#Set the following property to Y to enable the indexing of ZIP files present in
AM_ATTACHMENTS table in Create Index on Data Vault job
#informia.indexZIPFiles=Y
#Set the following property to Y to disable CSRF check
#informia.disableCSRFCheck=Y
#Set wait time to open DAT after an failed attempt in milli seconds
#informia.loadWaitTimeForFOS=5000

#Enter a custom file name and location for the master key file. For example: C:\\ILM\
\Masterkey\\key.file. The default  value is the Data Archive installation directory.
#informia.encryption.masterKeyPath=

#Property to enable or disable SQL Safe Validation. Default value is Y. To disable, set
to N.
informia.SQlSafeValidation=Y
```

Updating the Password Encryption Algorithm

Data Archive user passwords and source and target connection passwords are encrypted and stored in the AMHOME tables. The passwords are encrypted with the RC4 algorithm.

You can change the algorithm used to encrypt the passwords to the AES256 algorithm. The AES256 algorithm adds an additional layer of encryption with the master key. By default, the master key is stored in a key.file in the Data Archive installation directory. You can change the name and location of the file if required.

Note: You need the master key to decrypt and access the host key which provides access to the passwords. You cannot decrypt the passwords if you lose or edit the master key. Take a backup of the file and store the file in a secure location.

Use the AESEncryptionUtility to encrypt passwords with the AES256 algorithm and to update password encryption when required.

Note: After each successful run of the encryption utility, users must download Enterprise Data Manager.

Updating the Encryption Algorithm to AES256

The first time you run the AESEncryption utility, the utility updates the encryption algorithm to AES256 and creates a master key.

Perform the following tasks before you run the utility:

- Back up the AMHOME data.

1. Stop the Data Archive server.
2. Ensure that the Java path is set to the Azul JDK available in the Data Archive installation directory.
3. Optional. You can change the default name and location of the master key file that the utility generates. By default, the utility generates a Key.file in the Data Archive installation directory. To change the file name or location, perform the following steps:
   a. Open the following file: `<Data Archive installation directory>/conf.properties`
   b. Uncomment the *Informia.encryption.masterKeyPath* property and enter the required file name with the complete path.

c. Save the changes.

   The utility creates a master key with the file name and location that you specify.

4. Browse to the following location: `<Data Archive installation folder>/optional/`
5. Run the AESEncryptionUtility.bat file on Windows or the AESEncryptionUtility.sh file on UNIX.
6. The utility verifies the current encryption algorithm. If the algorithm is RC4, the utility prompts you to confirm whether you want to change the encryption algorithm. Enter Y to continue.

   The utility encrypts the passwords with the AES256 algorithm and creates the host key in the database.
7. Choose if you want to use a random master key. If you enter Y to use a random master key, the utility creates the key and stores the encrypted key in the `<Data Archive installation directory>/key.file` or the custom name and location you specified in step 2.
8. To enter a custom master key, enter N and then enter an input string.

   The utility encrypts the string and stores the encrypted key in the `<Data Archive installation directory>/key.file` or the custom name and location you specified in step 2. The utility indicates that the task was completed successfully and exits.

   The following section is an example of the utility run with a custom master key:

```
E:\INFAUSER\NEW\651_108\ILM\optional>AESEncryptionUtility.bat

E:\INFAUSER\NEW\651_108\ILM\optional>set PWD=E:\INFAUSER\NEW\651_108\ILM\optional\

E:\INFAUSER\NEW\651_108\ILM\optional>if "E:\INFAUSER\NEW\651_108\ILM\optional\\..
\java" == "" (set JAVA_HOME=E:\INFAUSER\NEW\651_108\ILM\optional\\..\java )

E:\INFAUSER\NEW\651_108\ILM\optional>set CLASSPATH=.;.;E:\INFAUSER\NEW\651_108\ILM
\optional\\..\webapp\WEB-INF\lib\infafas-6.5.1.jar;E:\INFAUSER\NEW\651_108\ILM
\optional\\..\webapp\WEB-INF\
lib\commons-collections-3.2.2.jar;E:\INFAUSER\NEW\651_108\ILM\optional\\..\webapp
\WEB-INF\lib\bc-fips-1.0.2.1.jar;

E:\INFAUSER\NEW\651_108\ILM\optional>cd ..\webapp\WEB-INF\lib\

E:\INFAUSER\NEW\651_108\ILM\webapp\WEB-INF\lib>"E:\INFAUSER\NEW\651_108\ILM\optional\
\..\java\bin\java.exe" -Dfile.encoding=UTF-8
com.applimation.util.AESEncryptionUitlity
log4j:WARN No appenders could be found for logger
(org.springframework.core.env.StandardEnvironment).
log4j:WARN Please initialize the log4j system properly.
log4j:WARN See http://logging.apache.org/log4j/1.2/faq.html#noconfig for more info.
Do you want to change the encryption algorithm to the new algorithm? (press y/n):
y
Do you want to use a random Master key? (press y/n):
n
Enter the Master key:
sajmkda873489^%@&H*!X4jm8xcmcmskl
Update of encrypted values completed successfully.

E:\INFAUSER\NEW\651_108\ILM\webapp\WEB-INF\lib>cd ../../../optional

E:\INFAUSER\NEW\651_108\ILM\optional>pause
Press any key to continue . . .
```

Note: Users must download Enterprise Data Manager each time that you run the encryption utility.

Updating Encrypted Passwords

You can run the utility to update the encryption of passwords encrypted with the AES256 algorithm.

Perform the following tasks before you run the utility:

- Back up the AMHOME data.

1. Stop the Data Archive server.
2. Ensure that the Java path is set to the Azul JDK available in the Data Archive installation directory.
3. Applicable if you store the master key in a custom location. Verify that the master key file is available at the location specified in the conf.properties file.
4. Browse to the following location: `<Data Archive installation folder>/optional/`
5. Run the AESEncryptionUtility.bat file on Windows or the AESEncryptionUtility.sh file on UNIX.
6. The utility verifies that the current encryption algorithm is AES256 and prompts you to choose to change the master key.

   Enter Y to change the master key or N to continue without changing the master key.
7. If you choose to change the master key, choose if you want to use a random master key. If you enter Y to use a random master key, the utility creates the key and stores the encrypted key in the `<Data Archive installation folder>/key.file` or the custom location specified in the conf.properties file.

   Note: The utility creates a backup of the existing master key file by adding the time stamp and creates a new master key file for the current run. You can use the backup file if you need to resume the utility.
8. To enter a custom master key, enter N and then enter an input string.

   The utility encrypts the string and stores the encrypted key in the `<Data Archive installation folder>/key.file` or the custom path specified. The utility renames the existing master key file by adding the timestamp.
9. Choose whether you want to change the host key. Changing the host key updates the encryption of passwords in the database.

   Enter Y to change the host key. The utility indicates that the task was completed successfully and exits.

   The following section is an example of the utility run to change the host key without changing the master key:

```
E:\INFAUSER\NEW\651_108\ILM\optional>AESEncryptionUtility.bat

E:\INFAUSER\NEW\651_108\ILM\optional>set PWD=E:\INFAUSER\NEW\651_108\ILM\optional\

E:\INFAUSER\NEW\651_108\ILM\optional>if "E:\INFAUSER\NEW\651_108\ILM\optional\\..
\java" == "" (set JAVA_HOME=E:\INFAUSER\NEW\651_108\ILM\optional\\..\java )

E:\INFAUSER\NEW\651_108\ILM\optional>set CLASSPATH=.;.;E:\INFAUSER\NEW\651_108\ILM
\optional\\..\webapp\WEB-INF\lib\infafas-6.5.1.jar;E:\INFAUSER\NEW\651_108\ILM
\optional\\..\webapp\WEB-INF\
lib\commons-collections-3.2.2.jar;E:\INFAUSER\NEW\651_108\ILM\optional\\..\webapp
\WEB-INF\lib\bc-fips-1.0.2.1.jar;

E:\INFAUSER\NEW\651_108\ILM\optional>cd ..\webapp\WEB-INF\lib\

E:\INFAUSER\NEW\651_108\ILM\webapp\WEB-INF\lib>"E:\INFAUSER\NEW\6.5.1_104BLD\ILM\java
\bin\java.exe" -Dfile.encoding=UTF-8 com.applimation.util.AESEncryptionUitlity
log4j:WARN No appenders could be found for logger
(org.springframework.core.env.StandardEnvironment).
log4j:WARN Please initialize the log4j system properly.
log4j:WARN See http://logging.apache.org/log4j/1.2/faq.html#noconfig for more info.
Do you want to rotate the Master key? (press y/n):
n
You have entered "N". The utility will not convert the passwords.
Do you want to rotate the Host key? (press y/n):
```

```
y
Host key rotation is completed successfully.

E:\INFAUSER\NEW\651_108\ILM\webapp\WEB-INF\lib>cd ../../../optional

E:\INFAUSER\NEW\651_108\ILM\optional>pause
Press any key to continue . . .

E:\INFAUSER\NEW\651_108\ILM\optional>
```

Note: Users must download Enterprise Data Manager each time that you run the encryption utility.

Resuming the Encryption Utility

If you have issues while running the utility or if the utility fails to complete, see the following log file for information: `<Data Archive installation directory>/optional/AESEncryption.txt`

1. Stop the Data Archive server.
2. Browse to the following location: `<Data Archive installation folder>/optional/`
3. Run the AESEncryptionUtility.bat file on Windows or the AESEncryptionUtility.sh file on UNIX.
4. The utility prompts you to choose to resume the previous flow. Enter Y to resume the flow.

   The utility indicates that the task was completed successfully and exits.

Troubleshooting the Encryption Utility

Consider the following troubleshooting tips when you use the encryption utility or if you encounter issues after you run the utility:

javax.crypto.BadPaddingException: Error finalising cipher data: pad block corrupted at org.bouncycastle.jcajce.provider.BaseCipher.engineDoFinal(Unknown Source) at javax.crypto.Cipher.doFinal(Cipher.java:2168) at com.applimation.util.AES256Utility.cbcDecryptBc(AES256Utility.java:382) at com.applimation.dao.impl.AES256EncryptorDaoImpl.decryptThePassword(AES256EncryptorDaoImpl.java:66)

Use case

You might notice this issue when you log in to Data Archive or when you run the utility. The following issues generate the error:

- The master key is old.
- The master key is corrupt.

Resolution

If the master key is old, use the new master key.

If the master key is corrupt, use the backup master key file.

java.lang.ArrayIndexOutOfBoundsException2021-10-14 20:50:09,562 ERROR [com.applimation.services.thread.AMServiceThread] [JOB$:786-JOBSTEP$:1642]

Use case

This error occurs if the AESEncryption utility does not run successfully because of session interruption or unexpected session closure while utility is running.

Resolution

Run the utility again and choose the Resume flow.

java.security.InvalidKeyException: Illegal key size or default parameters3. Caused by: java.security.InvalidKeyException: Illegal key size or default parameters at javax.crypto.Cipher

Use case

This error occurs if you use an unsupported JDK version to run the utility.

Resolution

Use the JDK bundled with the installer and available in the <Data Archive install directory>.

java.io.FileNotFoundException: C:\Installation_Directory\key.file (The system cannot find the file specified) at java.io.FileInputStream.open0(Native Method) at java.io.FileInputStream.open(FileInputStream.java:195)

Use case

The error occurs if the master key is not available in the location specified in the conf.properties file or in the installation directory.

Resolution

Copy the master key to the Data Archive installation directory or to the custom path that you specified in the conf.properties file and run the utility.

System Profile

The system profile contains global properties and properties that are specific to features.

Changes to the system profile are dynamic and do not require you to restart the ILM application server.

System profile properties are organized in tabs on the **Configuration Settings** page. Each tab corresponds to a system function.

General System Profile Properties

General system profile properties include email properties and other global properties.

Configure the following system profile properties on the **General** tab on the **Configuration Settings** page:

Mail Server TLS Enabled

If the mail server is TLS enabled, select this check box and provide the corresponding mail server details.

Currency

The currency code of your business locale. For example, USD.

Mail Server

IP address or host name of the company mail server. Required for email notifications related to jobs. For example, `mail.yourcompany.com`.

Mail Server Port

Port used to communicate with the host where the ILM repository schema is installed. Required for email notifications related to jobs.

Mail Server User

Email address of the default sender. Required for email notifications related to jobs. When you set up email notifications for jobs, emails are sent from this email address. For example, `admin@yourcompany.com`.

If you leave the *Mail Server User* field blank, you will receive an email from the `AMHome@informatica.com` address after the Test Email Server Configuration job runs.

Mail Server Password

Password for the company mail server. Required for email notifications related to jobs.

Company Logo

Absolute path for the image file of the company logo.

Job Auto Refresh Time Period

Time interval, in seconds, that determines how often the job status refreshes when you monitor the list of jobs.

Host

Machine where the ILM repository is located.

Port

Port number of the host where the ILM repository is installed.

Service Name

Name of the host where the ILM repository is installed.

ILM Repository Administrator User

Administrator of database that hosts the ILM repository.

External Database Status

Enables EXTERNAL_STORAGE as a connection type.

Profiling and Discovery System Profile Properties

Profiling and discovery system profile properties include properties related to Data Quality, such as domain name and authentication type.

Configure the following system profile properties on the **Profiling and Discovery** tab on the **Configuration Settings** page:

Domain Name

Name of the Informatica domain.

Host

Name of the machine hosting the master gateway node.

Port

Port on which the Data Integration Service accepts requests.

Default is 6005.

User Name

Name of the user with access to the Model Repository Service.

Password

Password for the user name.

Authentication Type

Informatica domain authentication method.

Use Native or LDAP authentication.

Model Repository Service

Name of the Model Repository Service.

Data Integration Service

Name of the Data Integration Service.

Node Name

Name of the PowerCenter node.

Data Discovery Portal System Profile Properties

Data Discovery Portal system profile properties include search and result date format properties.

Configure the following system profile properties on the **Data Discovery Portal** tab on the **Configuration Settings** page.

Date Format for Search Results

The default display format for datetime values in data discovery search results.

Default is `dd-MMM-yyyy HH:mm:ss.f`.

Date Format for Search Criteria

The required format for datetime values in data discovery searches. When you select a datetime column in a data discovery search field, Data Archive displays the default discovery search date format. For example, you want to search the Data Vault for fields with a specific order date. In the search field, Data Archive displays the ORDER_DATE column as "ORDER_DATE (mm/dd/yyyy)."

Default is `mm/dd/yyyy`.

Default Maximum Number of Records in Results

The maximum number of records that you want to view in the Search Within an Entity in Data Vault search results. Data Archive uses the value you specify in this property as the default value in a search query. However, a user can override the default value at the time of the search by specifying another value on the **Search Within an Entity in Data Vault** page.

For example, you enter 50 as the default maximum number of records you want Data Archive to display in the results. A user wants to view records with Amsterdam in the City column. When the user queries the Data Vault, Data Archive searches the entity until it finds the first 50 records that match the search criteria. If the user changes the value to 200 on the **Search Within an Entity in Data Vault** page, then Data Archive searches for the first 200 records that match the criteria. Future search queries will default to a maximum of 50 records in the search results.

Note: When you export results, Data Archive exports all records that meet the search criteria regardless of the value in the Default Maximum Number of Records in Results property.

Enter a value from one to 10,000. Do not use a comma, period, or space in the value.

Default is 1000.

Design Object Repository System Profile Properties

Design object repository system profile properties include properties related to the ILM repository.

Configure the following system profile properties on the **Design Object Repository** tab on the **Configuration Settings** page.

Design Object ID Repository Host

IP address or host name of the ILM repository schema.

Default is the ILM repository value that was used during the installation. Keep the default value. If you change the value, users might not be able to log in to Data Archive.

Design Object ID Repository Port

Port number of the machine that hosts the ILM repository.

Default is the ILM repository value that was used during the installation. Keep the default value. If you change the value, users might not be able to log in to Data Archive.

Design Object ID Repository Service

Database service name or database name of the ILM repository.

Default is the ILM repository value that was used during the installation. Keep the default value. If you change the value, users might not be able to log in to Data Archive.

Design Object ID Repository User Name

User name of the ILM repository schema.

Default is the ILM repository value that was used during the installation. Keep the default value. If you change the value, users might not be able to log in to Data Archive.

Design Object ID Repository User Password

Password of the ILM repository schema.

Security System Profile Properties

Security system profile properties include properties related to security groups and password expiration.

Configure the following system profile properties on the **Security** tab on the **Configuration Settings** page.

User Password Must be Changed After

Number of days before a user must change the login password.

If the user logs in after the password expires, the **Change Password** window appears requesting the user to enter a new password.

Enforce Enhanced Security Model

Determines if users must be assigned to a security group before they can view or access a destination database.

By default, users can view and access any destination database without being assigned to a security group. Enable if you want to limit destination database access to users assigned to a security group. To access a source database, users must always be assigned to a security group.

Data Visualization System Profile Properties

Configure the following system profile property on the **Data Visualization** tab on the **Configuration Settings** page:

User cache timeout

The number of minutes that user credentials are cached in the Data Visualization server memory without querying the ILM repository. If privileges change during the time the credentials are cached, the user continues to have access to the data. In production environments where privileges are relatively

constant, set a high value to improve the time it takes data visualization to respond to requests. If you prefer to check user credentials with every query to the ILM repository, set the value to 0 to turn off caching.

Default is 5 minutes.

Configuring the System Profile

Configure the system profile to specify global settings for databases and users.

1. Click **Administration** > **System Profile**.
2. Enter the system profile properties.
3. Click **Save.**

Email Notification

When you schedule a job, you can request email notifications for changes to the job status. You configure notifications based on the job status, such as when a job completes, terminates, or includes an error.

When you configure a job to send email notifications, Data Archive sends the email from the mail server user defined in the system profile. Data Archive sends the email to the email address of the user that schedules the job. The user account includes the email address. You can override the default recipient email address and specify another email address when you schedule the job.

Test Email Server Configuration Job

Run the Test Email Server Configuration job to verify that Data Archive can connect to the mail server. Run the job after you define the mail server properties in the system profile. After you verify the connection, you can configure email notifications when you schedule jobs or projects.

When you run the Test Email Server Configuration job, the job uses the mail server properties defined in the system profile to connect to the mail server. If the connection is successful, the job sends an email to the email address that you specify in the job parameters. Data Archive sends the email from the mail server user address defined in the system properties. If the connection is not successful, you can find an error message in the job log stating that the mail server connection failed.

Email recipient security is determined by the mail server policies defined by your organization. You can enter any email address that the mail server supports as a recipient. For example, if company policy allows you to send emails only within the organization, you can enter any email address within the organization. You cannot send email to addresses at external mail servers.

Setting Up Email Notification

Before you can configure jobs to send email notifications, you must set up the system for email notification.

1. Configure the mail server properties in the system profile.
2. Schedule the Test Email Server Configuration standalone job to test the mail server connection.

After you set up email notifications, all users that have authorization to schedule jobs can configure email notifications.

Environment Variables

Verify that the locale setting is compatible with the code page for the Data Vault Service, if you archive to the Data Vault. If the locale setting is not compatible with the code page, then non-English characters might not appear correctly from the Data Discovery portal or other reporting tools.

Verify that the LANG environment variable is set to UTF-8 character encoding on the Data Vault Service. For example, you can use `en_US.UTF-8` if the encoding is installed on the operating system.

Use the following command to verify what type of encoding the operating system has:

```
local -a
```

The command returns the languages installed on the operating system and the existing locale settings.

In addition, verify that all client tools, such as PuTTY and web browsers, use UTF-8 character encoding.

To configure UTF-8 character encoding in PuTTY, perform the following steps:

1. In the PuTTY configuration screen, click **Window** > **Translation**.
2. In the **Received data assumed to be in which character set** field, choose `UTF-8`.

Importing SSL Certificates to the Enterprise Data Manager Keystore

If you enable SSL authentication for Data Archive, you must import the SSL certificate into the Enterprise Data Manager keystore.

Note: To run the commands in the task, you must use the JDK shipped in the EDM.zip folder.

1. Log in to the Data Archive application using HTTPS.
2. Download the certificate from the browser.
3. Open a command prompt window.
4. Run the following commands to set the JAVA_HOME and PATH variables:

   ```
   set JAVA_HOME=C:\Users\<Administrator>\Downloads\EDM\windows_java
   set PATH=%JAVA_HOME%\bin;%PATH%
   ```

5. Run the following command to import the certificate to the Enterprise Data Manager keystore:

   ```
   keytool -import -trustcacerts -alias <server name> -file <file name of the downloaded certificate with full path> -keystore <Enterprise Data Manager keystore path with keystore name>
   ```

   For example: `-import -trustcacerts -alias INVR80DSG1431 -file C:\InfaPrd\Certificates\1431certificate.cer -keystore C:\Users\infauser\Downloads\EDM1431\windows_java\jre\lib\security\cacerts`

CHAPTER 3

Database Users and Privileges

This chapter includes the following topics:

Database Users and Privileges Overview

After you install Data Archive, set up the required database users and configure additional privileges.

Data Archive uses different database users during the archive process. Before you run an archive or retirement project, set up the required database users. Verify that the users have the correct authorizations.

You may need to configure additional privileges depending on the source database or the source application that you want to archive data from. Set up additional privileges to archive from IBM DB2 sources or to retire SAP applications.

Database Users

Data Archive uses different database users during the installation and archive process. Some processes use existing database users in your source database, such as the administration user. Before you install the product and run a Data Archive project, create the required database users and verify that the users have the correct authorizations.

The installation requires both an Administration user and an ILM repository user.

Depending on the task, Data Archive requires one or more of the following users:

- Administration user
- Production application user
- Staging user
- History read-only user
- History application user
- Combined user and archive user for seamless access

Administration User

The administration database user has DBA rights to the database, including the ability to run DDL and access system-level objects.

The administration user typically resides on both the source and target databases. During the archive process, the administration user queries the catalog tables. You can use the administration user to create and drop the public database link if the database link to the repository does not exist.

The administration user requires the following user privileges for the installation:

Oracle Privileges	Microsoft SQL Server Privileges	IBM Db2 Privileges
ALTER SESSION	ALTER EVENT SESSION	ALTER SESSION
ALTER USER	ALTER USER	
CREATE SESSION	CREATE EVENT SESSION	CREATE EVENT SESSION
CREATE USER	CREATE USER	
CREATE TABLESPACE	Not applicable	
GRANT CONNECT	GRANT CONNECT/CREATE USER	
SELECT_CATALOG_ROLE	SELECT	
Grant any privilege TO *user_name* with admin option.	Grant any privilege TO *user_name* with admin option.	Grant any privilege TO *user_name* with admin option.

The administration user typically requires the following user privileges for archive processes:

- SELECT DBA_SEGMENTS. Required for the generate candidates row count report.
- SELECT ANY TABLE. Not required if the administration user has access to application user tables.
- CREATE PUBLIC DATABASE LINK. Not required if the archive connection definition includes the database link name.
- DROP PUBLIC DATABASE LINK. Not required if the archive connection definition includes the database link name.

The administration user requires select privileges on the following views in order for the DGA Data Collection job to run successfully:

- DBA_DATA_FILES
- DBA_FREE_SPACE
- DBA_INDEXES
- DBA_SEGMENTS
- DBA_TABLES
- DBA_TABLESPACES

Microsoft SQL Server Requirements

On Microsoft SQL Server, the Administration user must have the serveradmin role.

ILM Repository User

The ILM repository user creates and owns the database objects in the ILM repository schema. The database objects contain repository data such as application metadata, users, jobs, and workbench definitions. The ILM repository user is also called the AMHOME user.

Data Archive requires the ILM repository user. You can create the user before the installation, or the installation can create the user and grant the required privileges.

For example, you can run the following script to create the ILM repository user:

```
create user username identified by username default tablespace tablespace_name temporary
tablespace TEMP;
```

The following table lists the privileges required for the ILM repository:

Oracle Privileges	Microsoft SQL Server Privileges	IBM DB2 Privileges
ALTER SESSION	ALTER EVENT SESSION	ALTER SESSION
CONNECT	For information about creating connections to the SQL Server database, see http://msdn.microsoft.com/en-us/library/s4yys16a(v=vs.90).aspx	CONNECT
CREATE DATABASE LINK	Not applicable	Not applicable
CREATE PROCEDURE	CREATE PROCEDURE	CREATE PROCEDURE
CREATE SEQUENCE	CREATE SEQUENCE	CREATE SEQUENCE
CREATE SESSION	CREATE EVENT SESSION	CREATE EVENT SESSION
CREATE SYNONYM	CREATE SYNONYM	CREATE SYNONYM
CREATE TABLE	CREATE TABLE	CREATE TABLE
CREATE TYPE	CREATE TYPE	CREATE TYPE
CREATE VIEW	CREATE VIEW	CREATE VIEW
DEFAULT ROLE ALL	Default roles enabled when user logs in.	Not applicable

IBM DB2 Privileges

You must create the database schema for the ILM repository before the installation. The user that has authorization to the ILM repository schema must have access to the catalog tables.

Installation on SSL Enabled Databases

If you are using an SSL enabled database to install the ILM repository, you must import the SSL certificates from the database server into the Java truststore on the ILM installation host machine.

You must import SSL certificates to the Azul Java truststore available in the following path:

```
<DA Installer build>/DA/<windows/linux java>
```

Production Application User

The production application user resides on the source database and is either the owner of the tables that you want to archive or partition, or the user that has full rights to the tables that you want to archive or partition. During the archive process, you can use the production application user to delete from the source. Alternatively, you can configure the source connection to use the staging user to delete data from the source.

The production application user typically requires the following user privileges:

- Select, insert, update, and delete privileges on all application tables.
- CREATE ANY TRIGGER and ALTER ANY TRIGGER. Required if triggers on application tables are owned by a different user.
- SELECT ANY TABLE. Required to access the staging tables only if you use the production application user for the delete from source step.
- Read-only access to the directory that contains binary files if the source data contains BFILE datatypes.

Staging User

The staging database user stores interim tables for business rule validation and generates candidates during the archive and partitioning processes. If you enable staging for the archive source connection, the staging user also creates staging tables. The staging tables temporarily store a copy of the archived data before the archive job moves the data to the archive or history database. The staging user typically resides in the source database.

Required Privileges

The following table lists the privileges required for the staging database user:

Oracle Privileges	Microsoft SQL Server Privileges	IBM DB2 Privileges
ALTER SESSION	ALTER EVENT SESSION	ALTER SESSION
CONNECT	For information about creating connections to the SQL Server database, see http://msdn.microsoft.com/en-us/library/s4yys16a(v=vs.90).aspx	CONNECT
CREATE DATABASE LINK	Not applicable	Not applicable
CREATE PROCEDURE	CREATE PROCEDURE	CREATE PROCEDURE
CREATE SEQUENCE	CREATE SEQUENCE	CREATE SEQUENCE
CREATE SESSION	CREATE EVENT SESSION	CREATE EVENT SESSION
CREATE SYNONYM	CREATE SYNONYM	CREATE SYNONYM
CREATE TABLE	CREATE TABLE	CREATE TABLE
CREATE TRIGGER	CREATE TRIGGER	CREATE TRIGGER

Oracle Privileges	Microsoft SQL Server Privileges	IBM DB2 Privileges
CREATE TYPE	CREATE TYPE	CREATE TYPE
CREATE VIEW	CREATE VIEW	CREATE VIEW

Additional Privileges

The following privileges may be required depending on your specific configuration:

- DELETE ANY TABLE. Required to access the staging tables if you use the staging user for the delete from source step. If you cannot grant this privilege, you can do a direct delete grant on all application tables that are referenced in the metadata. Or, you can configure the source connection to use the application user for the delete from source step. If you use the application user, the application user needs access to the staging tables.
- SELECT ANY TABLE. Not required if the staging user has access to the application user tables. If you cannot grant this privilege, you can do direct select grants on all application tables that are referenced in the metadata.
- UPDATE ANY TABLE. If you cannot grant this privilege, you can do direct update grants on all application tables that are referenced in the metadata.
- EXECUTE ANY PROCEDURE. Required if the archive entities execute any procedures in the application user schema.
- ALTER ANY TRIGGER. Required if you use the staging user for the delete from source step.
- CREATE ANY TRIGGER. Required if you use the staging user for the delete from source step. If you cannot grant this privilege, then the staging user needs direct select privileges on the application tables instead. For example, grant select on ORDERS to amstage.
- EXECUTE ANY TYPE. Required to create staging tables with user defined types that are owned by a different user.

Oracle Applications Privileges

To access Oracle applications, users must have insert privileges on the following tables:

- AP.AP_HISTORY_CHECKS_ALL
- AP.AP_HISTORY_INV_PAYMENTS_ALL
- AP.AP_HISTORY_INVOICES_ALL
- PO.PO_HISTORY_POS_ALL
- PO.PO_HISTORY_RECEIPTS
- PO.PO_HISTORY_REQUISITIONS_ALL

Oracle Partition Exchange Privileges

If you configure the archive job to use Oracle partition exchange to delete source data, the staging user creates a table to identify the records to keep in the source. The archive job uses the partition in the table to swap the original source segment when the job deletes data from the source.

The following user privileges are required to use Oracle partition exchange to delete data from sources on Oracle databases:

- ALTER ANY TABLE
- DROP ANY TABLE
- ALTER USER QUOTA UNLIMITED

History Read-Only User

The history read-only user has read-only access to all of the archive tables and resides on the target database. You can use the history read-only user to create the database link for seamless access.

The history read-only user requires the following user privileges:

Oracle Privileges	Microsoft SQL Server Privileges	IBM DB2 Privileges
ALTER SESSION	ALTER EVENT SESSION	ALTER SESSION
CREATE SESSION	CREATE EVENT SESSION	CREATE EVENT SESSION
SELECT ANY TABLE Or, grant table level permissions for all archive tables.	SELECT TABLE	

History Application User

The history application user owns all of the archive tables and resides on the target database.

Required Privileges

The following table lists the user privileges that the history application user requires:

Oracle Privileges	Microsoft SQL Server Privileges	IBM DB2 Privileges
ALTER SESSION	ALTER EVENT SESSION	ALTER SESSION
CREATE DATABASE LINK	Not applicable	Not applicable
CREATE PROCEDURE	CREATE PROCEDURE	CREATE PROCEDURE
CREATE SEQUENCE	CREATE SEQUENCE	CREATE SEQUENCE
CREATE SESSION	CREATE EVENT SESSION	CREATE EVENT SESSION
CREATE SYNONYM	CREATE SYNONYM	CREATE SYNONYM
CREATE TABLE	CREATE TABLE	CREATE TABLE
CREATE TRIGGER	CREATE TRIGGER	CREATE TRIGGER
CREATE TYPE	CREATE TYPE	CREATE TYPE
CREATE VIEW	CREATE VIEW	CREATE VIEW

IBM DB2 Privileges

If you use the IBM DB2 load client utility or the import utility to insert data into the history database, then the history application user requires load and insert privileges.

Restore Privileges

When you restore archived data, the history application user functions like the production application user during an archive cycle. To restore archived data, the history application user requires the SELECT ANY TABLE privilege.

Combined User and Archive User for Seamless Access

The combined user and archive user have query-only access to archived data. The users are required for seamless access for combined and history data. The combined user has access to view both the current and history data. The archive user has access to view the history data.

For Oracle, PeopleSoft, and Siebel applications, create both users on the production database for performance reasons.

Both users require the following user privileges:

Oracle Privileges	Microsoft SQL Server Privileges	IBM DB2 Privileges
ALTER SESSION	ALTER EVENT SESSION	ALTER SESSION
CREATE CLUSTER	CREATE CLUSTER Use the Create Cluster wizard.	
CREATE DATABASE LINK	Not applicable	Not applicable
CREATE PROCEDURE	CREATE PROCEDURE	CREATE PROCEDURE
CREATE SEQUENCE	CREATE SEQUENCE	CREATE SEQUENCE
CREATE SESSION	CREATE EVENT SESSION	CREATE EVENT SESSION
CREATE SYNONYM	CREATE SYNONYM	CREATE SYNONYM
CREATE TABLE	CREATE TABLE	CREATE TABLE
CREATE TRIGGER	CREATE TRIGGER	
CREATE VIEW	CREATE VIEW	CREATE VIEW
EXECUTE ANY PROCEDURE	EXECUTE PROCEDURE	
SELECT ANY SEQUENCE	SELECT SEQUENCE	
SELECT ANY TABLE	Not applicable	
SELECT_CATALOG_ROLE	SELECT	
SELECT ANY DICTIONARY	Not applicable	

The following user privileges are required for Oracle applications:

- AQ_ADMINISTRATOR_ROLE
- CTXAPP
- HR_REPORTING_USER
- OLAP_USER
- INSERT ANY TABLE (or specific to Oracle application tables where the report/form writes to the table)
- UPDATE ANY TABLE (or specific to Oracle application tables where the report/form writes to the table)
- DELETE ANY TABLE (or specific to Oracle application tables where the report/form writes to the table)

Use the following privileges to help improve Data Archive performance:

- CREATE ANY OUTLINE
- DROP ANY OUTLINE

IBM DB2 Privileges

If you archive from an IBM DB2 source, set up privileges to bind the packages and to access data.

Bind Packages Privileges

The user that you provide in the standalone job to bind IBM DB2 packages requires specific privileges.

The IBM DB2 Bind Package job requires an user ID as one of the parameters. The user that you provide runs the bind process on the source. Verify that the user has the required database access.

The following table lists the required authorizations for each platform:

Platform	Authorization
IBM DB2 for z/OS	The user requires one of the following privileges: - SYSADM - SYSCTRL - BINDADD and CREATE IN COLLECTION NULLID
IBM DB2 for Linux, UNIX, and Windows	DBADM
IBM DB2 for AS/400 iSeries	CHANGE authority or higher on the collection NULLID where the driver creates the packages.

Binding IBM DB2 for z/OS Packages Manually

If the IBM DB2 Bind Package job fails on the IBM DB2 for z/OS source, you can bind the packages manually on the source.

The following message appears if the IBM DB2 Bind Package job fails:

```
[informatica][DB2 JDBC Driver]Bind process is not active.
Please ensure that the user has permissions to create packages.
Packages cannot be created in an XA Connection.
```

To resolve the issue, perform the following steps:

1. Go to the directory where you extracted the Data Archive installation files.
2. Open the `Optional` folder.
3. Extract the `bind_files.zip` file.

   The `bind_files.zip` file contains the `CNTLFILE.XMIT` and `DBRMFILE.XMIT` files.
4. Log in to the source mainframe.
5. Create the following datasets.
   - `<userID>.CNTLFILE.XMIT`

- `<userID>.DBRMFILE.XMIT`

For example, if your user ID is `jane12`, then use the following name for the datasets:

- `JANE12.CNTLFILE.XMIT`
- `JANE12.DBRMFILE.XMIT`

Note: Use uppercase for the name of the datasets.

6. Enter the following properties for the datasets: `DSORG=PS, RECFM=FB, LRECL=80, BLKSIZE=3120`
7. Access the source mainframe through FTP from the machine that contains the `CNTLFILE.XMIT` and `DBRMFILE.XMIT` files. Use the same user ID and password for the FTP as the one you used to log in to the source mainframe in step 4.
8. Enter `BIN` to use the binary transfer mode.
9. Enter the `PUT` command to transfer the `CNTLFILE.XMIT` and `DBRMFILE.XMIT` files to the source mainframe.
10. Verify that the data from the `CNTLFILE.XMIT` and `DBRMFILE.XMIT` files transferred to the datasets you created in step 5.
11. Log in to the source mainframe.
12. From the ISPF option 6 or TSO Ready mode, enter the following `RECEIVE` command: `RECEIVE INDS('userid.CNTLFILE.XMIT')`
13. At the prompt, enter `DA('userid.DDJDBC.CNTL')`
14. Press `ENTER`.

    The system creates the partitioned data set `DDJDBC.CNTL` that contains the JCL and distribution packages.
15. From ISPF option 6 or TSO Ready mode, enter the following `RECEIVE` command: `RECEIVE INDS('userid.DBRMFILE.XMIT')`
16. At the prompt, enter `DA('userid.DDJDBC.DBRMLIB')`
17. Press `ENTER`.

    The system creates the partitioned data set `DDJDBC.DBRMLIB` that contains the JCL and distribution packages.
18. Edit the job `userid.DDJDBC.CNTL(BIND)`. Follow the instructions in the job to complete the following tasks:
    a. Bind the packages for the DataDirect Connect for JDBC DB2 driver.
    b. Grant the user the `EXECUTE` privilege.

Login User Privileges

The user that you provide to connect or log in to IBM DB2 requires specific privileges to access data. You provide a login user when you define source and target connections and when you define connections to mine data in the Enterprise Data Manager.

When you define source and target connections for IBM DB2, you provide the user for the admin login, application login, and staging login names. All login names require the same privileges. When you define a connection to mine data in EDM, you provide the user to connect to IBM DB2.

The following table lists the privileges required for the user to connect to IBM DB2:

Platform	Privileges
IBM DB2 on z/OS	The user requires SELECT access to the following system tables in the SYSIBM schema: - SYSTABLES - SYSCOLUMNS - SYSPROCEDURES - SYSPARMS - SYSCOLAUTH - SYSTABAUTH - SYSKEYS - SYSINDEXES - SYSSYNONYMS - SYSROUTINES - SYSFOREIGNKEYS - SYSSCHEMAAUTH
IBM DB2 on Linux, UNIX, or Windows	The user requires SELECT access to the following system tables in the SYSCAT schema: - SYTABLES - COLUMNS - PROCEDURES - PROCPARMS - COLAUTH - TABAUTH - KEYCOLUSE - INDEXES - INDEXCOLUSE - REFERENCES - SCHEMATA - SCHEMAAUTH - ROLEAUTH (only for post Linux, UNIX, and Windows 95)
IBM DB2 on AS400/i-Series	The user requires SELECT access to the following system tables in the QSYS2 schema: - SYSTABLES - SYSCOLUMNS - SYSINDEXES - SYSCST - SYSKEYCST - SYSREFCST - SYSROUTINES - SYSPARMS - SYSFUNCS - SYSPROCS - SQLTABLEPRIVILEGES

SAP Application Retirement Privileges

To retire SAP applications, you need an SAP system user that includes remote function call (RFC) connection authorization.

For tables that store data in a readable format in the database layer, the retirement job uses the standard SAP database user to access data. The job accesses data directly from the database that hosts the SAP

system. The SAP database user includes all of the required authorizations to retire an SAP application. You do not need to create or edit a database user.

For tables that store data in a readable format in the application layer, the retirement job uses an SAP system user to log in to the SAP system and access data.

After you install Data Archive and install the SAP transports for application retirement, you assign a role to the SAP system user. The SAP transports include the ZINFA_RETIREMENT_PREPARATION role. The role includes the rest of the required authorizations. You specify the SAP system user in the source connection properties.

ZINFA_RETIREMENT_PREPARATION Role

The ZINFA_RETIREMENT_PREPARATION role includes authorization to run the retirement job. The role is included in the Data Archive installation. You assign the role to the SAP system user that you plan to use to run the job.

The role includes the following authorization classes:

- AAAB (Cross-application authorization objects)
- BC_A (Basis: Administration)

AAAB Authorization Class

The authorizations in the AAAB authorization class allow the remote RFC user to run functions that prepare the database tables for retirement.

The following table lists the authorizations that the AAAB authorization class includes:

Authorization Object	Field	Values
S_RFC	ACTVT	16
S_RFC	RFC_NAME	*
S_RFC	RFC_TYPE	*

BC_A Authorization Class

The BC_A authorization class includes the following authorization objects:

- S_BTCH_JOB. The authorizations allow Data Archive to create, run, and release retirement jobs in the SAP system.
- S_C_FUNCT. The authorizations allow retirement jobs to create folders in the file system to process attachments.
- S_DATASET. The authorizations allow retirement jobs to read, write, and delete on the file system that contains the attachments.

The following table lists the authorizations that the BC_A authorization class includes:

Authorization Object	Field	Values
S_BTCH_JOB	JOBACTION	DELE, PROT, RELE, SHOW
S_BTCH_JOB	JOBGROUP	*

Authorization Object	Field	Values
S_C_FUNCT	ACTVT	16
S_C_FUNCT	CFUNCNAME	*
S_C_FUNCT	PROGRAM	*
S_DATASET	ACTVT	06, 33, 34
S_DATASET	FILENAME	*
S_DATASET	PROGRAM	*

CHAPTER 4

Source Connections

This chapter includes the following topics:

Source Connections Overview

Create a connection for each source database that you want to archive data from. The source database is the database that stores the transactional data that you want to archive. It is usually the production database or application database. For restore jobs, it is the history database, where the data was originally archived to.

To restore archived data, create a source connection for the target or history database. In the restore scenario, the target database now becomes the archive source.

Note: If a source connection is used as part of File Archive or retirement definitions, you cannot modify the application version.

To retire a SAP application, perform the following steps:

1. Create a source connection for the database in which the SAP application is installed.
2. Enter the SAP host, client, and other SAP details to create the connection. If you do not enter the SAP host correctly, the File Archive Loader cannot map the SAP datatypes to the Data Vault datatypes.

Source Database Requirements

Verify the source database requirements. The source database is the location where you archive data from and is typically referred to as the production database. You can verify the requirements before or after the installation.

You can archive data from the following source database types:

- IBM DB2
- Informix
- JDBC
- Microsoft SQL Server
- MongoDB
- Netezza
- Oracle
- Sybase ASE
- Teradata
- Salesforce

For other databases, for example MySQL, use a generic JDBC driver to connect to the source database as a generic JDBC source connection.

You may need to create a staging tablespace or have additional staging space available on the source database. The staging tablespace creates space for the data only and not the index.

Data Archive requires a staging tablespace in the following cases:

- Live archiving jobs.
- Archive jobs that use Oracle partition exchange to purge data from the source database.

Data Archive requires staging space for retirement jobs that retire SAP applications with attachments.

Live Archiving

Create a data staging tablespace in the source database for all live archive jobs. For example, AM_STAGING_D.

All live archive jobs require staging space. The archive job uses the staging space to store the following tables:

Interim tables

The archive job creates interim tables for business rule validation and candidate generation during the archive process.

The staging space requirements depend on the size of the interim tables. Configure the tablespace size to the number of archived rows from the largest table in a single archive job multiplied by the average row size. A general guideline is to use 1K for the average row size.

Staging tables

If you enable staging for the archive source connection, the archive job creates staging tables. The staging tables temporarily store a copy of the archived data before the archive job moves the data to the archive or history database.

The staging space requirements depend on the size of the data in the archive job. Configure the tablespace size to the largest amount of archived data in a single archive job.

Staging space is not required for retirement jobs. Retirement jobs do not validate business rules, generate candidates, or create staging tables because the jobs archive all source data.

Oracle Partition Exchange Purging

Verify that the staging tablespace includes space for archive jobs that use Oracle partition exchange to purge data from the source database. The amount of staging space depends on the amount of data that is going to be kept in the source table.

When the archive job runs the partition exchange procedure, the procedure only creates one keep table. When the database completes the partition exchange for the source table, the procedure drops the keep table. This process is repeated for every table that is configured for partition exchange. The size of the keep table depends on each source table. Verify that you have enough staging space for the maximum difference between the source and staging tables.

To estimate the amount of required staging space, configure the archive job to pause after the copy to staging step. Review the log to calculate the sizes of the staging tables and the source tables.

SAP Application Retirement

Verify that you have staging space for retirement jobs that archive SAP applications. File system staging space is required to archive attachments.

To retire attachments from SAP applications, you must have staging space in a file system that is accessible to the SAP system. The staging space requirements depend on the amount and size of the BCP files and attachments you plan to retire in the retirement project.

Verify that the staging space is at least four times larger than the size of the data that you plan to retire in a single retirement project.

By default, the retirement job saves the BCP files that it creates for the cluster and pool tables, to the staging space. The retirement job also downloads all attachments that are stored in the SAP database and in a file system to the staging space. For attachments that are stored in the SAP database, the job compresses and downloads the attachments to this staging space. For attachments that are stored in a file system, the job moves the attachments to the staging space.

You can leave the BCP files and attachments in the staging space. Or, you can move the BCP files attachments from the staging space to another file system, external storage, or the Data Vault. Optionally, you can configure the attachment entities to keep the attachments in the file system if you do not want to move attachments from the original file system to the staging space.

Connection Properties

When you create a connection, you define general properties and database properties.

Configure general properties that are relevant for all database connection types, such as the business application and application version that you want to archive data from.

Configure database properties that are relevant to the source database, such as how the archive job connects to the source and how the job archives data from the source. You can also specify properties to improve the archive job performance. The connection properties depend on the connection type that you choose for the database.

When you create an archive project, you select the source connection. The archive job uses the attributes that you defined for the connection.

General Connection Properties

Define general connection properties to set up an archive source connection.

You can configure the following general connection properties:

Connection Name

Archive source name. Use the name to differentiate archive source connections, such as Data Vault Source and Production Source.

When you create an archive project and you assign a source, you select the connection name that you define here. You can use the connection name to display a list of all connection types or filter the list of connections.

Connection Description

Long text description for the connection name. When you manage connections, you can use this field to filter the list of connections.

Connection Type

Database connection type that determines how you connect to the archive source database. The connection type that you choose depends on the database that contains the source data. Choose a compatible connection type for the database and database version. The connection type determines the connection properties that you need to define.

Application Version

Business application and application versions that the database connection type supports. The possible values depend on the database connection type.

You can choose a business application such as Oracle or a custom application for a solution.

For example, you can choose Oracle 10g as the connection type and Oracle Applications 11.5.10 from the Oracle E-Business Suite Product family as the application version.

IBM DB2 Source Connections

When you define an archive source connection for applications on IBM DB2 databases, you can choose from multiple connection types. The connection type that you choose depends on the database version.

Choose one of the following connection types to connect to an IBM DB2 database:

- DB2_Adapter z/OS. Use to connect to IBM DB2 on z/OS.
- DB2_Adapter UDB. Use to connect IBM DB2 on AIX, Linux, UNIX, or Windows (LUX).
- DB2_AS400_Adapter. Use to connect to IBM DB2 on AS/400.

The properties that you configure depend on the connection type that you choose. Some properties are not relevant for all connection types. Property names may vary between connection types.

When you use IBM DB2 native utilities for data movement, the archive job uses the connections that you configure in IBM DB2 Connect to connect to the source and target databases. When you run an archive job, the job gets the DB2 connect connection name from the archive source and target connections. Then, the job uses connection name to get the connection details from DB2 Connect.

To use IBM DB2 native utilities for data movement, create connections to the source and target in IBM DB2 Connect first. Then, create source and target connections in Data Archive and add the DB2 Connect connection name.

Depending on the connection type, you can configure the following source connection properties:

Host

IP address of the source application database server.

Port

Port of the source application database server.

Database or Service Name

Unique identifier or system identifier for the source application database server.

If the database is on AS/400 and includes CHAR for BIT DATA data types, append the following code to the service name:

```
CharsetFor65535=<code page>
```

For example, if your service name is `AU4SNDBX` and your code page is `CP037`. Enter the following code for the service name:

```
AU4SNDBX;CharsetFor65535=CP037
```

Admin Schema or ILM Repository Administrator User

Default administration database user for the source database server, such as SYSTEM.

The administration database user has DBA rights to the database, including the ability to run DDL and access system-level objects.

Admin User Name or Admin Login Name

Login name for the administration database user. This user does not require special permissions as it is only used to connect to the source database. You can provide any user name, such as a database connection user or a read-only user.

Password

Password for the administration login name.

Apps Schema or Application User Name

Database user that owns the application tables that you want archive, such as APPS for Oracle applications.

The production application user is either the owner of the tables that you want to archive or partition, or the user that has full rights to the tables that you want to archive or partition. During the archive process, you can use the production application user to delete from the source. Alternatively, you can configure the source connection to use the staging user to delete data from the source.

Use the default database user for the application or provide another value if you do not use the default user. If you provide a user that does not own or has full access to the application tables, then the archive job fails because the job cannot find the table or view.

The archive job uses the number of schemas in the source application to determine schema names. The archive job uses this attribute value for applications that have one schema, such as PeopleSoft and Siebel. The archive job uses the schema name from the mined application for applications that have multiple schemas, such as Oracle applications.

The ILM engine uses this field to generate the SQL SELECT statement to read data from the application tables. For example, if you enter SYSADMIN, the archive job generates the following SQL:

```
SELECT * from SYSADMIN.<OBJECT NAME>
```

Apps User Name or Application Login Name

Login name that connects to the source database that contains the data you want to archive. This user does not require special permissions as it is only used for the database connection. You can provide any user name, such as a database connection user or a read-only user.

Password

Password for the application login name.

Staging Schema or Staging User Name

Staging database user for the source database.

The staging database user stores interim tables for business rule validation and generates candidates during the archive and partitioning processes. If you enable staging for the archive source connection, the staging user also creates staging tables. The staging tables temporarily store a copy of the archived data before the archive job moves the data to the archive or history database. The staging user typically resides in the source database.

Staging User Name or Staging Login Name

Login name for the staging database.

Password

Password for the staging login name.

Staging Tablespace

Tablespace in the staging database user that stores the interim and staging tables when you run an archive cycle. The archive job always stores the interim tables in this tablespace. The archive job stores staging tables if you enable staging.

Note: For IBM DB2 databases, verify that the staging tablespace is segmented and the tables in each segment are not partitioned.

Index Storage Group or Index Tablespace

Tablespace in the staging database user that stores the interim table indexes when you run an archive cycle.

Use Staging

Determines whether the archive job temporarily stores data in a staging area on the archive source before it moves data to the target. The archive job includes heavy select and delete operations on the tables that are selected in the archive job. The operations may impact performance on any user operations on the tables. For example, if you run an archive job that archives purchase orders, users may see a performance impact when they create a purchase order. The main benefit of staging is to reduce the number of operations on the production tables.

If enabled, the archive job copies data to the staging area after the archive job evaluates the business rules. The archive job copies data that meets the business rules from the archive source to the staging area. Then, the job deletes data from the source and moves data from the staging area to the target system.

If disabled, the job extracts data directly from the archive source to the archive target. The extraction and deletion occur simultaneously. You may want to disable staging if you do not want to create additional storage space on the production database.

If your source data contains the Decfloat data type, do not enable **Use Staging**.

Default is disabled.

JDBC Fetch Size

Controls the number of rows that the system fetches in one batch from one table if you use JDBC to connect to the source database.

Only change the default value if you receive an out-of-memory error when the archive job copies data to the destination. You may get the out-of-memory error if the source table has large objects in each row, such as CLOBS or BLOBS data types. If you receive this error, then reduce the fetch size.

Default is 1000.

Database Link to Production (Restore Only)

Database link name that connects the history database to the production database. This attribute is required when you create the history database as a source and you use transaction or cycle restore. For restore scenarios, the history database is the source, and the production database is the target.

During the restore job, the system uses this attribute for the database link value. If you do not provide a value, then the system dynamically creates the database link from the source to the ILM repository. After the restore job completes, the system drops the database link.

If the system needs to create the database link, then the administration or application database users need to have create and drop database link permissions.

Use Staging User for Deletes

Determines which database user the archive job uses to delete data from the source database.

If enabled, the archive job uses the staging database user to delete from the source.

If disabled, the archive job uses the application database user to delete from the source.

Whether you enable this option depends on the authorizations that the application or staging database users have. For example, you may want to use the staging user for deletes if you have a read-only application database user.

Default is disabled.

Target Attachment Location

Target location for external attachments. Enter the location where you want to archive the external attachments to. You must have write access to the directory.

After the retirement job moves attachments to this location, you can use this location as the final destination that stores the attachments. Or, you can move the attachments from this location to a different file system, external storage, or the Data Vault.

Source/Staging Attachment Location

Source location of the external attachments. Enter the current location where the attachments exist. You must have read access to the directory.

For SAP application retirement, enter the location of the SAP application server file system. The retirement job uses an SAP function module to generate BCP files for data in transparent HR and STXL tables, ADK files, and attachments.

Enter the full path of the location. For example, `\\10.1.10.10\interfaces\CCO\`.

The location must be accessible to the SAP application server and in the same drive where the ILM engine is installed. If the SAP system and Data Archive are on the same operating systems, the path is the same as the Staging Directory property in the Data Vault target connection.

If the SAP system and Data Archive are on different operating systems, then the paths are different.

Staging Script Location

Temporary location that stores the scripts that the archive job generates to move external attachments from the source to the target.

Enter a location that you have read and write access to. For Siebel attachments, enter a location that is accessible to the Data Vault Service for External Attachments server.

This attribute only applies to external attachments.

Move Attachments in Synchronous Mode

Determines whether the archive job automatically archives external attachments or whether you run a standalone job to move the attachments after the archive job completes. If you provide a source file location for attachments, the archive job creates SQL scripts in the file server location and associates the scripts with the archive job ID.

If enabled, the archive job runs the scripts during the archive process. The run procedures configuration in the entity determines when the archive job archives the attachments.

If disabled, you must initiate the movement after the archive job completes. You can manually run the scripts in the file server location or you can run a standalone job to move the attachments. If you run the standalone job, you must provide the archive job ID. The job then looks for the scripts that are associated to the archive job ID.

This attribute only applies to external attachments.

SAP Fetch Size

Number of rows that the retirement job extracts at a time from the SAP cluster and pool tables to write to the BCP file.

Required for SAP application retirement only. Default is 2,000 rows.

SAP Host

Host of the SAP application that you want to retire.

Required for SAP application retirement only.

SAP Client

Client in which the user logs in. Note that all clients in the SAP application are retired.

Required for SAP application retirement only.

SAP System Number

System number in which the user logs in.

Required for SAP application retirement only.

SAP Language

Language in which the user logs in. Note that all languages in the SAP application are retired.

Required for SAP application retirement only.

SAP User

User that logs in to the SAP application. The user must be assigned to the ZINFA_RETIREMENT_PREPARATION role and include RFC connection authorizations.

Required for SAP application retirement only.

SAP User Password

Password for the SAP user.

Required for SAP application retirement only.

Parallel Entity Run

Determines the interim table names in the entity if you want to run the entity in parallel archive jobs. For example, you may extract data in parallel from two different organizations in the same human resource entity. The system creates interim tables in the staging database user. However, the tables need to be differentiated so the archive jobs know which interim tables are relevant for that job.

By default, the system creates interim table names with the same name in EDM. If you enable this attribute, then the system creates the interim table names with the following syntax:

```
<table name>_<job ID>
```

If you enable this attribute and the staging database user does not have authorization to create tables, then the archive job fails.

Default is disabled.

IBM DB2 Connect Data Source Name

Data source name that you configured for the source connection in IBM DB2 Connect. Required if you use IBM DB2 native utilities for data movement. The archive job uses the connections that you configure in IBM DB2 Connect to connect to the source and target databases. When you run an archive job, the job gets the IBM DB2 connect connection name from the archive source and target connections. Then, the job uses the connection name to get the connection details from IBM DB2 Connect.

If you do not configure a data source name and you use IBM DB2 native utilities for data movement, the archive job fails. The job can only connect to the database from IBM DB2 Connect.

Use Imported Schema Name

The name of the schema when you imported the table metadata from the source.

You must select **Use Imported Schema Name** when the tables in the entity are from more than one schema.

FTP User

User name to connect to the FTP server.

Required for SAP application retirement if you set up an FTP connection between the SAP application server and Data Archive.

FTP Password

Password for the FTP user.

Required for SAP application retirement if you set up an FTP connection between the SAP application server and Data Archive.

FTP Host

Host name of the FTP server.

Required for SAP application retirement if you set up an FTP connection between the SAP application server and Data Archive.

FTP Port

Port number of the FTP server. Default port is 21.

When you specify a port number, enable that port number for FTP on the host machine.

Required for SAP application retirement if you set up an FTP connection between the SAP application server and Data Archive.

FTP Folder Location

Name of the FTP folder on the Data Archive server. For example, `ERP\`.

Required for SAP application retirement if you set up an FTP connection between the SAP application server and Data Archive.

Compressed

For SAP application retirement projects, if you want Data Archive to save the BCP files in a compressed file format, enable the check box. When you enable the check box, Data Archive saves the BCP files as `.gz` files in the staging area.

Binding Packages on IBM DB2

Before you can use DataDirect JDBC drivers to connect to IBM DB2 sources, you must bind the packages on the source. To bind packages, use the IBM DB2 Bind Package standalone job.

The IBM DB2 Bind Package standalone job creates a set of catalog and packages on the source. The set of catalog and packages creates an optimized access path for the driver to use when the driver processes IBM DB2 SQL statements. The packages allow the driver to process both static and dynamic SQL on IBM DB2.

Bind Packages Privileges

The user that you provide in the standalone job to bind IBM DB2 packages requires specific privileges.

The IBM DB2 Bind Package job requires an user ID as one of the parameters. The user that you provide runs the bind process on the source. Verify that the user has the required database access.

The following table lists the required authorizations for each platform:

Platform	Authorization
IBM DB2 for z/OS	The user requires one of the following privileges: - SYSADM - SYSCTRL - BINDADD and CREATE IN COLLECTION NULLID
IBM DB2 for Linux, UNIX, and Windows	DBADM
IBM DB2 for AS/400 iSeries	CHANGE authority or higher on the collection NULLID where the driver creates the packages.

IBM DB2 Bind Package Job Parameters

Enter the job parameters when you run the IBM DB2 Bind Package job.

The following table describes the job parameters:

Parameter	Description
DB2 Host Address	Name or IP address of the server that hosts the IBM DB2 source.
DB2 Port Number	Port number of the server that hosts the IBM DB2 source.
DB2 Location/Database Name	Location or name of the database.
User ID	User that logs in to the database and runs the bind process on the source. Choose a user that has the required authorizations.
User Password	Password for the user ID.

Running the IBM DB2 Bind Package Job

Run the IBM DB2 Bind Package standalone job to bind packages on the source.

By default, when you run the job, the driver creates all the packages in the NULLID collection. You must run the job for all platforms and IBM DB2 versions. The job does not impact performance on IBM DB2.

1. Access the DB2 Bind Package job.
2. Enter the job parameters.
3. Schedule the job.
4. View the job log to verify if the binding was successful.

Binding IBM DB2 for z/OS Packages Manually

If the IBM DB2 Bind Package job fails on the IBM DB2 for z/OS source, you can bind the packages manually on the source.

The following message appears if the IBM DB2 Bind Package job fails:

```
[informatica][DB2 JDBC Driver]Bind process is not active.
Please ensure that the user has permissions to create packages.
Packages cannot be created in an XA Connection.
```

To resolve the issue, perform the following steps:

1. Go to the directory where you extracted the Data Archive installation files.
2. Open the `Optional` folder.
3. Extract the `bind_files.zip` file.

   The `bind_files.zip` file contains the `CNTLFILE.XMIT` and `DBRMFILE.XMIT` files.
4. Log in to the source mainframe.
5. Create the following datasets.

   - `<userID>.CNTLFILE.XMIT`
   - `<userID>.DBRMFILE.XMIT`

   For example, if your user ID is `jane12`, then use the following name for the datasets:

   - `JANE12.CNTLFILE.XMIT`
   - `JANE12.DBRMFILE.XMIT`

   Note: Use uppercase for the name of the datasets.
6. Enter the following properties for the datasets: `DSORG=PS, RECFM=FB, LRECL=80, BLKSIZE=3120`
7. Access the source mainframe through FTP from the machine that contains the `CNTLFILE.XMIT` and `DBRMFILE.XMIT` files. Use the same user ID and password for the FTP as the one you used to log in to the source mainframe in step 4.
8. Enter `BIN` to use the binary transfer mode.
9. Enter the `PUT` command to transfer the `CNTLFILE.XMIT` and `DBRMFILE.XMIT` files to the source mainframe.
10. Verify that the data from the `CNTLFILE.XMIT` and `DBRMFILE.XMIT` files transferred to the datasets you created in step 5.
11. Log in to the source mainframe.
12. From the ISPF option 6 or TSO Ready mode, enter the following `RECEIVE` command: `RECEIVE INDS('userid.CNTLFILE.XMIT')`
13. At the prompt, enter `DA('userid.DDJDBC.CNTL')`
14. Press `ENTER`.

    The system creates the partitioned data set `DDJDBC.CNTL` that contains the JCL and distribution packages.
15. From ISPF option 6 or TSO Ready mode, enter the following `RECEIVE` command: `RECEIVE INDS('userid.DBRMFILE.XMIT')`
16. At the prompt, enter `DA('userid.DDJDBC.DBRMLIB')`
17. Press `ENTER`.

    The system creates the partitioned data set `DDJDBC.DBRMLIB` that contains the JCL and distribution packages.

18. Edit the job `userid.DDJDBC.CNTL(BIND)`. Follow the instructions in the job to complete the following tasks:
    a. Bind the packages for the DataDirect Connect for JDBC DB2 driver.
    b. Grant the user the `EXECUTE` privilege.

Validate the Connection Settings

You can validate the connection settings. The Enterprise Data Manager and Data Archive validate the connections settings when you create connections.

Use one of the following tools to validate the connection settings:

- Enterprise Data Manager. The Enterprise Data Manager validates the connection settings when you create a connection to the source database. You use the connection to mine data from the source. You mine databases only if you have a custom application on the database. If you use a prepackaged application, such as PeopleSoft, you do not mine data from the application. Instead, test the connection from Data Archive.
- Data Archive. Data Archive validates the connection settings when you create a source or target connection. You use the source or target connections when you define an archive project.
- JDBC query tool. You can use a third-party tool, such as Toad, SQL Developer, or Aqua Data Studio, to validate the connection settings.

Generic JDBC Source Connections

Use a generic JDBC source connection to connect to a source database using a generic JDBC driver.

To use a generic JDBC source connection, such as a MySQL database, you must first add the JDBC driver JAR files.

After you add the files, restart the ILM application server on the machine that hosts Data Archive. You can then create the source connection and specify the connection properties.

Adding JDBC Driver JAR Files

Add a JDBC driver JAR file so that you can connect to a source database using a generic JDBC driver.

1. Copy the JAR file to the following locations:
    - `<ILM Products Installation Directory>\webapp\WEB-INF\lib`
    - `<ILM Products Installation Directory>\webapp\edm\lib`
2. Restart the ILM application server.
3. Download Enterprise Data Manager.
4. Close all Enterprise Data Manager windows and start the downloaded Enterprise Data Manager.

Generic JDBC Connection Properties

Enter the generic JDBC connection properties so that you can connect to a source database through the JDBC drivers that you uploaded. Set the connection type to GENERIC_JDBC 1.0.

You can configure the following source connection properties:

Driver Name

Name of the JDBC driver.

JDBC URL

URL for the JDBC driver. Use the URL that is supplied with the driver.

Admin Schema Name

Default administration database user for the source database server, such as SYSTEM.

The administration database user has DBA rights to the database, including the ability to run DDL and access system-level objects.

Admin Login Name

Login name for the administration database user. This user does not require special permissions as it is only used to connect to the source database. You can provide any user name, such as a database connection user or a read-only user.

Password

Password for the administration login name.

Application Login Name

Login name that connects to the source database that contains the data you want to archive. This user does not require special permissions as it is only used for the database connection. You can provide any user name, such as a database connection user or a read-only user.

Application User Name

Database user that owns the application tables that you want archive, such as APPS for Oracle applications.

The production application user is either the owner of the tables that you want to archive or partition, or the user that has full rights to the tables that you want to archive or partition. During the archive process, you can use the production application user to delete from the source. Alternatively, you can configure the source connection to use the staging user to delete data from the source.

Use the default database user for the application or provide another value if you do not use the default user. If you provide a user that does not own or has full access to the application tables, then the archive job fails because the job cannot find the table or view.

The archive job uses the number of schemas in the source application to determine schema names. The archive job uses this attribute value for applications that have one schema, such as PeopleSoft and Siebel. The archive job uses the schema name from the mined application for applications that have multiple schemas, such as Oracle applications.

The ILM engine uses this field to generate the SQL SELECT statement to read data from the application tables. For example, if you enter SYSADMIN, the archive job generates the following SQL:

```
SELECT * from SYSADMIN.<OBJECT NAME>
```

Password

Password for the application login name.

Staging User Name

Staging database user for the source database.

The staging database user stores interim tables for business rule validation and generates candidates during the archive and partitioning processes. If you enable staging for the archive source connection, the staging user also creates staging tables. The staging tables temporarily store a copy of the archived data before the archive job moves the data to the archive or history database. The staging user typically resides in the source database.

Staging Login Name

Login name for the staging database.

Password

Password for the staging login name.

Staging Tablespace

Tablespace in the staging database user that stores the interim and staging tables when you run an archive cycle. The archive job always stores the interim tables in this tablespace. The archive job stores staging tables if you enable staging.

Use Staging

Determines whether the archive job temporarily stores data in a staging area on the archive source before it moves data to the target. The archive job includes heavy select and delete operations on the tables that are selected in the archive job. The operations may impact performance on any user operations on the tables. For example, if you run an archive job that archives purchase orders, users may see a performance impact when they create a purchase order. The main benefit of staging is to reduce the number of operations on the production tables.

If enabled, the archive job copies data to the staging area after the archive job evaluates the business rules. The archive job copies data that meets the business rules from the archive source to the staging area. Then, the job deletes data from the source and moves data from the staging area to the target system.

If disabled, the job extracts data directly from the archive source to the archive target. The extraction and deletion occur simultaneously. You may want to disable staging if you do not want to create additional storage space on the production database.

JDBC Fetch Size

Controls the number of rows that the system fetches in one batch from one table if you use JDBC to connect to the source database.

Only change the default value if you receive an out-of-memory error when the archive job copies data to the destination. You may get the out-of-memory error if the source table has large objects in each row, such as CLOBS or BLOBS data types. If you receive this error, then reduce the fetch size.

Default is 1000.

Transactional Commit (Restore Only)

Determines when the system issues commits for restore jobs.

If enabled, the system issues a single commit after it restores all tables.

If disabled, the system issues a commit after it restores each table.

Default is disabled.

Use Staging User for Deletes

Determines which database user the archive job uses to delete data from the source database.

If enabled, the archive job uses the staging database user to delete from the source.

If disabled, the archive job uses the application database user to delete from the source.

Whether you enable this option depends on the authorizations that the application or staging database users have. For example, you may want to use the staging user for deletes if you have a read-only application database user.

Default is disabled.

Target Attachment Location

Target location for external attachments. Enter the location where you want to archive the external attachments to. You must have write access to the directory.

This attribute only applies to external attachments.

Source Attachment Location

Source location of the external attachments. Enter the current location where the attachments exist. You must have read access to the directory.

This attribute only applies to external attachments.

Staging Script Location

Temporary location that stores the scripts that the archive job generates to move external attachments from the source to the target.

Enter a location that you have read and write access to. For Siebel attachments, enter a location that is accessible to the Data Vault Service for External Attachments server.

This attribute only applies to external attachments.

Move Attachments in Synchronous Mode

Determines whether the archive job automatically archives external attachments or whether you run a standalone job to move the attachments after the archive job completes. If you provide a source file location for attachments, the archive job creates SQL scripts in the file server location and associates the scripts with the archive job ID.

If enabled, the archive job runs the scripts during the archive process. The run procedures configuration in the entity determines when the archive job archives the attachments.

If disabled, you must initiate the movement after the archive job completes. You can manually run the scripts in the file server location or you can run a standalone job to move the attachments. If you run the standalone job, you must provide the archive job ID. The job then looks for the scripts that are associated to the archive job ID.

This attribute only applies to external attachments.

Parallel Entity Run

Determines the interim table names in the entity if you want to run the entity in parallel archive jobs. For example, you may extract data in parallel from two different organizations in the same human resource entity. The system creates interim tables in the staging database user. However, the tables need to be differentiated so the archive jobs know which interim tables are relevant for that job.

By default, the system creates interim table names with the same name in EDM. If you enable this attribute, then the system creates the interim table names with the following syntax:

```
<table name>_<job ID>
```

If you enable this attribute and the staging database user does not have authorization to create tables, then the archive job fails.

Create Interim Table with No Logging Mode

Determines if the archive job creates interim tables without logging.

If enabled, the archive job creates the interim tables in no logging mode. You may want to enable if you have a limited amount of logging space in the source. However, if you create tables without logging, then recovery ability is at risk.

If disabled, the archive job creates the interim tables with logging.

Default is disabled.

Maintain Source Compression on Target

When you archive to an Oracle target, determines if the archive job creates target tables with the same compression as the source.

If enabled, the archive job creates target tables with the same compression as the source tables if the source tables have COMPRESS FOR DIRECT_LOAD OPERATIONS table compression.

If disabled, the archive job does not create target tables with compression.

Default is disabled.

Maintain Source Partitions on Target

When you archive to an Oracle target, determines if the archive job creates target tables with the same partitioning as the source.

If enabled, the archive job creates target tables with partitioning if the source tables include list and range partition types.

If disabled, the archive job does not create target tables with partitioning.

Default is disabled.

Use Imported Schema Name

The name of the schema when you imported the table metadata from the source.

You must select **Use Imported Schema Name** when the tables in the entity are from more than one schema.

Informix Source Connections

Define archive source connection properties to connect to applications on Informix databases.

Depending on the connection type, you can configure the following source connection properties:

Host

IP address of the source application database server.

Port

Port of the source application database server.

Server Name

Unique identifier or system identifier for the source application database server.

You must enter both the server name and the database name. Use the following format:

```
<server_name>;databasename=<dbname>
```

ILM Repository Administrator User

Default administration database user for the source database server, such as SYSTEM.

The administration database user has DBA rights to the database, including the ability to run DDL and access system-level objects.

Admin Login Name

Login name for the administration database user. This user does not require special permissions as it is only used to connect to the source database. You can provide any user name, such as a database connection user or a read-only user.

Password

Password for the administration login name.

Application User Name

Database user that owns the application tables that you want archive, such as APPS for Oracle applications.

The production application user is either the owner of the tables that you want to archive or partition, or the user that has full rights to the tables that you want to archive or partition. During the archive process, you can use the production application user to delete from the source. Alternatively, you can configure the source connection to use the staging user to delete data from the source.

Use the default database user for the application or provide another value if you do not use the default user. If you provide a user that does not own or has full access to the application tables, then the archive job fails because the job cannot find the table or view.

The archive job uses the number of schemas in the source application to determine schema names. The archive job uses this attribute value for applications that have one schema, such as PeopleSoft and Siebel. The archive job uses the schema name from the mined application for applications that have multiple schemas, such as Oracle applications.

The ILM engine uses this field to generate the SQL SELECT statement to read data from the application tables. For example, if you enter SYSADMIN, the archive job generates the following SQL:

```
SELECT * from SYSADMIN.<OBJECT NAME>
```

Application Login Name

Login name that connects to the source database that contains the data you want to archive. This user does not require special permissions as it is only used for the database connection. You can provide any user name, such as a database connection user or a read-only user.

Password

Password for the application login name.

Staging User Name

Staging database user for the source database.

The staging database user stores interim tables for business rule validation and generates candidates during the archive and partitioning processes. If you enable staging for the archive source connection, the staging user also creates staging tables. The staging tables temporarily store a copy of the archived data before the archive job moves the data to the archive or history database. The staging user typically resides in the source database.

Staging Login Name

Login name for the staging database.

Password

Password for the staging login name.

Staging Tablespace

Tablespace in the staging database user that stores the interim and staging tables when you run an archive cycle. The archive job always stores the interim tables in this tablespace. The archive job stores staging tables if you enable staging.

Note for Informix sources, Data Archive uses "exclusive" mode locking for staging and interim tables.

Use Copy To Staging

Determines whether the archive job temporarily stores data in a staging area on the archive source before it moves data to the target. The archive job includes heavy select and delete operations on the tables that are selected in the archive job. The operations may impact performance on any user operations on the tables. For example, if you run an archive job that archives purchase orders, users may see a performance impact when they create a purchase order. The main benefit of staging is to reduce the number of operations on the production tables.

If enabled, the archive job copies data to the staging area after the archive job evaluates the business rules. The archive job copies data that meets the business rules from the archive source to the staging area. Then, the job deletes data from the source and moves data from the staging area to the target system.

If disabled, the job extracts data directly from the archive source to the archive target. The extraction and deletion occur simultaneously. You may want to disable staging if you do not want to create additional storage space on the production database.

Enable staging if you want to use the "Delete Commit Interval" option when you create and run an archive and purge project for an Informix source

For Informix sources, Data Archive uses "exclusive" mode locking for staging and interim tables.

JDBC Fetch Size

Controls the number of rows that the system fetches in one batch from one table if you use JDBC to connect to the source database.

Only change the default value if you receive an out-of-memory error when the archive job copies data to the destination. You may get the out-of-memory error if the source table has large objects in each row, such as CLOBS or BLOBS data types. If you receive this error, then reduce the fetch size.

Default is 1000.

Database Link to Production (Restore Only)

Database link name that connects the history database to the production database. This attribute is required when you create the history database as a source and you use transaction or cycle restore. For restore scenarios, the history database is the source, and the production database is the target.

During the restore job, the system uses this attribute for the database link value. If you do not provide a value, then the system dynamically creates the database link from the source to the ILM repository. After the restore job completes, the system drops the database link.

If the system needs to create the database link, then the administration or application database users need to have create and drop database link permissions.

Parallel Entity Run

Determines the interim table names in the entity if you want to run the entity in parallel archive jobs. For example, you may extract data in parallel from two different organizations in the same human resource

entity. The system creates interim tables in the staging database user. However, the tables need to be differentiated so the archive jobs know which interim tables are relevant for that job.

By default, the system creates interim table names with the same name in the EDM. If you enable this attribute, then the system creates the interim table names with the following syntax:

```
<table name>_<job ID>
```

If you enable this attribute and the staging database user does not have authorization to create tables, then the archive job fails.

SAP Host

Host of the SAP application that you want to retire.

Required for SAP application retirement only.

SAP Client

Client in which the user logs in. Note that all clients in the SAP application are retired.

Required for SAP application retirement only.

SAP System Number

System number in which the user logs in.

Required for SAP application retirement only.

SAP Language

Language in which the user logs in. Note that all languages in the SAP application are retired.

Required for SAP application retirement only.

SAP User

User that logs in to the SAP application. The user must be assigned to the ZINFA_RETIREMENT_PREPARATION role and include RFC connection authorizations.

Required for SAP application retirement only.

SAP User Password

Password for the SAP user.

Required for SAP application retirement only.

FTP User

User name to connect to the FTP server.

Required for SAP application retirement if you set up an FTP connection between the SAP application server and Data Archive.

FTP Password

Password for the FTP user.

Required for SAP application retirement if you set up an FTP connection between the SAP application server and Data Archive.

FTP Host

Host name of the FTP server.

Required for SAP application retirement if you set up an FTP connection between the SAP application server and Data Archive.

FTP Port

Port number of the FTP server. Default port is 21.

When you specify a port number, enable that port number for FTP on the host machine.

Required for SAP application retirement if you set up an FTP connection between the SAP application server and Data Archive.

FTP Folder Location

Name of the FTP folder on the Data Archive server. For example, `ERP\`.

Required for SAP application retirement if you set up an FTP connection between the SAP application server and Data Archive.

Compressed

For SAP application retirement projects, if you want Data Archive to save the BCP files in a compressed file format, enable the check box. When you enable the check box, Data Archive saves the BCP files as `.gz` files in the staging area.

Delete Wait Time

Used by the **Delete from Source** step to determine the number of seconds to wait before deleting the next batch in a table and before deleting the next table.

Specify the number of seconds or leave blank.

For example, if you specify 300, Data Archive waits 5 minutes between deleting batches in a table and between deleting tables.

If you do not specify a number or if you enter 0, Data Archive deletes all rows continuously without a time delay.

Select First Max Rows

Determines the number of records, starting from the first record, to archive in the current archive job. Specify the number of rows or leave blank.

For example, if you specify 900, Data Archive copies the first 900 rows from the driving table to the interim table and archives these rows.

If you do not specify a number or if you enter 0, Data Archive copies all the rows from the driving table to the interim table and archives these rows.

INTERVAL Data Types

If the Informix data contains INTERVAL data types, you must use the Informix native driver to load data to Data Vault. Once you have uploaded the driver, run the following commands in the ILM repository:

```
update am_platforms set DRIVER_NAME = 'com.informix.jdbc.IfxDriver' where platform_id =
24
update am_platforms set JDBC_URL = 'jdbc:informix-sqli://$HOST:$PORT/
$DBNAME:INFORMIXSERVER=$SID;DELIMIDENT=Y' where platform_id = 24
```

Then restart the Data Archive server.

Legacy Adapter Connection

Define source connection properties to use legacy adapters, such as iWay, to connect to databases.

You can configure the following source connection properties:

Host

IP address of the source application database server.

Port

Port of the source application database server.

Server

Unique identifier or system identifier for the source application database server.

The default server name for the Legacy Adapter is `EDASERVER`.

ILM Repository Administrator User

Default administration database user for the source database server, such as SYSTEM.

The administration database user has DBA rights to the database, including the ability to run DDL and access system-level objects.

Admin Login Name

Login name for the administration database user. This user does not require special permissions as it is only used to connect to the source database. You can provide any user name, such as a database connection user or a read-only user.

Password

Password for the administration login name.

Application User Name

Database user that owns the application tables that you want archive, such as APPS for Oracle applications.

The production application user is either the owner of the tables that you want to archive or partition, or the user that has full rights to the tables that you want to archive or partition. During the archive process, you can use the production application user to delete from the source. Alternatively, you can configure the source connection to use the staging user to delete data from the source.

Use the default database user for the application or provide another value if you do not use the default user. If you provide a user that does not own or has full access to the application tables, then the archive job fails because the job cannot find the table or view.

The archive job uses the number of schemas in the source application to determine schema names. The archive job uses this attribute value for applications that have one schema, such as PeopleSoft and Siebel. The archive job uses the schema name from the mined application for applications that have multiple schemas, such as Oracle applications.

The ILM engine uses this field to generate the SQL SELECT statement to read data from the application tables. For example, if you enter SYSADMIN, the archive job generates the following SQL:

```
SELECT * from SYSADMIN.<OBJECT NAME>
```

Application Login Name

Login name that connects to the source database that contains the data you want to archive. This user does not require special permissions as it is only used for the database connection. You can provide any user name, such as a database connection user or a read-only user.

Password

Password for the application login name.

Staging User Name

Staging database user for the source database.

The staging database user stores interim tables for business rule validation and generates candidates during the archive and partitioning processes. If you enable staging for the archive source connection, the staging user also creates staging tables. The staging tables temporarily store a copy of the archived data before the archive job moves the data to the archive or history database. The staging user typically resides in the source database.

Staging Login Name

Login name for the staging database.

Password

Password for the staging login name.

Staging Tablespace

Tablespace in the staging database user that stores the interim and staging tables when you run an archive cycle. The archive job always stores the interim tables in this tablespace. The archive job stores staging tables if you enable staging.

Use Staging

Determines whether the archive job temporarily stores data in a staging area on the archive source before it moves data to the target. The archive job includes heavy select and delete operations on the tables that are selected in the archive job. The operations may impact performance on any user operations on the tables. For example, if you run an archive job that archives purchase orders, users may see a performance impact when they create a purchase order. The main benefit of staging is to reduce the number of operations on the production tables.

If enabled, the archive job copies data to the staging area after the archive job evaluates the business rules. The archive job copies data that meets the business rules from the archive source to the staging area. Then, the job deletes data from the source and moves data from the staging area to the target system.

If disabled, the job extracts data directly from the archive source to the archive target. The extraction and deletion occur simultaneously. You may want to disable staging if you do not want to create additional storage space on the production database.

JDBC Fetch Size

Controls the number of rows that the system fetches in one batch from one table if you use JDBC to connect to the source database.

Only change the default value if you receive an out-of-memory error when the archive job copies data to the destination. You may get the out-of-memory error if the source table has large objects in each row, such as CLOBS or BLOBS data types. If you receive this error, then reduce the fetch size.

Default is 1000.

Database Link to Production (Restore Only)

Database link name that connects the history database to the production database. This attribute is required when you create the history database as a source and you use transaction or cycle restore. For restore scenarios, the history database is the source, and the production database is the target.

During the restore job, the system uses this attribute for the database link value. If you do not provide a value, then the system dynamically creates the database link from the source to the ILM repository. After the restore job completes, the system drops the database link.

If the system needs to create the database link, then the administration or application database users need to have create and drop database link permissions.

Parallel Entity Run

Determines the interim table names in the entity if you want to run the entity in parallel archive jobs. For example, you may extract data in parallel from two different organizations in the same human resource entity. The system creates interim tables in the staging database user. However, the tables need to be differentiated so the archive jobs know which interim tables are relevant for that job.

By default, the system creates interim table names with the same name in EDM. If you enable this attribute, then the system creates the interim table names with the following syntax:

```
<table name>_<job ID>
```

If you enable this attribute and the staging database user does not have authorization to create tables, then the archive job fails.

Default is enabled.

Microsoft SQL Server Source Connections

When you define an archive source connection for applications on Microsoft SQL Server databases, you can choose from multiple connection types. The connection type that you choose depends on the database version.

Choose one of the following connection types to connect to a Microsoft SQL Server database:

- Microsoft SQL Server 2000
- Microsoft SQL Server 2005
- Microsoft SQL Server 2008
- Microsoft SQL Server 2012
- Microsoft SQL Server 2014
- Microsoft SQL Server 2016
- Microsoft SQL Server 2017

The properties that you configure depend on the connection type that you choose. Some properties are not relevant for all connection types. Property names may vary between connection types.

Depending on the connection type, you can configure the following source connection properties:

Host

IP address of the source application database server.

Port

Port of the source application database server.

Admin Database

Default administration database for the source database server. Default is master.

Admin Database Owner

Administration database owner that has DBA rights to the database, including the ability to execute DDL and access system-level objects.

Default is dbo.

Admin Login Name

Login name for the administration database user. This user does not require special permissions as it is only used to connect to the source database. You can provide any user name, such as a database connection user or a read-only user.

Password

Password for the administration login name.

Apps Database Owner

Application database owner that owns the application tables you want to archive.

The production application database owner is either the owner of the tables that you want to archive, or is the user that has full rights to the tables that you want to archive. During the archive process, you can use the production database owner to delete from the source. Alternatively, you can use the staging database owner to delete from the source.

Use the default database owner for the application or provide another value if you do not use the default owner. If you provide an owner that does not own or has full access to the application tables, then the archive job fails because the job cannot find the table or view.

The archive job uses the number of schemas in the source application to determine schema names. The archive job uses this attribute value for applications that have one schema. The archive job uses the schema name from the mined application for applications that have multiple schemas.

The ILM engine uses this field to generate the SQL SELECT statement to read data from the application tables.

Default is dbo.

Application Database

Database that contains the application tables that you want to archive.

Application Login Name

Login name that connects to the source database that contains the data you want to archive. This user does not require special permissions as it is only used for the database connection. You can provide any user name, such as a database connection user or a read-only user.

Password

Password for the application login name.

Staging Database

Staging database that stores interim tables for business rule validation and generates candidates during the archive process. The database also stores staging tables if you enable staging for the archive source connection.

Staging Database Owner

Staging database owner that has privileges to create tables in the staging database. The staging database owner creates interim tables. If you enable staging for the archive source connection, the staging owner also creates staging tables. The staging tables temporarily store a copy of the archived data before the archive job moves the data to the archive or history database. The staging owner can delete data from the source. Or, you can use the application owner to delete from the source.

Default is dbo.

Staging Login Name

Login name for the staging database.

Password

Password for the staging login name.

Staging Filegroup

Tablespace in the staging database user that stores the interim and staging tables when you run an archive cycle. The archive job always stores the interim tables in this tablespace. The archive job stores staging tables if you enable staging.

Use Staging

Determines whether the archive job temporarily stores data in a staging area on the archive source before it moves data to the target. The archive job includes heavy select and delete operations on the tables that are selected in the archive job. The operations may impact performance on any user operations on the tables. For example, if you run an archive job that archives purchase orders, users may see a performance impact when they create a purchase order. The main benefit of staging is to reduce the number of operations on the production tables.

If enabled, the archive job copies data to the staging area after the archive job evaluates the business rules. The archive job copies data that meets the business rules from the archive source to the staging area. Then, the job deletes data from the source and moves data from the staging area to the target system.

If disabled, the job extracts data directly from the archive source to the archive target. The extraction and deletion occur simultaneously. You may want to disable staging if you do not want to create additional storage space on the production database.

JDBC Fetch Size

Controls the number of rows that the system fetches in one batch from one table if you use JDBC to connect to the source database.

Only change the default value if you receive an out-of-memory error when the archive job copies data to the destination. You may get the out-of-memory error if the source table has large objects in each row, such as CLOBS or BLOBS data types. If you receive this error, then reduce the fetch size.

Default is 1000.

Linked Server Name to Home or Database Link to ILM Repository

Database link name that connects the source database to the ILM repository. The ILM repository is commonly referred to as the home schema. This attribute is valid if the entity that is included in the archive job has run procedure steps. For example, the entity may have a run procedure step to call an external routine.

When the archive job runs procedures, the system uses this attribute for the database link value. If you do not provide a value, then the system dynamically creates the database link from the source to the ILM repository. After the archive job runs procedures, the system drops the database link.

If the system needs to create the database link, then the administration or application database users need to have create and drop database link permissions.

Linked Server Name to Production (Restore Only) or Database Link to Production (Restore Only)

Database link name that connects the history database to the production database. This attribute is required when you create the history database as a source and you use transaction or cycle restore. For restore scenarios, the history database is the source, and the production database is the target.

During the restore job, the system uses this attribute for the database link value. If you do not provide a value, then the system dynamically creates the database link from the source to the ILM repository. After the restore job completes, the system drops the database link.

If the system needs to create the database link, then the administration or application database users need to have create and drop database link permissions.

Disable Triggers

Determines whether the system disables insert, update, and delete triggers when the archive job deletes rows from tables.

If enabled, the system disables triggers when the archive job deletes data from the source.

If disabled, the system retains the triggers.

Default is enabled.

Target Attachment Location

Target location for external attachments. Enter the location where you want to archive the external attachments to. You must have write access to the directory.

After the retirement job moves attachments to this location, you can use this location as the final destination that stores the attachments. Or, you can move the attachments from this location to a different file system, external storage, or the Data Vault.

Source/Staging Attachment Location

Source location of the external attachments. Enter the current location where the attachments exist. You must have read access to the directory.

For SAP application retirement, enter the location of the SAP application server file system. The retirement job uses an SAP function module to generate BCP files for data in transparent HR and STXL tables, ADK files, and attachments.

Enter the full path of the location. For example, `\\10.1.10.10\interfaces\CCO\`.

The location must be accessible to the SAP application server and in the same drive where the ILM engine is installed. If the SAP system and Data Archive are on the same operating systems, the path is the same as the Staging Directory property in the Data Vault target connection.

If the SAP system and Data Archive are on different operating systems, then the paths are different.

Staging Script Location

Temporary location that stores the scripts that the archive job generates to move external attachments from the source to the target.

Enter a location that you have read and write access to. For Siebel attachments, enter a location that is accessible to the Data Vault Service for External Attachments server.

This attribute only applies to external attachments.

Move Attachments in Synchronous Mode

Determines whether the archive job automatically archives external attachments or whether you run a standalone job to move the attachments after the archive job completes. If you provide a source file location for attachments, the archive job creates SQL scripts in the file server location and associates the scripts with the archive job ID.

If enabled, the archive job runs the scripts during the archive process. The run procedures configuration in the entity determines when the archive job archives the attachments.

If disabled, you must initiate the movement after the archive job completes. You can manually run the scripts in the file server location or you can run a standalone job to move the attachments. If you run the standalone job, you must provide the archive job ID. The job then looks for the scripts that are associated to the archive job ID.

This attribute only applies to external attachments.

Compile ILM Functions

Determines if the system compiles user-defined functions during the archive job.

By default, the staging database user needs privileges to compile user-defined functions on Microsoft SQL Server. The user-defined functions are required to use application metadata to archive data. Application retirement does not use metadata to archive data. Therefore, the staging user does not need any additional privileges besides select permissions, to compile the user-defined functions.

You must clear the **Compile ILM Functions** check box before you start a retirement job on a Microsoft SQL Server source application. This ensures that the staging database user has read-only access and will not be able to modify the source application.

Default is enabled.

Parallel Entity Run

Determines the interim table names in the entity if you want to run the entity in parallel archive jobs. For example, you may extract data in parallel from two different organizations in the same human resource entity. The system creates interim tables in the staging database user. However, the tables need to be differentiated so the archive jobs know which interim tables are relevant for that job.

By default, the system creates interim table names with the same name in EDM. If you enable this attribute, then the system creates the interim table names with the following syntax:

```
<table name>_<job ID>
```

If you enable this attribute and the staging database user does not have authorization to create tables, then the archive job fails.

SAP Fetch Size

Number of rows that the retirement job extracts at a time from the SAP cluster and pool tables to write to the BCP file.

Required for SAP application retirement only. Default is 2,000 rows.

SAP Host

Host of the SAP application that you want to retire.

Required for SAP application retirement only.

SAP Client

Client in which the user logs in. Note that all clients in the SAP application are retired.

Required for SAP application retirement only.

SAP System Number

System number in which the user logs in.

Required for SAP application retirement only.

SAP Language

Language in which the user logs in. Note that all languages in the SAP application are retired.

Required for SAP application retirement only.

SAP User

User that logs in to the SAP application. The user must be assigned to the ZINFA_RETIREMENT_PREPARATION role and include RFC connection authorizations.

Required for SAP application retirement only.

SAP User Password

Password for the SAP user.

Required for SAP application retirement only.

Use Imported Schema Name

The name of the schema when you imported the table metadata from the source.

If you want to archive an entity that contains tables from multiple schemas, you must select the **Use Imported Schema Name** check box. If you do not, the archive job fails at the generate candidates step.

If you are upgrading Data Archive from a version before 6.2 HotFix 2, do not select the check box. Review the Application Database property to confirm that you are connecting to the correct database.

FTP User

User name to connect to the FTP server.

Required for SAP application retirement if you set up an FTP connection between the SAP application server and Data Archive.

FTP Password

Password for the FTP user.

Required for SAP application retirement if you set up an FTP connection between the SAP application server and Data Archive.

FTP Host

Host name of the FTP server.

Required for SAP application retirement if you set up an FTP connection between the SAP application server and Data Archive.

FTP Port

Port number of the FTP server. Default port is 21.

When you specify a port number, enable that port number for FTP on the host machine.

Required for SAP application retirement if you set up an FTP connection between the SAP application server and Data Archive.

FTP Folder Location

Name of the FTP folder on the Data Archive server. For example, `ERP\`.

Required for SAP application retirement if you set up an FTP connection between the SAP application server and Data Archive.

Compressed

For SAP application retirement projects, if you want Data Archive to save the BCP files in a compressed file format, enable the check box. When you enable the check box, Data Archive saves the BCP files as `.gz` files in the staging area.

Creating an SSL Connection

To create an SSL connection between a Microsoft SQL Server database and Data Archive, create or edit the source connection.

1. Log in to the Data Archive UI.
2. If a connection is available, edit the connection through the **Administration** > **Manage Connections** menu.
3. If a connection is not available, click **Administration** > **New Source Connection**.
4. Select a Microsoft SQL Server as the connection type.
5. Enter the connection details to the Microsoft SQL Server.
6. Add the following attributes to the **Port** property:

   ```
   [Database_Port=xxxx];encryptionMethod=SSL;ValidateServerCertificate=false/
   true;CryptoProtocolVersion=TLSv1.2
   ```

   For example:

   ```
   1433;encryptionMethod=SSL;ValidateServerCertificate=false/
   true;CryptoProtocolVersion=TLSv1.2
   ```

   `CryptoProtocolVersion` is the name of the protocol supported by the database server.

   Data Archive uses the **Port** value to create an SSL connection between the Microsoft SQL Server database and Data Archive.
7. Click **Save**.
8. To validate the SSL certificate for the Microsoft SQL Server, you must import the certificate from the SSL enabled Microsoft SQL Server and add it to the respective Java `cacerts` file where Data Archive is running.

MongoDB Source Connections

If you installed the MongoDB accelerator, you can create a source connection to MongoDB. When you configure a MongoDB connection, you give the connection a name, description, and select **MONGODB 1** as the connection type.

Configure the following source connection properties:

Application Version

Name of the MongoDB application version in the Enterprise Data Manager. For example, MongoDB Application.

Host

Host of the MongoDB environment.

Port

Port number of the MongoDB server. The default is 27017.

Database Name

Database name for the source application database server.

Application Login Name

MongoDB user name.

Password

MongoDB password.

Confirm Password

Confirm the MongoDB password.

Netezza Source Connections

Define archive source connection properties to connect to applications on Netezza databases.

A Netezza adapter is available for both Data Archive and the Enterprise Data Manager. When you configure a Netezza connection you must use the Netezza adapter.

Netezza connections do not require specific privileges for database users. Netezza databases have only one database user that is granted all required privileges.

When you create a Netezza source connection, the user name and password for the administration, application, and staging users are identical. When you create a source or target connection, you must select the check box **Parallel Entity Run** and deselect the check box **Use Staging**.

To create an entity on a Netezza source, you must enter a value for all default interim columns in the Enterprise Data Manager. The **Select Clause** field cannot be empty, unlike Oracle.

You can configure the following source connection properties:

Host

IP address of the source application database server.

Port

Port of the source application database server.

Service Name

Database name for the source application database server.

ILM Repository Administrator User

Default administration database user for the source database server. Used to connect to the source database. You can provide any user name, such as database connection user or a read-only user.

Admin Login Name

Login name for the administration database user.

Password

Password for the administration login name.

Application User Name

Database user that owns the application tables that you want archive. The application user name is identical to the administration user name.

Application Login Name

Login name that connects to the source database that contains the data you want to archive. Only used for the database connection.

The application login name is identical to the administration login name.

Password

Password for the application login name.

The password is identical to the administration login password.

Staging User Name

Staging database user for the source database. The staging user name is identical to the administration user name.

The staging database user stores interim tables for business rule validation and generates candidates during the archive process. The staging user typically resides in the source database.

Staging Login Name

Login name for the staging database.

The staging login name is identical to the administration login name.

Password

Password for the staging login name.

The password is identical to the administration login password.

Staging Tablespace

Not applicable. You can enter any value in this field.

Index Tablespace

Not applicable. You can enter any value in this field.

Use Staging

You must clear this check box.

JDBC Fetch Size

Controls the number of rows that the system fetches in one batch from one table if you use JDBC to connect to the source database.

Only change the default value if you receive an out-of-memory error when the archive job copies data to the destination. You may get the out-of-memory error if the source table has large objects in each row, such as CLOBS or BLOBS data types. If you receive this error, then reduce the fetch size.

Default is 1000.

Database Link to Production (Restore Only)

Database link name that connects the history database to the production database. This attribute is required scenarios, the history database is the source, and the production database is the target.

Enter "NONE" in this field because Netezza does not create dblinks.

Use Staging User for Deletes

Not applicable.

Default is disabled.

Target Attachment Location

Target location for external attachments. Enter the location where you want to archive the external attachments to. You must have write access to the directory.

This attribute only applies to external attachments.

Source/Staging Attachment Location

Required to restore archived attachments from the history database or from the Data Vault to the source database. The definition depends on the archived attachment location and whether the attachments are encrypted.

To restore attachments from the history database and non-encrypted attachments from the Data Vault, enter the current location of the archived attachments. This is the location where you originally archived the attachments to. You must have read access to the directory.

To restore encrypted attachments, such as Siebel attachments, from the Data Vault, enter a temporary location that is accessible to the ILM engine and the Data Vault Service for External Attachments component. The restore job moves the attachments from the Data Vault Service AM_ATTACHMENTS table to this temporary location.

Staging Script Location

Required to restore archived attachments from the history database or from the Data Vault to the source database. Enter a temporary location to store the script that the restore job generates.

For attachments in the history database and non-encrypted attachments from the Data Vault, the script moves the attachments from the source attachment location to the target attachment location.

For encrypted attachments in the Data Vault, the script uses the source application encryption utility to encrypt the attachments back into the source application proprietary format. Then, the script moves the attachments from the staging attachment location to the target attachment location.

Move Attachments in Synchronous Mode

Determines whether the archive job automatically archives external attachments or whether you run a standalone job to move the attachments after the archive job completes. If you provide a source file location for attachments, the archive job creates SQL scripts in the file server location and associates the scripts with the archive job ID.

If enabled, the archive job runs the scripts during the archive process. The run procedures configuration in the entity determines when the archive job archives the attachments.

If disabled, you must initiate the movement after the archive job completes. You can manually run the scripts in the file server location or you can run a standalone job to move the attachments. If you run the standalone job, you must provide the archive job ID. The job then looks for the scripts that are associated to the archive job ID.

This attribute only applies to external attachments.

Parallel Entity Run

Determines the interim table names in the entity if you want to run the entity in parallel archive jobs. For example, you may extract data in parallel from two different organizations in the same human resource entity. The system creates interim tables in the staging database user. However, the tables need to be differentiated so the archive jobs know which interim tables are relevant for that job.

By default, the system creates interim table names with the same name in EDM. If you enable this attribute, then the system creates the interim table names with the following syntax:

```
<table name>_<job ID>
```

You must enable this check box.

Oracle Source Connections

When you define an archive source connection for applications on Oracle databases, you can choose from multiple connection types. The connection type that you choose depends on the database version.

Choose one of the following connection types to connect to an Oracle database:

- Oracle 8i
- Oracle 9i
- Oracle 10g
- Oracle 11g
- Oracle 12c
- Oracle 18c

The properties that you configure depend on the connection type that you choose. Some properties are not relevant for all connection types. Property names may vary between connection types.

Depending on the connection type, you can configure the following source connection properties:

Host

IP address of the source application database server. When the source application server is in an Oracle Real Application Cluster, enter the IP address of node1 and the IP address of node 2, separated by a comma.

Port

Port of the source application database server. When the source application server is in an Oracle RAC, enter the port number of node 1 and port number of node 2, separated by a comma.

Service Name

Unique identifier or system identifier for the source application database server.

When you define an Oracle RAC source database server, enter the service name.

Admin Schema Name

Default administration database user for the source database server, such as SYSTEM.

The administration database user has DBA rights to the database, including the ability to run DDL and access system-level objects.

Admin Login Name

Admin Login Name is the same as the schema name.

Password

Password for the administration login name.

Apps Schema Name

Database user that owns the application tables that you want archive, such as APPS for Oracle applications.

The production application user is either the owner of the tables that you want to archive or partition, or the user that has full rights to the tables that you want to archive or partition. During the archive process, you can use the production application user to delete from the source. Alternatively, you can configure the source connection to use the staging user to delete data from the source.

Use the default database user for the application or provide another value if you do not use the default user. If you provide a user that does not own or has full access to the application tables, then the archive job fails because the job cannot find the table or view.

The archive job uses the number of schemas in the source application to determine schema names. The archive job uses this attribute value for applications that have one schema, such as PeopleSoft and Siebel. The archive job uses the schema name from the mined application for applications that have multiple schemas, such as Oracle applications.

The ILM engine uses this field to generate the SQL SELECT statement to read data from the application tables. For example, if you enter SYSADMIN, the archive job generates the following SQL:

```
SELECT * from SYSADMIN.<OBJECT NAME>
```

Application Login Name

Login name that connects to the source database that contains the data you want to archive. This user does not require special permissions as it is only used for the database connection. You can provide any user name, such as a database connection user or a read-only user.

Password

Password for the application login name.

Staging Schema Name

Staging database user for the source database.

The staging database user stores interim tables for business rule validation and generates candidates during the archive and partitioning processes. If you enable staging for the archive source connection, the staging user also creates staging tables. The staging tables temporarily store a copy of the archived data before the archive job moves the data to the archive or history database. The staging user typically resides in the source database.

Staging Login Name

Login name for the staging database.

Password

Password for the staging login name.

Staging Tablespace

Tablespace in the staging database user that stores the interim and staging tables when you run an archive cycle. The archive job always stores the interim tables in this tablespace. The archive job stores staging tables if you enable staging.

Use Staging

Determines whether the archive job temporarily stores data in a staging area on the archive source before it moves data to the target. The archive job includes heavy select and delete operations on the tables that are selected in the archive job. The operations may impact performance on any user operations on the tables. For example, if you run an archive job that archives purchase orders, users may see a performance impact when they create a purchase order. The main benefit of staging is to reduce the number of operations on the production tables.

If enabled, the archive job copies data to the staging area after the archive job evaluates the business rules. The archive job copies data that meets the business rules from the archive source to the staging area. Then, the job deletes data from the source and moves data from the staging area to the target system.

If disabled, the job extracts data directly from the archive source to the archive target. The extraction and deletion occur simultaneously. You may want to disable staging if you do not want to create additional storage space on the production database.

Use Row ID for Delete

Determines how the archive job removes data from the source database. This attribute improves performance when the archive job deletes data from the source.

If enabled, the system uses the row ID and the primary key to delete records from the source database. The system creates temporary tables in the staging database user to store the row ID and the row number or primary key. To create temporary tables in the staging database user, you must enable staging for the connection. This option is the fastest method for deletion.

If disabled, the system deletes records based on the primary key or the WHERE clause from the archive job. If the Enterprise Data Manager has a primary key defined for the table, the system uses the primary key to delete records for the table. If the Enterprise Data Manager does not have a primary key defined for the table, the system uses the same WHERE clause that the archive job generated to extract the archive data from the source.

Default is enabled.

JDBC Fetch Size

Controls the number of rows that the system fetches in one batch from one table if you use JDBC to connect to the source database.

Only change the default value if you receive an out-of-memory error when the archive job copies data to the destination. You may get the out-of-memory error if the source table has large objects in each row, such as CLOBS or BLOBS data types. If you receive this error, then reduce the fetch size.

Default is 1000.

Use Row ID for File Gen

Determines how the archive job generates the file from the source database. If you provide multiple threads to extract from the same table, this attribute improves the archive job performance when the archive job copies data to the destination.

If enabled, the system issues multiple SELECT statements against the same table. The system creates temporary tables in the staging database user to store the row ID and the row number to control what each worker thread processes. The staging database user requires write access or both the create table and create object privileges. Disable this option if you cannot grant the privileges to the staging database user.

If disabled, the system uses one worker thread for each table.

Use Oracle Parallel DML for Delete

Determines if the system uses Oracle or Java workers when the archive job deletes data from the source. This attribute only applies if you use the row ID and primary key for deletion.

If enabled, the system uses the Oracle parallel Data Manipulation Language (DML) for delete operations. Oracle parallel DML locks the source tables during deletion. During the deletion, no other insert or update operations are allowed on the tables. This option may improve performance. You may want to use this option if you archive during system downtime or lower activity periods.

If disabled, the system uses Java workers to delete data from the source. Java workers do not lock the tables during deletion.

Database Link to ILM Repository

Database link name that connects the source database to the ILM repository. The ILM repository is commonly referred to as the home schema. This attribute is valid if the entity that is included in the archive job has run procedure steps. For example, the entity may have a run procedure step to call an external routine.

When the archive job runs procedures, the system uses this attribute for the database link value. If you do not provide a value, then the system dynamically creates the database link from the source to the ILM repository. After the archive job runs procedures, the system drops the database link.

If the system needs to create the database link, then the administration or application database users need to have create and drop database link permissions.

Database Link to Production (Restore Only)

Database link name that connects the history database to the production database. This attribute is required when you create the history database as a source and you use transaction or cycle restore. For restore scenarios, the history database is the source, and the production database is the target.

During the restore job, the system uses this attribute for the database link value. If you do not provide a value, then the system dynamically creates the database link from the source to the ILM repository. After the restore job completes, the system drops the database link.

If the system needs to create the database link, then the administration or application database users need to have create and drop database link permissions.

Transactional Commit (Restore Only)

Determines when the system issues commits for restore jobs.

If enabled, the system issues a single commit after it restores all tables.

If disabled, the system issues a commit after it restores each table.

Default is disabled.

Use Staging User for Deletes

Determines which database user the archive job uses to delete data from the source database.

If enabled, the archive job uses the staging database user to delete from the source.

If disabled, the archive job uses the application database user to delete from the source.

Whether you enable this option depends on the authorizations that the application or staging database users have. For example, you may want to use the staging user for deletes if you have a read-only application database user.

Default is disabled.

Disable Triggers

Determines whether the system disables insert, update, and delete triggers when the archive job deletes rows from tables.

If enabled, the system disables triggers when the archive job deletes data from the source.

If disabled, the system retains the triggers.

Default is enabled.

Target Attachment Location

Target location for external attachments. Enter the location where you want to archive the external attachments to. You must have write access to the directory.

After the retirement job moves attachments to this location, you can use this location as the final destination that stores the attachments. Or, you can move the attachments from this location to a different file system, external storage, or the Data Vault.

Source/Staging Attachment Location

Source location of the external attachments. Enter the current location where the attachments exist. You must have read access to the directory.

For SAP application retirement, enter the location of the SAP application server file system. The retirement job uses an SAP function module to generate BCP files for data in transparent HR and STXL tables, ADK files, and attachments.

Enter the full path of the location. For example, `\\10.1.10.10\interfaces\CCO\`.

The location must be accessible to the SAP application server and in the same drive where the ILM engine is installed. If the SAP system and Data Archive are on the same operating systems, the path is the same as the Staging Directory property in the Data Vault target connection.

If the SAP system and Data Archive are on different operating systems, then the paths are different.

Staging Script Location

Temporary location that stores the scripts that the archive job generates to move external attachments from the source to the target.

Enter a location that you have read and write access to. For Siebel attachments, enter a location that is accessible to the Data Vault Service for External Attachments server.

This attribute only applies to external attachments.

Move Attachments in Synchronous Mode

Determines whether the archive job automatically archives external attachments or whether you run a standalone job to move the attachments after the archive job completes. If you provide a source file location for attachments, the archive job creates SQL scripts in the file server location and associates the scripts with the archive job ID.

If enabled, the archive job runs the scripts during the archive process. The run procedures configuration in the entity determines when the archive job archives the attachments.

If disabled, you must initiate the movement after the archive job completes. You can manually run the scripts in the file server location or you can run a standalone job to move the attachments. If you run the standalone job, you must provide the archive job ID. The job then looks for the scripts that are associated to the archive job ID.

This attribute only applies to external attachments.

Parallel Entity Run

Determines the interim table names in the entity if you want to run the entity in parallel archive jobs. For example, you may extract data in parallel from two different organizations in the same human resource entity. The system creates interim tables in the staging database user. However, the tables need to be differentiated so the archive jobs know which interim tables are relevant for that job.

By default, the system creates interim table names with the same name in EDM. If you enable this attribute, then the system creates the interim table names with the following syntax:

```
<table name>_<job ID>
```

If you enable this attribute and the staging database user does not have authorization to create tables, then the archive job fails.

Create Interim Table with No Logging Mode

Determines if the archive job creates interim tables without logging.

If enabled, the archive job creates the interim tables in no logging mode. You may want to enable if you have a limited amount of logging space in the source. However, if you create tables without logging, then recovery ability is at risk.

If disabled, the archive job creates the interim tables with logging.

Default is disabled.

Maintain Source Compression on Target

When you archive to an Oracle target, determines if the archive job creates target tables with the same compression as the source.

If enabled, the archive job creates target tables with the same compression as the source tables if the source tables have COMPRESS FOR DIRECT_LOAD OPERATIONS table compression.

If disabled, the archive job does not create target tables with compression.

Default is disabled.

Maintain Source Partitions on Target

When you archive to an Oracle target, determines if the archive job creates target tables with the same partitioning as the source.

If enabled, the archive job creates target tables with partitioning if the source tables include list and range partition types.

If disabled, the archive job does not create target tables with partitioning.

Default is disabled.

SAP Fetch Size

Number of rows that the retirement job extracts at a time from the SAP cluster and pool tables to write to the BCP file.

Required for SAP application retirement only. Default is 2,000 rows.

SAP Host

Host of the SAP application that you want to retire.

Required for SAP application retirement only.

SAP Client

Client in which the user logs in. Note that all clients in the SAP application are retired.

Required for SAP application retirement only.

SAP System Number

System number in which the user logs in.

Required for SAP application retirement only.

SAP Language

Language in which the user logs in. Note that all languages in the SAP application are retired.

Required for SAP application retirement only.

SAP User

User that logs in to the SAP application. The user must be assigned to the ZINFA_RETIREMENT_PREPARATION role and include RFC connection authorizations.

Required for SAP application retirement only.

SAP User Password

Password for the SAP user.

Required for SAP application retirement only.

Use Imported Schema Name

The name of the schema when you imported the table metadata from the source.

You must select **Use Imported Schema Name** when the tables in the entity are from more than one schema.

FTP User

User name to connect to the FTP server.

Required for SAP application retirement if you set up an FTP connection between the SAP application server and Data Archive.

FTP Password

Password for the FTP user.

Required for SAP application retirement if you set up an FTP connection between the SAP application server and Data Archive.

FTP Host

Host name of the FTP server.

Required for SAP application retirement if you set up an FTP connection between the SAP application server and Data Archive.

FTP Port

Port number of the FTP server. Default port is 21.

When you specify a port number, enable that port number for FTP on the host machine.

Required for SAP application retirement if you set up an FTP connection between the SAP application server and Data Archive.

FTP Folder Location

Name of the FTP folder on the Data Archive server. For example, `ERP\`.

Required for SAP application retirement if you set up an FTP connection between the SAP application server and Data Archive.

Compressed

For SAP application retirement projects, if you want Data Archive to save the BCP files in a compressed file format, enable the check box. When you enable the check box, Data Archive saves the BCP files as `.gz` files in the staging area.

SSL Enabled

When enabled, creates an SSL connection to the Oracle database. This property is supported for Oracle 11g and later.

ASO Encryption Type

Specifies a list of encryption algorithms used when a client, or another server acting as a client, connects to the Oracle server. The Oracle server uses this list to negotiate a mutually acceptable algorithm with the client. If an algorithm that is not installed is specified, the connection fails with the ORA-12650 error message. This property is supported for Oracle 11g and later.

This property accepts single or multiple comma separated values. For example: AES256,AES192,AES128,3DES112

Based on the negotiation with the server, the Oracle server picks the best algorithm to establish the connection.

ASO Encryption Level

Specifies the data integrity behavior when a client, or another server acting as a client, connects to the Oracle server. The behavior partially depends on the setting used in the source database connection. This property is supported for Oracle 11g and later.

Use one the following values:

- ACCEPTED
- REJECTED
- REQUESTED
- REQUIRED

Default is ACCEPTED.

If you specify a value other than the above values, the connection fails with the following error:

Invalid parameter, use one of ACCEPTED, REJECTED, REQUESTED and REQUIRED

ASO Checksum Type

Specifies a list of data integrity algorithms that when a client, or another server acting as a client, connects to the Oracle server. The list of data integrity algorithms are listed in order of intended use. This list is used to negotiate a mutually acceptable algorithm with the other end of the connection. Each algorithm is checked against the list of available client algorithm types until a match is found. If an algorithm is specified that is not installed, the connection fails with the ORA-12650 error message. This property is supported for Oracle 11g and later. This property accepts single or multiple comma separated values. For example:

MD5, SHA1

ASO Checksum Level

Specifies the data integrity behavior when a client, or another server acting as a client, connects to this server. The behavior partially depends on the setting used in the source database connection. This property is supported for Oracle 11g and later.

Use one the following values:

- ACCEPTED
- REJECTED
- REQUESTED
- REQUIRED

Default is ACCEPTED.

If you specify a value other than the preceding values, the connection fails with the following error:

Invalid parameter, use one of ACCEPTED, REJECTED, REQUESTED and REQUIRED

OID Enabled

Oracle Internet Directory (OID) is an LDAP directory that uses an Oracle database for storage. When you enable OID, you must provide the OID database host name for the Host parameter. You must also provide the service name that specifies the distinguished name of the database, which is configured in OID.

For example:

```
ORA12C,cn=OracleContext,dc=informatica,dc=com
```

This property is available for Oracle 11g and 12c.

PowerExchange ODBC Connections

PowerExchange users can connect to a z/OS source in order to retire non-relational mainframe data such as IMS and VSAM. Data Archive uses a JDBC-ODBC bridge to connect to the z/OS source data through the ODBC drivers that are included with PowerExchange.

You can configure the following source connection properties:

Data Source Name

Unique identifier or system identifier for the source application database server.

ILM Repository Administrator User

Default administration database user for the source database server, such as SYSTEM.

The administration database user has DBA rights to the database, including the ability to run DDL and access system-level objects.

Admin Login Name

Login name for the administration database user. This user does not require special permissions as it is only used to connect to the source database. You can provide any user name, such as a database connection user or a read-only user.

Password

Password for the administration login name.

Application User Name

Database user that owns the application tables that you want archive, such as APPS for Oracle applications.

The production application user is either the owner of the tables that you want to archive or partition, or the user that has full rights to the tables that you want to archive or partition. During the archive process, you can use the production application user to delete from the source. Alternatively, you can configure the source connection to use the staging user to delete data from the source.

Use the default database user for the application or provide another value if you do not use the default user. If you provide a user that does not own or has full access to the application tables, then the archive job fails because the job cannot find the table or view.

The archive job uses the number of schemas in the source application to determine schema names. The archive job uses this attribute value for applications that have one schema, such as PeopleSoft and Siebel. The archive job uses the schema name from the mined application for applications that have multiple schemas, such as Oracle applications.

The ILM engine uses this field to generate the SQL SELECT statement to read data from the application tables. For example, if you enter SYSADMIN, the archive job generates the following SQL:

```
SELECT * from SYSADMIN.<OBJECT NAME>
```

Application Login Name

Login name that connects to the source database that contains the data you want to archive. This user does not require special permissions as it is only used for the database connection. You can provide any user name, such as a database connection user or a read-only user.

Password

Password for the application login name.

Staging User Name

Staging database user for the source database.

The staging database user stores interim tables for business rule validation and generates candidates during the archive and partitioning processes. If you enable staging for the archive source connection, the staging user also creates staging tables. The staging tables temporarily store a copy of the archived data before the archive job moves the data to the archive or history database. The staging user typically resides in the source database.

Staging Login Name

Login name for the staging database.

Password

Password for the staging login name.

Staging Tablespace

Tablespace in the staging database user that stores the interim and staging tables when you run an archive cycle. The archive job always stores the interim tables in this tablespace. The archive job stores staging tables if you enable staging.

Use Staging

Determines whether the archive job temporarily stores data in a staging area on the archive source before it moves data to the target. The archive job includes heavy select and delete operations on the tables that are selected in the archive job. The operations may impact performance on any user operations on the tables. For example, if you run an archive job that archives purchase orders, users may see a performance impact when they create a purchase order. The main benefit of staging is to reduce the number of operations on the production tables.

If enabled, the archive job copies data to the staging area after the archive job evaluates the business rules. The archive job copies data that meets the business rules from the archive source to the staging area. Then, the job deletes data from the source and moves data from the staging area to the target system.

If disabled, the job extracts data directly from the archive source to the archive target. The extraction and deletion occur simultaneously. You may want to disable staging if you do not want to create additional storage space on the production database.

For z/OS data retirement, clear this check box.

JDBC Fetch Size

Controls the number of rows that the system fetches in one batch from one table if you use JDBC to connect to the source database.

Only change the default value if you receive an out-of-memory error when the archive job copies data to the destination. You may get the out-of-memory error if the source table has large objects in each row, such as CLOBS or BLOBS datatypes. If you receive this error, then reduce the fetch size.

Default is 1000.

For z/OS source data retirement, enter 0.

Database Link to ILM Repository

Database link name that connects the source database to the ILM repository. The ILM repository is commonly referred to as the home schema. This attribute is valid if the entity that is included in the archive job has run procedure steps. For example, the entity may have a run procedure step to call an external routine.

When the archive job runs procedures, the system uses this attribute for the database link value. If you do not provide a value, then the system dynamically creates the database link from the source to the ILM repository. After the archive job runs procedures, the system drops the database link.

If the system needs to create the database link, then the administration or application database users need to have create and drop database link permissions.

Database Link to Production (Restore Only)

Database link name that connects the history database to the production database. This attribute is required when you create the history database as a source and you use transaction or cycle restore. For restore scenarios, the history database is the source, and the production database is the target.

During the restore job, the system uses this attribute for the database link value. If you do not provide a value, then the system dynamically creates the database link from the source to the ILM repository. After the restore job completes, the system drops the database link.

If the system needs to create the database link, then the administration or application database users need to have create and drop database link permissions.

Use Staging User for Deletes

Determines which database user the archive job uses to delete data from the source database.

If enabled, the archive job uses the staging database user to delete from the source.

If disabled, the archive job uses the application database user to delete from the source.

Whether you enable this option depends on the authorizations that the application or staging database users have. For example, you may want to use the staging user for deletes if you have a read-only application database user.

Default is disabled.

Salesforce Source Connections

If you have installed the Salesforce accelerator, you can create a source connection to Salesforce. When you configure a Salesforce connection, you give the connection a name, description, and select "Salesforce1" as the connection type.

Configure the following source connection properties:

Application Version

Name of the Salesforce application version in the Enterprise Data Manager. For example, "Salesforce Sales."

Host

Host URL of the Salesforce environment, for example:

- For the Salesforce production environment: login.salesforce.com
  "login.salesforce.com" is the default value.
- For the Salesforce sandbox environment: test.salesforce.com
- For a Salesforce customized URL, enter the customized URL without the "https://". For example: my-domain.salesforce.com

Application Schema

Application schema name. Enter "SFORCE".

"SFORCE" is the default value.

Application Login Name

Salesforce username.

Password

Salesforce password.

Confirm Password

Confirm the Salesforce password.

App Security Token

Salesforce security token. Optional field.

Confirm App Security Token

Confirm the Salesforce security token if applicable.

Interim Schema

Location where you want to create the interim schema table. Enter one of the following values:

- PUBLIC. Data Archive creates the interim table in a location based on the "driver database" parameter. The PUBLIC value is the default value.
- SFORCE. Data Archive creates the interim table in Salesforce. Verify that there is sufficient storage space in Salesforce before you run an archive or purge job. If you choose to create the interim table in Salesforce, performance of the archive or purge job will be affected. The Salesforce JDBC driver takes longer to create the interim table in Salesforce than if you set the value to PUBLIC. Creating the interim table in Salesforce also requires an extra Salesforce API request to insert the data into to the interim table. You must have the necessary permissions in Salesforce to create a table. For more information about required permissions, see the chapter "Salesforce Archiving Administrator Tasks" in the *Data Archive Administrator Guide*.

Driver Database (Name | Path+Name)

When the Salesforce JDBC driver connects to Salesforce, it creates an embedded database for internal use. The driver database parameter specifies the file name prefix, or file absolute path plus the file name

prefix, that the driver uses to create or locate the set of files that define the embedded local database. Enter one of the following values:

- You can leave the parameter blank, if the system user has permission to create files in the location where the ILM application server runs. By default a blank value is taken as "<host>_<application_login_name>." For example, if the host is "login.salesforce.com," and the application login name is "abc," then the default value is "login.salesforce.com_abc." Files will be created with the same prefix. For example, "login.salesforce.com_abc.config."
- Enter any name for the database. The driver creates a local database with the given name. For example, if you enter the name "Test," the driver creates a local database with file names such as "Test.config" and "Test.properties" in the location where the ILM application server runs.
- Enter an absolute file path plus a database name. The driver creates a local database with the given database name in the given path. For example, if you enter the value as "C:\ILM_DA\Test," the driver creates a local database with file names such as "Test.config" and "Test.properties" under the path "C:\ILM_DA."

The driver database name must be unique to the source connection; different source connections should not use the same driver database name.

Include Salesforce Archived Records

Select this check box if you want to archive records that have already been archived in Salesforce.

Salesforce has an internal archive functionality that archives eligible records from Salesforce objects. This feature is enabled by default for both the Task and Event entities. If you want to archive or purge the Task and Event entities, you must select this check box in the source connection.

Sybase Source Connections

Define archive source connection properties to connect to applications on Sybase databases.

You can configure the following source connection properties:

Host

IP address of the source application database server.

Port

Port of the source application database server.

Service Name

Unique identifier or system identifier for the source application database server.

ILM Repository Administrator User

Default administration database user for the source database server, such as SYSTEM.

The administration database user has DBA rights to the database, including the ability to run DDL and access system-level objects.

Admin Login Name

Login name for the administration database user. This user does not require special permissions as it is only used to connect to the source database. You can provide any user name, such as a database connection user or a read-only user.

Password

Password for the administration login name.

Application User Name

Database user that owns the application tables that you want archive, such as APPS for Oracle applications.

The production application user is either the owner of the tables that you want to archive or partition, or the user that has full rights to the tables that you want to archive or partition. During the archive process, you can use the production application user to delete from the source. Alternatively, you can configure the source connection to use the staging user to delete data from the source.

Use the default database user for the application or provide another value if you do not use the default user. If you provide a user that does not own or has full access to the application tables, then the archive job fails because the job cannot find the table or view.

The archive job uses the number of schemas in the source application to determine schema names. The archive job uses this attribute value for applications that have one schema, such as PeopleSoft and Siebel. The archive job uses the schema name from the mined application for applications that have multiple schemas, such as Oracle applications.

The ILM engine uses this field to generate the SQL SELECT statement to read data from the application tables. For example, if you enter SYSADMIN, the archive job generates the following SQL:

```
SELECT * from SYSADMIN.<OBJECT NAME>
```

Application Login Name

Login name that connects to the source database that contains the data you want to archive. This user does not require special permissions as it is only used for the database connection. You can provide any user name, such as a database connection user or a read-only user.

Password

Password for the application login name.

Staging User Name

Staging database user for the source database.

The staging database user stores interim tables for business rule validation and generates candidates during the archive and partitioning processes. If you enable staging for the archive source connection, the staging user also creates staging tables. The staging tables temporarily store a copy of the archived data before the archive job moves the data to the archive or history database. The staging user typically resides in the source database.

Staging Login Name

Login name for the staging database.

Password

Password for the staging login name.

Staging Tablespace

Tablespace in the staging database user that stores the interim and staging tables when you run an archive cycle. The archive job always stores the interim tables in this tablespace. The archive job stores staging tables if you enable staging.

Use Staging

Determines whether the archive job temporarily stores data in a staging area on the archive source before it moves data to the target. The archive job includes heavy select and delete operations on the

tables that are selected in the archive job. The operations may impact performance on any user operations on the tables. For example, if you run an archive job that archives purchase orders, users may see a performance impact when they create a purchase order. The main benefit of staging is to reduce the number of operations on the production tables.

If enabled, the archive job copies data to the staging area after the archive job evaluates the business rules. The archive job copies data that meets the business rules from the archive source to the staging area. Then, the job deletes data from the source and moves data from the staging area to the target system.

If disabled, the job extracts data directly from the archive source to the archive target. The extraction and deletion occur simultaneously. You may want to disable staging if you do not want to create additional storage space on the production database.

JDBC Fetch Size

Controls the number of rows that the system fetches in one batch from one table if you use JDBC to connect to the source database.

Only change the default value if you receive an out-of-memory error when the archive job copies data to the destination. You may get the out-of-memory error if the source table has large objects in each row, such as CLOBS or BLOBS data types. If you receive this error, then reduce the fetch size.

Default is 1000.

Database Link to Production (Restore Only)

Database link name that connects the history database to the production database. This attribute is required when you create the history database as a source and you use transaction or cycle restore. For restore scenarios, the history database is the source, and the production database is the target.

During the restore job, the system uses this attribute for the database link value. If you do not provide a value, then the system dynamically creates the database link from the source to the ILM repository. After the restore job completes, the system drops the database link.

If the system needs to create the database link, then the administration or application database users need to have create and drop database link permissions.

Parallel Entity Run

Determines the interim table names in the entity if you want to run the entity in parallel archive jobs. For example, you may extract data in parallel from two different organizations in the same human resource entity. The system creates interim tables in the staging database user. However, the tables need to be differentiated so the archive jobs know which interim tables are relevant for that job.

By default, the system creates interim table names with the same name in EDM. If you enable this attribute, then the system creates the interim table names with the following syntax:

```
<table name>_<job ID>
```

If you enable this attribute and the staging database user does not have authorization to create tables, then the archive job fails.

SAP Host

Host of the SAP application that you want to retire.

Required for SAP application retirement only.

SAP Client

Client in which the user logs in. Note that all clients in the SAP application are retired.

Required for SAP application retirement only.

SAP System Number

System number in which the user logs in.

Required for SAP application retirement only.

SAP Language

Language in which the user logs in. Note that all languages in the SAP application are retired.

Required for SAP application retirement only.

SAP User

User that logs in to the SAP application. The user must be assigned to the ZINFA_RETIREMENT_PREPARATION role and include RFC connection authorizations.

Required for SAP application retirement only.

SAP User Password

Password for the SAP user.

Required for SAP application retirement only.

Use Imported Schema Name

The name of the schema when you imported the table metadata from the source.

You must select **Use Imported Schema Name** when the tables in the entity are from more than one schema.

FTP User

User name to connect to the FTP server.

Required for SAP application retirement if you set up an FTP connection between the SAP application server and Data Archive.

FTP Password

Password for the FTP user.

Required for SAP application retirement if you set up an FTP connection between the SAP application server and Data Archive.

FTP Host

Host name of the FTP server.

Required for SAP application retirement if you set up an FTP connection between the SAP application server and Data Archive.

FTP Port

Port number of the FTP server. Default port is 21.

When you specify a port number, enable that port number for FTP on the host machine.

Required for SAP application retirement if you set up an FTP connection between the SAP application server and Data Archive.

FTP Folder Location

Name of the FTP folder on the Data Archive server. For example, `ERP\`.

Required for SAP application retirement if you set up an FTP connection between the SAP application server and Data Archive.

Compressed

For SAP application retirement projects, if you want Data Archive to save the BCP files in a compressed file format, enable the check box. When you enable the check box, Data Archive saves the BCP files as `.gz` files in the staging area.

Teradata Source Connections

Define archive source connection properties to connect to applications on Teradata databases.

You can configure the following source connection properties:

Host

IP address of the source application database server.

Database Name

Unique identifier or system identifier for the source application database server.

Admin Schema Name

Default administration database user for the source database server, such as SYSTEM.

The administration database user has DBA rights to the database, including the ability to run DDL and access system-level objects.

Admin Login Name

Login name for the administration database user. This user does not require special permissions as it is only used to connect to the source database. You can provide any user name, such as a database connection user or a read-only user.

Password

Password for the administration login name.

Apps Schema Name

Database user that owns the application tables that you want archive, such as APPS for Oracle applications.

The production application user is either the owner of the tables that you want to archive or partition, or the user that has full rights to the tables that you want to archive or partition. During the archive process, you can use the production application user to delete from the source. Alternatively, you can configure the source connection to use the staging user to delete data from the source.

Use the default database user for the application or provide another value if you do not use the default user. If you provide a user that does not own or has full access to the application tables, then the archive job fails because the job cannot find the table or view.

The archive job uses the number of schemas in the source application to determine schema names. The archive job uses this attribute value for applications that have one schema, such as PeopleSoft and Siebel. The archive job uses the schema name from the mined application for applications that have multiple schemas, such as Oracle applications.

The ILM engine uses this field to generate the SQL SELECT statement to read data from the application tables. For example, if you enter SYSADMIN, the archive job generates the following SQL:

```
SELECT * from SYSADMIN.<OBJECT NAME>
```

Application Login Name

Login name that connects to the source database that contains the data you want to archive. This user does not require special permissions as it is only used for the database connection. You can provide any user name, such as a database connection user or a read-only user.

Password

Password for the application login name.

Staging Schema Name

Staging database user for the source database.

The staging database user stores interim tables for business rule validation and generates candidates during the archive and partitioning processes. If you enable staging for the archive source connection, the staging user also creates staging tables. The staging tables temporarily store a copy of the archived data before the archive job moves the data to the archive or history database. The staging user typically resides in the source database.

Staging Login Name

Login name for the staging database.

Password

Password for the staging login name.

Use Staging

Determines whether the archive job temporarily stores data in a staging area on the archive source before it moves data to the target. The archive job includes heavy select and delete operations on the tables that are selected in the archive job. The operations may impact performance on any user operations on the tables. For example, if you run an archive job that archives purchase orders, users may see a performance impact when they create a purchase order. The main benefit of staging is to reduce the number of operations on the production tables.

If enabled, the archive job copies data to the staging area after the archive job evaluates the business rules. The archive job copies data that meets the business rules from the archive source to the staging area. Then, the job deletes data from the source and moves data from the staging area to the target system.

If disabled, the job extracts data directly from the archive source to the archive target. The extraction and deletion occur simultaneously. You may want to disable staging if you do not want to create additional storage space on the production database.

JDBC Fetch Size

Controls the number of rows that the system fetches in one batch from one table if you use JDBC to connect to the source database.

Only change the default value if you receive an out-of-memory error when the archive job copies data to the destination. You may get the out-of-memory error if the source table has large objects in each row, such as CLOBS or BLOBS data types. If you receive this error, then reduce the fetch size.

Default is 1000.

Maximum Parallel Sessions

Determines the maximum amount of parallel session requests that jobs can pass to Teradata Parallel Transporter or the Teradata JDBC FastExport when you export data from Teradata sources. Use the average number of Teradata Access Module Processors (AMPs) to determine the maximum parallel sessions to configure.

Default is 1.

Parallel Entity Run

Determines the interim table names in the entity if you want to run the entity in parallel archive jobs. For example, you may extract data in parallel from two different organizations in the same human resource entity. The system creates interim tables in the staging database user. However, the tables need to be differentiated so the archive jobs know which interim tables are relevant for that job.

By default, the system creates interim table names with the same name in EDM. If you enable this attribute, then the system creates the interim table names with the following syntax:

```
<table name>_<job ID>
```

If you enable this attribute and the staging database user does not have authorization to create tables, then the archive job fails.

Default is enabled.

Use Teradata Parallel Transporter

Determines the Teradata native utility that the archive or retirement job uses to export data from Teradata to the Data Vault.

If enabled, the job uses Teradata Parallel Transporter to export data. Teradata Parallel Transporter cannot process tables that include binary and large object data types, such as BLOB and CLOB. The job uses Teradata JDBC FastExport to process tables with binary and large object data types. If the length of all column data types exceeds 8000 characters, the job uses JDBC FastExport.

If disabled, the job uses Teradata JDBC FastExport to export data from all tables. Disable for source connections in restore jobs. Teradata Parallel Transporter is not available for restore.

Teradata Parallel Transporter is not supported in non-bulk mode.

Default is disabled.

Template Name (Teradata Parallel Transporter)

Name of the script template that the archive or retirement job uses to create script files when the job uses Teradata Parallel Transporter for data export. The name is the API package name in the Enterprise Data Manager.

The job uses the script template to create one script for each table the Teradata Parallel Transporter exports data from. During the script creation, the job replaces the parameters with table-specific values. The job uses the script to pass the query request to Teradata Parallel Transporter. Teradata Parallel Transporter rewrites the query requests to optimize the query for maximum performance and throughput.

Edit only if you customized a script template. The default script template is read-only. If you customize the script template, you must copy the script template and modify the copied version. Enter the name of the customized script template.

Default is TPT_EXTRACT_SCRIPT_TEMPLATE.

Compress File (Teradata Parallel Transporter)

Determines if Teradata Parallel Transporter compresses the BCP files that the utility creates.

If enabled, Teradata Parallel Transporter creates compressed, zipped BCP files. You may want to enable the compression if you have a limited amount of staging space for the BCP files. If you enable BCP file compression, the performance of the archive or retirement job may be impacted. Teradata Parallel Transporter creates one BCP file for each table. The time it takes to create the BCP file depends on the size of the data in the table. The compression time increases as the table size increases.

If disabled, Teradata Parallel Transporter creates uncompressed BCP files. Disable to optimize the performance of the archive or retirement job. You may want to disable the compression if you do not have a space limitation for the BCP file staging area.

Default is disabled.

Use Imported Schema Name

The name of the schema when you imported the table metadata from the source.

You must select **Use Imported Schema Name** when the tables in the entity are from more than one schema.

User-defined data types are not supported for Teradata source connections. The PERIOD (TIME WITH TIMEZONE) data type is not supported in the restore functionality.

Creating a Source Connection

Define source connections before you create an archive project.

Before you create an archive source connection, verify that the staging database user exists on the source database and the database user includes the required user privileges. The staging database user is required even if you do not want to use a staging area for archiving.

1. Click **Administration** > **New Source Connection**.
2. Enter the general connection properties.

   After you choose the connection type, the system dynamically refreshes the screen with the properties that are relevant for the connection type.
3. Enter the database-specific connection properties.

   For more information, see the connection properties for the source database.
4. Click **Save**.

   The system validates the database connection. For SAP application retirement, the retirement job validates the login details to the SAP application.

Copying a Source Connection

You can copy a connection if the property configuration is similar to the one that you need to create.

When you copy an existing connection, the system creates a connection. The system renames the connection to `Copy of <connection name>`. You can edit the connection name to accurately reflect the connection. The connection includes the same configuration settings and attribute values from the connection that you copied. You can keep or change the copied values.

1. Click **Administration** > **Manage Connections**.

2. Click the **Copy** icon next to the connection that you want to copy.

   The **New/Edit Archive Source** page appears with the default property values from the copied connection.
3. Optionally, change the connection name and connection property values.
4. Click **Save.**

Database User Password Changes

Your company security policy may require that you change database user passwords frequently. When you change database user passwords, you must update connections in the ILM repository.

You can edit the connection manually. Or, you can use an SQL script to update passwords in the ILM repository without logging in to the ILM application. Use the script to change the password for the administration, production application, and staging database users and for the optimized file target. The script is available as a batch file or a shell file and is located in the optional directory of the ILM installation.

The script provides a more efficient option to update the ILM repository when company policy requires you to update database user passwords on a regular basis. For example, you must change database user passwords every 20 days.

Properties File

The script uses values in a properties file to connect to the ILM repository and to change the passwords. You create and configure the properties file before you run the script.

The following table describes the required properties for the properties file:

Property	Description
amhome.connection.host	ILM repository host name.
amhome.connection.port	ILM repository port number.
amhome.connection.sid	ILM repository database service name or database name.
amhome.connection.username	ILM repository user name.
amhome.connection.password	ILM repository user password.
amhome.connection.schemaname	Unique database placeholder for tables.
amhome.connection.validate.sql	Script that determines if the installation repository is the ILM repository.
amhome.values.separator	Symbol that indicates the separation of multiple values for the other properties. Use a unique value that is not used in the database names or passwords. For example, use `@@`.

Property	Description
amhome.rep.names	ILM repository database schema name for the Admin, Apps, or Staging schemas. Use the same name as the **Schema Name** field in the Edit Archive Source or Edit Archive Target menu. To enter multiple values, use the separator that you defined for the amhome.values.separator property. For example, enter `SYSTEM@@STAGING` to change the password for the SYSTEM and STAGING database schemas.
amhome.rep.usernames	Login name for the database schema or the Data Vault user. Use the same name as the **Login Name** field in the Edit Archive Source or Edit Archive Target menu. To enter multiple values, use the separator that you defined for the amhome.values.separator property.
amhome.rep.newpasswords	New password for the login name. To enter multiple values, use the separator that you defined for the amhome.values.separator property.

Changing Database User Passwords

Configure the ILM repository connection details and the database user details in a properties file. The script uses the properties file to connect to the ILM repository and to validate the database user login name and password. If the database user login details are valid, the script updates the ILM repository with the new password.

The ILM application server does not need to be running when you run the script.

Note: If a Data Vault user password expires, you can update the ILM repository with a new password only if folders have been created in Data Vault with the Create Archive Folder job.

1. Navigate to the following directory:

   `<ILM installation directory>/optional`
2. Use the `PasswordChangesSampleProperty` file to create a property file.
3. Add the required properties to the property file.
4. Save and close the property file.
5. From the same directory, run one of the following files:
   - `PasswordChangeEnv.bat` for Windows.
   - `PasswordChangeEnv.sh` for Linux and UNIX.
6. Enter the location of the property file that you created.

   The script executes and creates the `PasswordChangeLog` log file in the same directory.
7. View the log file to show the script outcome.

CHAPTER 5

Target Connections

This chapter includes the following topics:

Target Connections Overview

Create a target connection for each database or storage type to which you want to archive data. The target connection can point to a different database, such as an archive or history database, or the connection can point to the Data Vault.

If you installed the Data Visualization component, the target connection is the connection that Data Archive uses to retrieve data for reports. When you create a connection to the Data Vault, Data Archive uses the target connection information to create a data source. If you change the target connection information, Data Archive updates the data source.

For restore jobs, the target database is the production database that you want to restore data to. If you want to restore archived data, create a new target connection for the source or production database. In the restore scenario, the source database becomes the archive target.

You can create target connections to the following databases:

- IBM DB2
- Informix

- JDBC
- Microsoft SQL Server
- Netezza
- ODBC
- Data Vault
- Oracle
- Sybase ASE
- Teradata

Note: If a target connection is used as part of File Archive or retirement definitions, you cannot modify the application version.

Target Database Requirements

Verify the target database requirements when you archive to another database. The target database is the location where you archive data to and is typically referred to as the history or archive database. You can verify the requirements before or after the installation.

The target database must be on the same platform and version as the source database. For example, if you archive from an Oracle 10g database, the target database must be Oracle 10g.

Create the following history tablespaces when you archive to another database:

- History data tablespace. For example, AM_HISTORY_D. The history data tablespace stores the archived data.
- History index tablespace. For example, AM_HISTORY_X. The history index tablespace stores the associated indexes of the archived data.

The history tablespace requirements depend on the size of the data and indexes in the archive job. Configure the tablespace size to at least the same size of the source instance to store all data and the associated indexes to be archived.

Connection Properties

When you create a connection, you define general properties and database-specific properties. You define the connection properties once.

Configure general properties that are relevant for all database connection types, such as the business application and application version that you want to archive data to.

Configure database properties that are relevant to the target database, such as how the archive job connects to the target. The connection properties depend on the connection type that you choose for the database.

When you create an archive project, you select the target connection. The archive job uses the attributes that you defined for the connection. Define target connections before you create an archive project.

General Connection Properties

Define general connection properties to set up an archive target connection.

You can configure the following general connection properties:

Connection Name

Archive target name. Use the name to differentiate archive target connections.

When you create an archive project and you assign a target, you select the connection name that you define here. You can use the connection name to display a list of all connection types or filter the list of connections by name.

Connection Description

Long text description for the connection name. When you manage connections, you can use this field to filter the list of connections.

Connection Type

Database connection type that determines how you connect to the archive target database. The connection type that you choose depends on the database that you want archive the data to. Choose a compatible connection type for the database and database version. The connection type determines which database-specific connection properties you need to define.

Application Version

Business application and application versions that the database connection type supports. The possible values depend on the database connection type.

You can choose a business application such as Oracle or an optimized file.

IBM DB2 Target Connections

When you define an archive target connection for IBM DB2 databases, you can choose from multiple connection types. The connection type you choose depends on the database version.

Choose one of the following connection types to connect to an IBM DB2 database:

- DB2_Adapter z/OS. Use to connect to IBM DB2 on z/OS.
- DB2_Adapter UDB. Use to connect IBM DB2 on AIX, Linux, UNIX, or Windows (LUX).
- DB2_AS400_Adapter. Use to connect to IBM DB2 on AS/400.

The properties that are available are based on the connection type that you choose. Also, the property names may vary slightly.

When you use IBM DB2 native utilities for data movement, the archive job uses the connections that you configure in IBM DB2 Connect to connect to the source and target databases. When you run an archive job, the job gets the DB2 connect connection name from the archive source and target connections. Then, the job uses connection name to get the connection details from DB2 Connect.

To use IBM DB2 native utilities for data movement, create connections to the source and target in IBM DB2 Connect first. Then, create source and target connections in Data Archive and add the DB2 Connect connection name.

Depending on the connection type, you can configure the following target connection properties:

Host

Host of the target application database server.

Port

Port of the target application database server.

Database or Service Name

Unique identifier or system identifier for the target database server.

Admin Schema or ILM Repository Administrator User

Default administration database user for the target database server, such as SYSTEM.

Admin User Name or Admin Login Name

Login name for the administration database server. This user does not require special permissions as it is only used to connect to the target database. You can provide any user name, such as a database connection user or a read-only user.

Password

Password for the administration login name.

Apps Schema or Application User Name

Application database user that owns the tables that you want to archive to. Also commonly referred to as the history application user.

Apps User Name or Application Login Name

Login name for the target application database user.

Password

Password for the application login name.

Data Tablespace

Tablespace name in the application database user that stores the history tables when you run an archive cycle. The archive job always stores the history tables in this tablespace. The archive job only stores staging tables in this tablespace if you configure the source to use staging.

Index Tablespace

Index tablespace name for the history tables.

Federated Server Name to Source or Database Link to Source

Database link name that connects the target database to the source database. The archive job uses the database link when the job copies data to the destination.

If you do not provide a value, then the archive job uses JDBC to move the data.

Target Attachment Location

Required to restore archived attachments from the history database or from the Data Vault to the source database.

Enter the target directory where you want to restore the external attachments. The directory is the location where the source application reads the attachments from. You must have write access to the directory.

Source/Staging Attachment Location

Required to restore archived attachments from the history database or from the Data Vault to the source database. The definition depends on the archived attachment location and whether the attachments are encrypted.

To restore attachments from the history database and non-encrypted attachments from the Data Vault, enter the current location of the archived attachments. This is the location where you originally archived the attachments to. You must have read access to the directory.

To restore encrypted attachments, such as Siebel attachments, from the Data Vault, enter a temporary location that is accessible to the ILM engine and the Data Vault Service for External Attachments component. The restore job moves the attachments from the Data Vault Service AM_ATTACHMENTS table to this temporary location.

Staging Script Location

Required to restore archived attachments from the history database or from the Data Vault to the source database. Enter a temporary location to store the script that the restore job generates.

For attachments in the history database and non-encrypted attachments from the Data Vault, the script moves the attachments from the source attachment location to the target attachment location.

For encrypted attachments in the Data Vault, the script uses the source application encryption utility to encrypt the attachments back into the source application proprietary format. Then, the script moves the attachments from the staging attachment location to the target attachment location.

Add-On URL

Required to restore encrypted attachments from the Data Vault to the source database. Enter the URL for the Data Vault Service for External Attachments. The Data Vault Service for External Attachments converts external attachments from the archived format into the source proprietary format.

Drop Indexes

Controls whether the archive job drops the index on the history database before it inserts data in the history database. Use to improve performance when the archive job copies data to the destination.

IBM DB2 Connect Data Source Name

Data source name that you configured for the target connection in IBM DB2 Connect. Required if you use IBM DB2 native utilities for data movement. The archive job uses the connections that you configure in IBM DB2 Connect to connect to the source and target databases. When you run an archive job, the job gets the IBM DB2 connect connection name from the archive source and target connections. Then, the job uses the connection name to get the connection details from IBM DB2 Connect.

If you do not configure a data source name and you use IBM DB2 native utilities for data movement, the archive job fails. The job can only connect to the database from IBM DB2 Connect.

Informix Target Connections

Define archive target connection properties to archive data to Informix databases.

You can configure the following target connection properties:

Host

Host of the target application database server.

Port

Port of the target application database server.

Server Name

Unique identifier or system identifier for the target database server.

You must enter both the server name and the database name. Use the following format:

```
<server_name>;databasename=<dbname>
```

ILM Repository Administrator User

Default administration database user for the target database server, such as SYSTEM.

Password

Password for the administration login name.

Application User Name

Application database user that owns the tables that you want to archive to. Also commonly referred to as the history application user.

Password

Password for the application login name.

Data Tablespace

Tablespace name in the application database user that stores the history tables when you run an archive cycle. The archive job always stores the history tables in this tablespace. The archive job only stores staging tables in this tablespace if you configure the source to use staging.

Index Tablespace

Index tablespace name for the history tables.

Database Link to Source

Database link name that connects the target database to the source database. The archive job uses the database link when the job copies data to the destination.

If you do not provide a value, then the archive job uses JDBC to move the data.

Drop Destination Indexes

Controls whether the archive job drops the index on the history database before it inserts data in the history database. Use to improve performance when the archive job copies data to the destination.

Microsoft SQL Server Target Connections

When you define an archive target connection for Microsoft SQL Server databases, you can choose from multiple connection types. The connection type you choose depends on the database version.

Choose one of the following connection types to connect to a Microsoft SQL Server database:

- Microsoft SQL Server 2000
- Microsoft SQL Server 2005
- Microsoft SQL Server 2008
- Microsoft SQL Server 2012
- Microsoft SQL Server 2014
- Microsoft SQL Server 2016
- Microsoft SQL Server 2017

The properties that you configure depend on the connection type that you choose. Some properties are not relevant for all connection types. Property names may vary between connection types.

Depending on the connection type, you can configure the following target connection properties:

Host

Host of the target application database server.

Port

Port of the target application database server.

Admin Database

Default administration database for the target database server. Default is master.

Admin Database Owner

Administration database owner that has DBA rights to the database, including the ability to execute DDL and access system-level objects.

Default is dbo.

Admin Login Name

Login name for the administration database server. This user does not require special permissions as it is only used to connect to the target database. You can provide any user name, such as a database connection user or a read-only user.

Password

Password for the administration login name.

Application Database

Database that contains the application tables that you want to archive to.

Apps Database Owner

Application database owner that owns the application tables you want to archive to. Default is dbo.

Application Login Name

Login name for the target application database user.

Password

Password for the application login name.

Data Filegroup

Tablespace name in the application database user that stores the history tables when you run an archive cycle. The archive job always stores the history tables in this tablespace. The archive job only stores staging tables in this tablespace if you configure the source to use staging.

Index Filegroup

Index tablespace name for the history tables.

Database Link to Source or Linked Server Name to Source

Database link name that connects the target database to the source database. The archive job uses the database link when the job copies data to the destination.

If you do not provide a value, then the archive job uses JDBC to move the data.

Drop Destination Indexes

Controls whether the archive job drops the index on the history database before it inserts data in the history database. Use to improve performance when the archive job copies data to the destination.

Disable Triggers

Determines whether the system disables insert, update, and delete triggers when the archive job deletes rows from tables.

If enabled, the system disables triggers when the archive job deletes data from the source.

If disabled, the system retains the triggers.

Default is enabled.

Only valid for restore archive jobs.

Target Attachment Location

Required to restore archived attachments from the history database or from the Data Vault to the source database.

Enter the target directory where you want to restore the external attachments. The directory is the location where the source application reads the attachments from. You must have write access to the directory.

Source/Staging Attachment Location

Required to restore archived attachments from the history database or from the Data Vault to the source database. The definition depends on the archived attachment location and whether the attachments are encrypted.

To restore attachments from the history database and non-encrypted attachments from the Data Vault, enter the current location of the archived attachments. This is the location where you originally archived the attachments to. You must have read access to the directory.

To restore encrypted attachments, such as Siebel attachments, from the Data Vault, enter a temporary location that is accessible to the ILM engine and the Data Vault Service for External Attachments component. The restore job moves the attachments from the Data Vault Service AM_ATTACHMENTS table to this temporary location.

Staging Script Location

Required to restore archived attachments from the history database or from the Data Vault to the source database. Enter a temporary location to store the script that the restore job generates.

For attachments in the history database and non-encrypted attachments from the Data Vault, the script moves the attachments from the source attachment location to the target attachment location.

For encrypted attachments in the Data Vault, the script uses the source application encryption utility to encrypt the attachments back into the source application proprietary format. Then, the script moves the attachments from the staging attachment location to the target attachment location.

Add-On URL

Required to restore encrypted attachments from the Data Vault to the source database. Enter the URL for the Data Vault Service for External Attachments. The Data Vault Service for External Attachments converts external attachments from the archived format into the source proprietary format.

Available for Microsoft SQL Server 2005 and Microsoft SQL Server 2008.

Creating an SSL Connection

To create an SSL connection between a Microsoft SQL Server database and Data Archive, create or edit the target connection.

1. Log in to the Data Archive UI.

2. If the connection is available, edit the connection through the **Administration** > **Manage Connections** menu.
3. If a connection is not available, click **Administration** > **New Target Connection**.
4. Select a Microsoft SQL Server as the connection type.
5. Enter the connection details to the Microsoft SQL Server.
6. Add the following attributes to the **Port** property:

   ```
   [Database_Port=xxxx];encryptionMethod=SSL;ValidateServerCertificate=false/
   true;CryptoProtocolVersion=TLSv1.2
   ```

   For example:

   ```
   1433;encryptionMethod=SSL;ValidateServerCertificate=false/
   true;CryptoProtocolVersion=TLSv1.2
   ```

   `CryptoProtocolVersion` is the name of the protocol supported by the database server.

   Data Archive uses the **Port** value to create an SSL connection between the Microsoft SQL Server database and Data Archive.
7. Click **Save**.
8. To validate the certificate, you must import the certificate from the SSL enabled Microsoft SQL Server and add it to the respective Java `cacerts` file where Data Archive is running.

Data Vault Target Connections

Define the properties for the target connection to the Data Vault. If the Data Vault is in an external storage platform, specify the connection properties for the platform.

You can configure the following target connection properties:

Staging Directory

Directory in which the Data Vault Loader temporarily stores data as it completes the archive process. Enter the absolute path for the directory.

The directory must be accessible to the ILM application server.

For SAP application retirement, based on the type of connection between the SAP application server and staging area on the Data Archive server, enter one of the following paths:

- If the connection is through FTP, enter the absolute path for the FTP folder on the Data Archive server.
- If the connections is through an NFS mount point, enter the absolute path of the staging folder on the SAP application server.

Number of Rows Per File

Maximum number of rows that the Data Vault Loader stores in a file in the Data Vault. Default is one million rows.

Data Vault Data Directory

Directory in which the Data Vault Loader creates the archive. Enter the absolute path for the directory. You can set up the directory on a local storage or use Network File System (NFS) to connect to a directory on any of the following types of storage devices:

- Direct-attached storage (DAS)

- Network-attached storage (NAS)
- Storage area network (SAN)

You can specify a different directory for each Data Vault target connection. The directory must be accessible to the ILM application server and the Data Vault Service.

If you select an archive store in the **Archive Store Type** property, the Data Vault Loader archives data to the archive store, not to the location specified in the **Data Vault Data Directory** property. Instead, the Data Vault Loader uses the Data Vault data directory as a staging location when it writes data to the archive store.

Data Vault Archive Folder Name

Name of the folder in the Data Vault in which to store the archived data. The Data Vault folder corresponds to the database in the archive source.

Data Vault Host

Host name or IP address of the machine that hosts the Data Vault Service.

Data Vault Port

Port number used by the ssasql command line program and other clients such as the Data Vault SQL Tool and ODBC applications to connect to the Data Vault. Default is 8500.

Data Vault Administration Port

Port number used by the Data Vault Agent and the Data Vault Administration Tool to connect to the Data Vault. Default is 8600.

Data Vault User

Name of the administrator user account to connect to the Data Vault Service.

You can use the default administrator user account created during the Data Vault installation. The user name for the default administrator user account is *dba*.

Data Vault User Password

Password for the administrator user account.

Confirm Password

Verification of the password for the administrator user account.

Add-On URL

URL for the Data Vault Service for External Attachments component. The Data Vault Service for External Attachments converts external attachments from the archived format to the source format. The URL is required to restore encrypted attachments from the Data Vault to the source database.

Maintain Imported Schema Name

Use schema names from the source data imported through the Enterprise Data Manager.

By default, this option is enabled. The Data Vault Loader creates a schema structure in the Data Vault folder that corresponds to the source schema structure imported through the Enterprise Data Manager. It adds the transactional tables to the schemas within the structure. The Data Vault Loader also creates a `dbo` schema and adds the metadata tables to the `dbo` schema.

The imported schema structure is based on the data source. If source connections contain similar structures but use different schema names, you must import the source schema structure for each source connection. For example, you import the schema structure from a development instance. You export metadata from the development instance and import the metadata into the production instance. If the schema names are different in development and production, you must import the schema structure

from the production instance. You cannot use the schema structure imported from the development instance.

If this option is not enabled, the Data Vault Loader creates the `dbo` schema in the Data Vault folder. The Data Vault Loader adds all transactional tables for all schemas and all metadata tables to the `dbo` schema.

Archive Store Type

Storage platform for the Data Vault.

You can select from one of the following archive stores:

- AWS S3
- EMC Atmos
- EMC Centera
- Hitachi Content Archive Platform
- Hadoop HDFS
- Microsoft Azure Storage
- S3 Storage

For best practices information about using cloud storage to archive data, see the How-To article.

If you do not select an archive store type, the Data Vault loader creates the Data Vault in the directory you specify in the **Data Vault Data Directory** property.

Application Owner

User name of the data custodian. Required if you integrate Data Archive with an e-discovery solution such as Exterro Fusion.

Application Owner Email ID

Email address of the data custodian. Required if you integrate Data Archive with an e-discovery solution such as Exterro Fusion.

Use Data Encryption

Select this option to enable data encryption on the compressed Data Vault files during load. If you select this option, you must also choose to use a random key generator provided by Informatica or your choice of a third-party key generator to generate an encryption key. Data Archive stores the encrypted key in the ILM repository for re-use. The encrypted key is unique to the connection and is generated only once for a connection. When this connection is used in an archive or retirement project, the key is passed to Data Vault as a job parameter and is not stored in any file.

Use Random Key Generator

Option to use a random key generator provided by Informatica when data encryption is enabled. When you select this option, the encryption key is generated by a random key generator provided by Informatica (javax.crypto.KeyGenerator).

Use Third Party

Option to use a third-party key generator when data encryption is enabled. If you select this option, you must configure the property "informia.encryptionkey.command" in the `conf.properties` file. Provide the command to run the third-party key generator.

Archive Store Properties

When you select an archive store for the Data Vault target connection, you must specify the connection properties to the archive store.

The archive store that you select determines the connection properties that you must set.

AWS S3

If Data Archive and Data Vault are installed on a Windows 64-bit or Red Hat Enterprise Linux 7 environment, you can set up keyless access to the Amazon Web Services (AWS) S3 archive store.

To access an AWS S3 archive store without specifying the access keys, select **AWS S3** as the **Archive Store Type**.

The following list describes the connection properties that you set for keyless access to the AWS S3 archive store:

Command

The AWS region URL that includes the bucket name. For example: `https://s3.amazonaws.com/testbucket/`

Profile Name

The profile name of the AWS S3 connection.

To connect a Data Archive instance that runs on Amazon EC2 to AWS S3, add an Identity Access Management role. Set **Profile Name** to `default`.

When Data Archive uses a service other than Amazon EC2, connect to AWS S3 in one of the following ways:

- Set the AWS_ACCESS_KEY_ID and AWS_SECRET_ACCESS_KEY properties as environment variables. Set **Profile Name** to `default`.
- Create a credentials file using AWS CLI. This file contains the user defined profile and the values for the AWS_ACCESS_KEY_ID and AWS_SECRET_ACCESS_KEY properties.

Note: You do not have to enter the AWS_ACCESS_KEY_ID and AWS_SECRET_ACCESS_KEY values in the `ssa.ini` file of the Data Vault server or plugin.

EMC Atmos

The following list describes the connection properties that you set for the EMC Atmos archive store:

POOL_ADDR

Pool address for the Atmos storage pool in the following format:

```
<Storage IP address>?<Path and filename of the PEA file>
```

The Pool Entry Authorization (PEA) file controls client application access to the Atmos storage. You must get the pool address from the Atmos system administrator.

The following example shows a pool address:

```
168.159.214.13?C:/SHARE/HW/CAS64/user1.pea
```

EMC Centera

The following list describes the connection properties that you set for the EMC Centera archive store:

POOL_ADDR

Pool address for the Centera storage pool in the following format:

```
<Storage IP address>?<Path and filename of the PEA file>
```

The Pool Entry Authorization (PEA) file controls client application access to the Centera storage. You must get the pool address from the Centera system administrator.

The following example shows a pool address:

```
168.159.214.13?C:/SHARE/HW/CAS64/user1.pea
```

Hitachi Content Archive Platform

The following list describes the properties that you set for the Hitachi Content Archive Platform archive store:

HCP Authentication Token

Authentication token for the namespace and tenant in the HCP server. Get the authentication token from the HCP administrator.

Command

Path to the Data Vault folder in HCP.

For example, the following path shows the location of the Data Vault archive folder named *infa_archive* in the REST space of the namespace *ns0* and tenant *ten1* in the HCP server *hcp.archivas.com*:

```
ns0.ten1.hcp.archivas.com/rest/infa_archive
```

Hadoop HDFS

The following list describes the properties that you set for the Hadoop HDFS archive store:

HDFS URL

Hostname or IP address for the HDFS server.

HDFS Port

Port number to connect to HDFS. The default HDFS port number is 54310.

Command

Path to the directory for the Data Vault in HDFS. Do not include the HDFS prefix or host name.

Microsoft Azure Storage

The following list describes the connection properties that you set for the Microsoft Azure Storage archive store:

Azure Key

Key for the account that has read and write access to the buckets.

Command

The location of the Microsoft Azure container that hosts the `sct` files.

For example: `https://idv.blob.core.windows.net/idvcontainer1/`

In this example, `https://idv.blob.core.windows.net/` is the URL to the Microsoft Azure Storage account and `idvcontainer1` is the name of the container.

S3 Storage

You can configure a simple storage service provided by Amazon or EMC as the Data Vault archive store. To configure, you must provide the Access Key ID and the Secret Access Key.

The following list describes the connection properties that you must set for the S3 archive store:

AWS Key

AWS-key-ID:AWS-secret-key.

For example: `AKIAI62XJDYYYXXXXXX:MFx94W+ZlGTdEXom+21BKBh4Y41y11x1x1x1x1`

Command

The AWS region URL along with the bucket name. For example: `https://s3.amazonaws.com/testbucket/`

In addition to the AWS Key and Command properties, you must install the libcurl API (libcurl.so.4.2.0) and the dependent libraries in the following directories:

Data Vault Service

On Windows, install the files in the root folder of the Data Vault Service directory.

On UNIX, install the files in the `/<File Archive Service Directory>/odbc` directory.

Data Vault Service Agent

On Windows or UNIX, install the files in the root folder of the Data Vault Service agent directory.

If the Data Vault Service agent is installed on multiple machines, install the files on each machine that hosts a Data Vault Service agent.

You must also configure the corresponding connection properties in the `ssa.ini` file of the Data Vault Server. For more information see the *Informatica Data Vault Administrator Guide*.

Oracle Target Connections

When you define an archive target connection for Oracle databases, you can choose from multiple connection types. The connection type that you choose depends on the database version.

Choose one of the following connection types to connect to an Oracle database:

- Oracle 8i
- Oracle 9i
- Oracle 10g
- Oracle 11g
- Oracle 12c
- Oracle 18c

Depending on the connection type, you can configure the following target connection properties:

Host

Host of the target application database server.

Port

Port of the target application database server.

Service Name

Unique identifier or system identifier for the target database server.

Admin Schema Name

Default administration database user for the target database server, such as SYSTEM.

Admin Login Name

Login name is the same as the schema name.

Password

Password for the administration login name.

Apps Schema Name

Application database user that owns the tables that you want to archive to. Also commonly referred to as the history application user.

Application Login Name

Login name for the target application database user.

Password

Password for the application login name.

Data Tablespace

Tablespace name in the application database user that stores the history tables when you run an archive cycle. The archive job always stores the history tables in this tablespace. The archive job only stores staging tables in this tablespace if you configure the source to use staging.

Index Tablespace

Index tablespace name for the history tables.

Database Link to Source

Database link name that connects the target database to the source database. The archive job uses the database link when the job copies data to the destination.

If you do not provide a value, then the archive job uses JDBC to move the data.

Target Attachment Location

Required to restore archived attachments from the history database or from the Data Vault to the source database.

Enter the target directory where you want to restore the external attachments. The directory is the location where the source application reads the attachments from. You must have write access to the directory.

Source/Staging Attachment Location

Required to restore archived attachments from the history database or from the Data Vault to the source database. The definition depends on the archived attachment location and whether the attachments are encrypted.

To restore attachments from the history database and non-encrypted attachments from the Data Vault, enter the current location of the archived attachments. This is the location where you originally archived the attachments to. You must have read access to the directory.

To restore encrypted attachments, such as Siebel attachments, from the Data Vault, enter a temporary location that is accessible to the ILM engine and the Data Vault Service for External Attachments component. The restore job moves the attachments from the Data Vault Service AM_ATTACHMENTS table to this temporary location.

Staging Script Location

Required to restore archived attachments from the history database or from the Data Vault to the source database. Enter a temporary location to store the script that the restore job generates.

For attachments in the history database and non-encrypted attachments from the Data Vault, the script moves the attachments from the source attachment location to the target attachment location.

For encrypted attachments in the Data Vault, the script uses the source application encryption utility to encrypt the attachments back into the source application proprietary format. Then, the script moves the attachments from the staging attachment location to the target attachment location.

Add-On URL

Required to restore encrypted attachments from the Data Vault to the source database. Enter the URL for the Data Vault Service for External Attachments. The Data Vault Service for External Attachments converts external attachments from the archived format into the source proprietary format.

Drop Destination Indexes

Controls whether the archive job drops the index on the history database before it inserts data in the history database. Use to improve performance when the archive job copies data to the destination.

Disable Triggers

Determines whether the system disables insert, update, and delete triggers when the archive job deletes rows from tables.

If enabled, the system disables triggers when the archive job deletes data from the source.

If disabled, the system retains the triggers.

Default is enabled.

Only valid for restore archive jobs.

SSL Enabled

Enable this property if you want to create an SSL connection to an Oracle database. This property is supported for Oracle 11g and later.

ASO Encryption Type

Specifies a list of encryption algorithms used when a client, or another server acting as a client, connects to the Oracle server. The Oracle server uses this list to negotiate a mutually acceptable algorithm with the client. If an algorithm that is not installed is specified, the connection fails with the ORA-12650 error message. This property is supported for Oracle 11g and later.

This property accepts single or multiple comma separated values. For example: AES256,AES192,AES128,3DES112

Based on the negotiation with the server, the Oracle server picks the best algorithm to establish the connection.

ASO Encryption Level

Specifies the data integrity behavior when a client, or another server acting as a client, connects to the Oracle server. The behavior partially depends on the setting used in the source database connection. This property is supported for Oracle 11g and later.

Use one the following values:

- ACCEPTED
- REJECTED
- REQUESTED
- REQUIRED

Default is ACCEPTED.

If you specify a value other than the above values, the connection fails with the following error:

Invalid parameter, use one of ACCEPTED, REJECTED, REQUESTED and REQUIRED

ASO Checksum Type

Specifies a list of data integrity algorithms that when a client, or another server acting as a client, connects to the Oracle server. The list of data integrity algorithms are listed in order of intended use. This list is used to negotiate a mutually acceptable algorithm with the other end of the connection. Each algorithm is checked against the list of available client algorithm types until a match is found. If an algorithm is specified that is not installed, the connection fails with the ORA-12650 error message. This property is supported for Oracle 11g and later. This property accepts single or multiple comma separated values. For example:

MD5, SHA1

ASO Checksum Level

Specifies the data integrity behavior when a client, or another server acting as a client, connects to this server. The behavior partially depends on the setting used in the source database connection. This property is supported for Oracle 11g and later.

Use one the following values:

- ACCEPTED
- REJECTED
- REQUESTED
- REQUIRED

Default is ACCEPTED.

If you specify a value other than the above values, the connection fails with the following error:

Invalid parameter, use one of ACCEPTED, REJECTED, REQUESTED and REQUIRED

OID Enabled

Oracle Internet Directory (OID) is an LDAP directory that uses an Oracle database for storage. When you enable OID, you must provide the OID database host name for the Host parameter. You must also provide the service name that specifies the distinguished name of the database, which is configured in OID.

For example:

```
ORA12C,cn=OracleContext,dc=informatica,dc=com
```

This property is available for Oracle 11g and 12c.

Teradata Target Connections

Define archive target connection properties to archive data to Teradata databases.

You can configure the following target connection properties:

Host

Host of the target application database server.

Database Name

Unique identifier or system identifier for the target database server.

Admin Schema Name

Default administration database user for the target database server, such as SYSTEM.

Admin Login Name

Login name for the administration database server. This user does not require special permissions as it is only used to connect to the target database. You can provide any user name, such as a database connection user or a read-only user.

Password

Password for the administration login name.

Apps Schema Name

Application database user that owns the tables that you want to archive to. Also commonly referred to as the history application user.

Application Login Name

Login name for the target application database user.

Password

Password for the application login name.

Drop Destination Indexes

Controls whether the archive job drops the index on the history database before it inserts data in the history database. Use to improve performance when the archive job copies data to the destination.

Disable Triggers

Determines whether the system disables insert, update, and delete triggers when the archive job deletes rows from tables.

If enabled, the system disables triggers when the archive job deletes data from the source.

If disabled, the system retains the triggers.

Default is enabled.

Only valid for restore archive jobs.

Netezza Target Connections

Define archive target connection properties to connect to applications on Netezza databases.

You can configure the following target connection properties:

Host

Host of the target application database server.

Port

Port of the target application database server.

Service Name

Unique identifier or system identifier for the target database server.

ILM Repository Administrator User

Default administration database user for the target database server.

Admin Login Name

Login name for the administration database server. This user does not require special permissions as it is only used to connect to the target database. You can provide any user name, such as a database connection user or a read-only user.

Password

Password for the administration login name.

Application User Name

Database user that owns the application tables on the target. The application user name is identical to the administration user name.

Application Login Name

Login name that connects to the target database. Only used for the database connection.

Password

Password for the application login name. The password is identical to the administration login password.

Data Tablespace

Not applicable.

Index Tablespace

Not applicable.

Database Link to Source

Not applicable.

Drop Destination Indexes

Not applicable.

Target Attachment Location

Required to restore archived attachments from the history database or from the Data Vault to the source database.

Enter the target directory where you want to restore the external attachments. The directory is the location where the source application reads the attachments from. You must have write access to the directory.

Source/Staging Attachment Location

Required to restore archived attachments from the history database or from the Data Vault to the source database. The definition depends on the archived attachment location and whether the attachments are encrypted.

To restore attachments from the history database and non-encrypted attachments from the Data Vault, enter the current location of the archived attachments. This is the location where you originally archived the attachments to. You must have read access to the directory.

To restore encrypted attachments, such as Siebel attachments, from the Data Vault, enter a temporary location that is accessible to the ILM engine and the Data Vault Service for External Attachments component. The restore job moves the attachments from the Data Vault Service AM_ATTACHMENTS table to this temporary location.

Move Attachments in Synchronous Mode

Determines whether the archive job automatically archives external attachments or whether you run a standalone job to move the attachments after the archive job completes. If you provide a source file location for attachments, the archive job creates SQL scripts in the file server location and associates the scripts with the archive job ID.

If enabled, the archive job runs the scripts during the archive process. The run procedures configuration in the entity determines when the archive job archives the attachments.

If disabled, you must initiate the movement after the archive job completes. You can manually run the scripts in the file server location or you can run a standalone job to move the attachments. If you run the standalone job, you must provide the archive job ID. The job then looks for the scripts that are associated to the archive job ID.

This attribute only applies to external attachments.

Staging Script Location

Required to restore archived attachments from the history database or from the Data Vault to the source database. Enter a temporary location to store the script that the restore job generates.

For attachments in the history database and non-encrypted attachments from the Data Vault, the script moves the attachments from the source attachment location to the target attachment location.

For encrypted attachments in the Data Vault, the script uses the source application encryption utility to encrypt the attachments back into the source application proprietary format. Then, the script moves the attachments from the staging attachment location to the target attachment location.

Add-On URL

Required to restore encrypted attachments from the Data Vault to the source database. Enter the URL for the Data Vault Service for External Attachments. The Data Vault Service for External Attachments converts external attachments from the archived format into the source proprietary format.

Creating a Target Connection

Define target connections before you create an archive project.

1. Click **Administration** > **New Target Connection**.
2. Enter the general connection properties.

   After you choose the connection type, the system dynamically refreshes the screen with the properties that are relevant for the connection type.
3. Enter the database-specific connection properties.

   For more information, see the connection properties for the target database.
4. Click **Save**.

Copying a Target Connection

You can copy a connection if the property configuration is similar to the one that you need to create.

When you copy an existing connection, the system creates a connection. The system renames the connection to `Copy of <connection name>`. You can edit the connection name to accurately reflect the connection. The connection includes the same configuration settings and attribute values from the connection that you copied. You can keep or change the copied values.

1. Click **Administration** > **Manage Connections**.
2. Click the **Copy** icon next to the connection that you want to copy.

   The **New/Edit Archive Target** page appears with the default property values from the copied connection.
3. Optionally, change the connection name and connection property values.
4. Click **Save.**

Database User Password Changes

Your company security policy may require that you change database user passwords frequently. When you change database user passwords, you must update connections in the ILM repository.

You can edit the connection manually. Or, you can use an SQL script to update passwords in the ILM repository without logging in to the ILM application. Use the script to change the password for the administration, production application, and staging database users and for the optimized file target. The script is available as a batch file or a shell file and is located in the optional directory of the ILM installation.

The script provides a more efficient option to update the ILM repository when company policy requires you to update database user passwords on a regular basis. For example, you must change database user passwords every 20 days.

Properties File

The script uses values in a properties file to connect to the ILM repository and to change the passwords. You create and configure the properties file before you run the script.

The following table describes the required properties for the properties file:

Property	Description
amhome.connection.host	ILM repository host name.
amhome.connection.port	ILM repository port number.
amhome.connection.sid	ILM repository database service name or database name.
amhome.connection.username	ILM repository user name.
amhome.connection.password	ILM repository user password.
amhome.connection.schemaname	Unique database placeholder for tables.

Property	Description
amhome.connection.validate.sql	Script that determines if the installation repository is the ILM repository.
amhome.values.separator	Symbol that indicates the separation of multiple values for the other properties. Use a unique value that is not used in the database names or passwords. For example, use `@@`.
amhome.rep.names	ILM repository database schema name for the Admin, Apps, or Staging schemas. Use the same name as the **Schema Name** field in the Edit Archive Source or Edit Archive Target menu. To enter multiple values, use the separator that you defined for the amhome.values.separator property. For example, enter `SYSTEM@@STAGING` to change the password for the SYSTEM and STAGING database schemas.
amhome.rep.usernames	Login name for the database schema or the Data Vault user. Use the same name as the **Login Name** field in the Edit Archive Source or Edit Archive Target menu. To enter multiple values, use the separator that you defined for the amhome.values.separator property.
amhome.rep.newpasswords	New password for the login name. To enter multiple values, use the separator that you defined for the amhome.values.separator property.

Changing Database User Passwords

Configure the ILM repository connection details and the database user details in a properties file. The script uses the properties file to connect to the ILM repository and to validate the database user login name and password. If the database user login details are valid, the script updates the ILM repository with the new password.

The ILM application server does not need to be running when you run the script.

Note: If a Data Vault user password expires, you can update the ILM repository with a new password only if folders have been created in Data Vault with the Create Archive Folder job.

1. Navigate to the following directory:

   `<ILM installation directory>/optional`
2. Use the `PasswordChangesSampleProperty` file to create a property file.
3. Add the required properties to the property file.
4. Save and close the property file.
5. From the same directory, run one of the following files:
   - `PasswordChangeEnv.bat` for Windows.
   - `PasswordChangeEnv.sh` for Linux and UNIX.
6. Enter the location of the property file that you created.

   The script executes and creates the `PasswordChangeLog` log file in the same directory.
7. View the log file to show the script outcome.

CHAPTER 6

Archive Store Configuration

This chapter includes the following topics:

Archive Store Configuration Overview

You can create a Data Vault archive on different storage platforms. To create a Data Vault archive on a storage platform, you must select the type of archive store to use and configure the Data Vault Service to work with the archive store.

You can archive to the following storage platforms:

- EMC Centera
- EMC Atmos
- Hitachi Content Platform (HCP)
- Hadoop Distributed File System (HDFS)
- Microsoft Azure Storage

You can archive different entities from the same application to different archive stores. For example, business entities in an application can have different compliance regulations in different states. To archive different business entities to different archive stores, create multiple target connections and configure a separate archive store for each connection. Then run a separate job to archive each set of business entities.

EMC Centera and EMC Atmos Configuration

You can use EMC Centera or EMC Atmos content addressable storage as an archive store in Data Archive.

To create an archive in Centera or Atmos, complete the following tasks:

1. Install the EMC Centera API.
2. Create a Data Vault target connection.

3. Run the Create Archive Folder job.
4. Copy the connection to other Data Vault Service configuration files.
5. Validate the connection to the Data Vault.

Step 1. Install the EMC Centera API Files

Data Archive uses the EMC Centera API to archive data to Centera or Atmos. Install the EMC Centera API files on the machines that host the Data Vault components.

You can download the Centera API files for your operating system from the EMC community web site:

`https://community.emc.com/community/edn/centera`

Download the Centera SDK version 3.3.718 or later.

Install the Centera SDK on the machines where the following Data Vault components are installed:

Data Vault Service

On Windows, install the files in the root of the Data Vault Service directory.

On UNIX, install the files in `<Data Vault Service Directory>/odbc`.

Data Vault Agent

On Windows or UNIX, install the files in the root of the Data Vault Agent directory.

If the Data Vault Agent is installed on multiple machines, install the files on each machine that hosts a Data Vault Agent.

Data Vault Service plug-in for Data Archive

On Windows, install the files in `<Data Archive Directory>\webapp\file_archive`.

On UNIX, install the files in `<Data Archive Directory>/webapp/file_archive/odbc`.

On UNIX, if the libraries are stored in a location other than the default, you must set the library path in the appropriate system variable, such as LD_LIBRARY_PATH, SHLIB_PATH, or LIB_PATH.

Step 2. Create the Target Connection

In Data Archive, create a target connection to the archive store. To create a target connection to Centera, set the archive store type to EMC Centera. To create a target connection to Atmos, set the archive store type to EMC Atmos.

When you select a Data Vault target connection, set the following the properties:

Staging Directory

Directory in which the Data Vault Loader temporarily stores data as it completes the archive process. Enter the absolute path for the directory.

The directory must be accessible to the ILM application server.

For SAP application retirement, based on the type of connection between the SAP application server and staging area on the Data Archive server, enter one of the following paths:

- If the connection is through FTP, enter the absolute path for the FTP folder on the Data Archive server.
- If the connections is through an NFS mount point, enter the absolute path of the staging folder on the SAP application server.

Number of Rows Per File

Maximum number of rows that the Data Vault Loader stores in a file in the Data Vault. Default is one million rows.

Data Vault Data Directory

Directory in which the Data Vault Loader creates the archive. Enter the absolute path for the directory. You can set up the directory on a local storage or use Network File System (NFS) to connect to a directory on any of the following types of storage devices:

- Direct-attached storage (DAS)
- Network-attached storage (NAS)
- Storage area network (SAN)

You can specify a different directory for each Data Vault target connection. The directory must be accessible to the ILM application server and the Data Vault Service.

If you select an archive store in the **Archive Store Type** property, the Data Vault Loader archives data to the archive store, not to the location specified in the **Data Vault Data Directory** property. Instead, the Data Vault Loader uses the Data Vault data directory as a staging location when it writes data to the archive store.

Data Vault Archive Folder Name

Name of the folder in the Data Vault in which to store the archived data. The Data Vault folder corresponds to the database in the archive source.

Data Vault Host

Host name or IP address of the machine that hosts the Data Vault Service.

Data Vault Port

Port number used by the ssasql command line program and other clients such as the Data Vault SQL Tool and ODBC applications to connect to the Data Vault. Default is 8500.

Data Vault Administration Port

Port number used by the Data Vault Agent and the Data Vault Administration Tool to connect to the Data Vault. Default is 8600.

Data Vault User

Name of the administrator user account to connect to the Data Vault Service.

You can use the default administrator user account created during the Data Vault installation. The user name for the default administrator user account is *dba*.

Data Vault User Password

Password for the administrator user account.

Confirm Password

Verification of the password for the administrator user account.

Add-On URL

URL for the Data Vault Service for External Attachments component. The Data Vault Service for External Attachments converts external attachments from the archived format to the source format. The URL is required to restore encrypted attachments from the Data Vault to the source database.

Maintain Imported Schema Name

Use schema names from the source data imported through the Enterprise Data Manager.

By default, this option is enabled. The Data Vault Loader creates a schema structure in the Data Vault folder that corresponds to the source schema structure imported through the Enterprise Data Manager. It adds the transactional tables to the schemas within the structure. The Data Vault Loader also creates a `dbo` schema and adds the metadata tables to the `dbo` schema.

The imported schema structure is based on the data source. If source connections contain similar structures but use different schema names, you must import the source schema structure for each source connection. For example, you import the schema structure from a development instance. You export metadata from the development instance and import the metadata into the production instance. If the schema names are different in development and production, you must import the schema structure from the production instance. You cannot use the schema structure imported from the development instance.

If this option is not enabled, the Data Vault Loader creates the `dbo` schema in the Data Vault folder. The Data Vault Loader adds all transactional tables for all schemas and all metadata tables to the `dbo` schema.

Archive Store Type

Storage platform for the Data Vault. For a target connection to Centera, select the archive store type **EMC Centera**. For a target connection to Atmos, select the archive store type **EMC Atmos**.

POOL_ADDR

Pool address for the Centera or Atmos storage pool in the following format:

```
<Storage IP address>?<Path and filename of the PEA file>
```

The Pool Entry Authorization (PEA) file controls client application access to the Centera or Atmos storage. You must get the pool address from the Centera or Atmos system administrator.

The following example shows a pool address:

```
168.159.214.13?C:/SHARE/HW/CAS64/user1.pea
```

Step 3. Run the Create Archive Folder Job

In Data Archive, run the Create Archive Folder job to create the Data Vault folder and the connection to EMC Centera or EMC Atmos.

The Create Archive Folder job creates the Data Vault folder and adds a content addressable storage (CAS) connection entry to the ssa.ini file in the Data Vault Service plug-in in Data Archive. The job sets the name of the Data Vault folder and the name of the CAS connection based on the folder name property specified in the target connection.

For example, the **Data Vault Archive Folder Name** property in the target connection is set to *CAS_Sales*. The Create Archive Folder job creates a Data Vault archive folder named *CAS_Sales* and adds a CAS connection named *CAS_Sales* to the ssa.ini file.

The following example shows an entry for a CAS connection named *CAS_Sales* in the ssa.ini file:

```
[CAS_CONNECTION CAS_Sales]
POOL_ADDR=168.159.214.19?c:\ilm-fas\informatica.pea
```

Step 4. Copy the CAS Connection to Other Data Vault Service Configuration Files

The CAS connection definition on the machine that hosts Data Archive must match the CAS connection definition on the machines that host other Data Vault Service components. Copy the CAS connection

definition from the ssa.ini file on the machine that hosts Data Archive to the ssa.ini files on the machines that host the Data Vault Service and Data Vault Agent.

After you run the Create Archive Folder job, go to the Data Vault Service plug-in directory in Data Archive and find the ssa.ini file. Copy the CAS connection definition to the ssa.ini file on the machine that hosts the Data Vault Service. Additionally, if you have installed the Data Vault Agent on another machine, copy the CAS connection definition to the ssa.ini file on the machine that hosts the Data Vault Agent. If you have installed the Data Vault Agent on multiple machines, copy the CAS connection definition to the ssa.ini file on each machine that hosts a Data Vault Agent.

Step 5. Validate the Connection

Verify the connection to Centera or Atmos from the Data Vault Service and from the Data Vault Service plug-in in Data Archive.

Use the Data Vault Service *ssadrv* administration command to validate the connection to the Centera or Atmos Data Vault store. On the machine that hosts the Data Vault Service, run the following command:

```
ssadrv –a cas://<Connection Name>/
```

For example, run the following command:

```
ssadrv –a cas://CAS_Sales/
```

In this command, *CAS_Sales* is the name of the connection defined in the ssa.ini.

You can run the same command on the machine that hosts Data Archive to test the connection from Data Archive to Centera and Atmos.

Hitachi Content Platform Configuration

You can use Hitachi Content Platform (HCP) as an archive store in Data Archive.

To set up an archive in HCP, complete the following tasks:

1. Get the authentication token from the HCP administrator.
2. Create a folder in HCP for the archive.
3. Install the cURL library.
4. Add the HCP server to the hosts file.
5. Create a Data Vault target connection.
6. Run the Create Archive Folder job.
7. Copy the connection to the ssa.ini file for other Data Vault components.
8. Validate the connection to the Data Vault.

Step 1. Get the Authentication Token

Data Archive requires an authentication token as authorization to communicate with the HCP server.

You can get the authentication token for the namespace that you want to access in the HCP server from the HCP server administrator. The authentication token is stored in the HCP_AUTH_TOKEN variable.

Typically, the value of the token looks like the following string:

```
hcp-ns-auth=bnMwZGF0YQ==:79e262a81dd19d40ae008f74eb59edce
```

When you create the target connection to the HCP archive store in Data Archive, you must specify the authentication token for the namespace in the HCP server.

Step 2. Create a Folder in HCP

In HCP, create a folder for the archive in the namespace that you want to use.

Use the HCP console to create the Data Vault folder in the REST space of the namespace and tenant that you want to use for Data Archive.

The path to a folder in a REST space in HCP includes the following elements:

```
<Namespace>.<Tenant Name>.<HCP Server>/rest/<Folder Name>
```

The following example shows the path to the folder *infa_archive* in the REST space of the namespace *ns0* and tenant *ten1*:

```
ns0.ten1.hcp.archivas.com/rest/infa_archive
```

Step 3. Install the cURL Library

Data Archive uses a URL transfer library named cURL to archive data to HCP. Install the cURL library files on the machines that host the Data Vault components.

You can download the cURL library for your operating system from the cURL download web site:

```
http://curl.haxx.se/download.html
```

Download the cURL library version 7.27 or later. Informatica does not ship the cURL library files with Data Archive.

If you plan to use the secure HTTPS protocol, you must download a version of the cURL library that supports SSL.

Install the cURL library files on the machines where the following Data Vault Service components are installed:

Data Vault Service

On Windows, install the files in the root of the Data Vault Service directory.

On UNIX, install the files in `<Data Vault Service Directory>/odbc`.

Data Vault Agent

On Windows or UNIX, install the files in the root of the Data Vault Agent directory.

If the Data Vault Agent is installed on multiple machines, install the cURL library files on each machine that hosts a Data Vault Agent.

Data Vault Service plug-in for Data Archive

On Windows, install the files in `<Data Archive Directory>\webapp\file_archive`.

On UNIX, install the files in `<Data Archive Directory>/webapp/file_archive/odbc`.

On UNIX, if the libraries are stored in a location other than the default, you must set the library path in the appropriate system variable, such as LD_LIBRARY_PATH, SHLIB_PATH, or LIB_PATH.

Step 4. Add the HCP Server to the Hosts File

The hosts files on the machines where the Data Vault Service components are installed must include the hostname and IP address of the HCP server.

Add the HCP server address to the hosts file on the machines where the following Data Vault components are installed:

- Data Vault Service
- Data Vault Agent
- Data Vault Service Data Archive plug-in

On Windows, the hosts file is in the following directory: `c:\Windows\System32\drivers\etc`

On UNIX, the hosts file is in the following directory: `/etc`

For example, the following entry in the hosts file provides access to the namespace *ns0* of tenant **ten1** in the HCP server:

```
# access to HCP server
10.17.0.49 ns0.ten1.hcp.archivas.com     # REST HTTP data access to namespace ns0 of
tenant ten1
```

Step 5. Create the Data Vault Target Connection

In Data Archive, create a target connection to HCP and set the archive store type to **Hitachi Content Archive Platform**.

The following list describes the properties that you need to set for the target connection:

Staging Directory

Directory in which the Data Vault Loader temporarily stores data as it completes the archive process. Enter the absolute path for the directory.

The directory must be accessible to the ILM application server.

For SAP application retirement, based on the type of connection between the SAP application server and staging area on the Data Archive server, enter one of the following paths:

- If the connection is through FTP, enter the absolute path for the FTP folder on the Data Archive server.
- If the connections is through an NFS mount point, enter the absolute path of the staging folder on the SAP application server.

Number of Rows Per File

Maximum number of rows that the Data Vault Loader stores in a file in the Data Vault. Default is one million rows.

Data Vault Data Directory

Directory in which the Data Vault Loader creates the archive. Enter the absolute path for the directory. You can set up the directory on a local storage or use Network File System (NFS) to connect to a directory on any of the following types of storage devices:

- Direct-attached storage (DAS)
- Network-attached storage (NAS)
- Storage area network (SAN)

You can specify a different directory for each Data Vault target connection. The directory must be accessible to the ILM application server and the Data Vault Service.

If you select an archive store in the **Archive Store Type** property, the Data Vault Loader archives data to the archive store, not to the location specified in the **Data Vault Data Directory** property. Instead, the Data Vault Loader uses the Data Vault data directory as a staging location when it writes data to the archive store.

Data Vault Archive Folder Name

Name of the folder in the Data Vault in which to store the archived data. The Data Vault folder corresponds to the database in the archive source.

Data Vault Host

Host name or IP address of the machine that hosts the Data Vault Service.

Data Vault Port

Port number used by the ssasql command line program and other clients such as the Data Vault SQL Tool and ODBC applications to connect to the Data Vault. Default is 8500.

Data Vault Administration Port

Port number used by the Data Vault Agent and the Data Vault Administration Tool to connect to the Data Vault. Default is 8600.

Data Vault User

Name of the administrator user account to connect to the Data Vault Service.

You can use the default administrator user account created during the Data Vault installation. The user name for the default administrator user account is *dba*.

Data Vault User Password

Password for the administrator user account.

Confirm Password

Verification of the password for the administrator user account.

Add-On URL

URL for the Data Vault Service for External Attachments component. The Data Vault Service for External Attachments converts external attachments from the archived format to the source format. The URL is required to restore encrypted attachments from the Data Vault to the source database.

Maintain Imported Schema Name

Use schema names from the source data imported through the Enterprise Data Manager.

By default, this option is enabled. The Data Vault Loader creates a schema structure in the Data Vault folder that corresponds to the source schema structure imported through the Enterprise Data Manager. It adds the transactional tables to the schemas within the structure. The Data Vault Loader also creates a `dbo` schema and adds the metadata tables to the `dbo` schema.

The imported schema structure is based on the data source. If source connections contain similar structures but use different schema names, you must import the source schema structure for each source connection. For example, you import the schema structure from a development instance. You export metadata from the development instance and import the metadata into the production instance. If the schema names are different in development and production, you must import the schema structure from the production instance. You cannot use the schema structure imported from the development instance.

If this option is not enabled, the Data Vault Loader creates the `dbo` schema in the Data Vault folder. The Data Vault Loader adds all transactional tables for all schemas and all metadata tables to the `dbo` schema.

Archive Store Type

Storage platform for the Data Vault. Select the **Hitachi Content Archive Platform** archive store.

HCP Authentication Token

Authentication token for the namespace and tenant in the HCP server. Get the authentication token from the HCP administrator.

Command

Path to the Data Vault folder in HCP.

For example, the following path shows the location of the Data Vault archive folder named *infa_archive* in the REST space of the namespace *ns0* and tenant *ten1* in the HCP server *hcp.archivas.com*:

```
ns0.ten1.hcp.archivas.com/rest/infa_archive
```

Step 6. Run the Create Archive Folder Job

In Data Archive, run the Create Archive Folder job to create the Data Vault folder and the connection to HCP.

The Create Archive Folder job creates the Data Vault folder and adds an HCP connection entry to the ssa.ini file in the Data Vault Service plug-in in Data Archive. The job sets the name of the Data Vault folder and the HCP connection to the name specified in the target connection.

For example, the Data Vault Archive Folder Name property in the target connection is set to *HCPSales*. The Create Archive Folder job creates a Data Vault archive folder named *HCPSales* and adds an HCP connection named *HCPSales* to the ssa.ini file.

The following example shows an entry for an HCP connection named *HCPSales* in the ssa.ini file:

```
[HCP_CONNECTION HCPSales]
HCP_AUTH_TOKEN=hcp-ns-auth=bnMwZGF0YQ==:79e262a81dd19d40ae008f74eb59edce
RETENTION_PERIOD=10
```

Step 7. Copy the HCP Connection to Other Data Vault Service Configuration Files

The HCP connection definition on the machine that hosts Data Archive must match the HCP connection definition on the machines that host other Data Vault Service components. Copy the HCP connection definition from the ssa.ini file on the machine that hosts Data Archive to the ssa.ini files on the machines that host the Data Vault Service and Data Vault Agent.

After you run the Create Archive Folder job, go to the Data Vault Service plug-in directory in Data Archive and find the ssa.ini file. Copy the HCP connection definition to the ssa.ini file on the machine that hosts the Data Vault Service. Additionally, if you have installed the Data Vault Agent on another machine, copy the HCP connection definition to the ssa.ini file on the machine that hosts the Data Vault Agent. If you have installed the Data Vault Agent on multiple machines, copy the HCP connection definition to the ssa.ini file on each machine that hosts a Data Vault Agent.

Step 8. Validate the Connection to HCP

Verify the connection to HCP from the Data Vault Service and from the Data Vault Service plug-in in Data Archive.

Use the Data Vault Service *ssadrv* administration command to validate the connection to the HCP archive store. On the machine that hosts the Data Vault Service, run the following command:

```
ssadrv -a hcp://<Connection Name>/<HCP path to REST space>/<Data Vault Folder>
```

For example, run the following command:

```
ssadrv -a hcp://HCPConnect/ns0.ten1.hcp.archivas.com/rest/infa_archive
```

In this command, *HCPConnect* is the name of the connection to HCP defined in the ssa.ini and *ns0.ten1.hcp.archivas.com/rest* is the HCP namespace and REST root directory. *infa_archive* is the name of the Data Vault folder in the HCP archive store.

You can also run the same command on the machine that hosts Data Archive to test the connection from Data Archive to HCP.

Hadoop Distributed File System Configuration

You can use a Hadoop Distributed File System (HDFS) that runs on Linux as an archive store in Data Archive.

To create an archive in HDFS, complete the following tasks:

1. Install the libhdfs API files.
2. Create a directory in HDFS.
3. Create a Data Vault target connection.
4. Run the Create Archive Folder job.
5. Copy the connection to other Data Vault Service configuration files.
6. Validate the connection to HDFS.

Step 1. Install the libhdfs API Files

The libhdfs API provides access to files in a Hadoop file system. Data Archive requires the libhdfs API files to access an archive in HDFS. The Hadoop installation includes the libhdfs API.

The Data Vault Service requires the following libhdfs files:

- commons-logging-api-1.0.4.jar
- hadoop-0.20.2-core.jar
- libhdfs.so

To install the libhdfs API, copy the libhdfs files to the machines where the following Data Vault Service components are installed:

Data Vault Service

On Windows, copy the files to the root of the Data Vault Service directory.

On UNIX, copy the files to `<Data Vault Service Directory>/odbc`.

Data Vault Agent

On Windows or UNIX, copy the files to the root of the Data Vault Agent directory.

If the Data Vault Agent is installed on multiple machines, copy the libhdfs API files to all machines that host a Data Vault Agent.

Data Vault Service plug-in for Data Archive

On Windows, copy the files to `<Data Archive Directory>\webapp\file_archive`.

On UNIX, copy the files to `<Data Archive Directory>/webapp/file_archive/odbc`.

After the installation, verify that the CLASSPATH environment variable includes the location of the libhdfs files.

Step 2. Create a Directory in HDFS

In HDFS, create a directory for the archive.

Step 3. Create the Target Connection

In Data Archive, create a target connection to HDFS and set the archive store type to **Hadoop HDFS**.

The following list describes the properties that you need to set for the target connection:

Staging Directory

Directory in which the Data Vault Loader temporarily stores data as it completes the archive process. Enter the absolute path for the directory.

The directory must be accessible to the ILM application server.

For SAP application retirement, based on the type of connection between the SAP application server and staging area on the Data Archive server, enter one of the following paths:

- If the connection is through FTP, enter the absolute path for the FTP folder on the Data Archive server.
- If the connections is through an NFS mount point, enter the absolute path of the staging folder on the SAP application server.

Number of Rows Per File

Maximum number of rows that the Data Vault Loader stores in a file in the Data Vault. Default is one million rows.

Data Vault Data Directory

Directory in which the Data Vault Loader creates the archive. Enter the absolute path for the directory. You can set up the directory on a local storage or use Network File System (NFS) to connect to a directory on any of the following types of storage devices:

- Direct-attached storage (DAS)
- Network-attached storage (NAS)
- Storage area network (SAN)

You can specify a different directory for each Data Vault target connection. The directory must be accessible to the ILM application server and the Data Vault Service.

If you select an archive store in the **Archive Store Type** property, the Data Vault Loader archives data to the archive store, not to the location specified in the **Data Vault Data Directory** property. Instead, the Data Vault Loader uses the Data Vault data directory as a staging location when it writes data to the archive store.

Data Vault Archive Folder Name

Name of the folder in the Data Vault in which to store the archived data. The Data Vault folder corresponds to the database in the archive source.

Data Vault Host

Host name or IP address of the machine that hosts the Data Vault Service.

Data Vault Port

Port number used by the ssasql command line program and other clients such as the Data Vault SQL Tool and ODBC applications to connect to the Data Vault. Default is 8500.

Data Vault Administration Port

Port number used by the Data Vault Agent and the Data Vault Administration Tool to connect to the Data Vault. Default is 8600.

Data Vault User

Name of the administrator user account to connect to the Data Vault Service.

You can use the default administrator user account created during the Data Vault installation. The user name for the default administrator user account is *dba*.

Data Vault User Password

Password for the administrator user account.

Confirm Password

Verification of the password for the administrator user account.

Add-On URL

URL for the Data Vault Service for External Attachments component. The Data Vault Service for External Attachments converts external attachments from the archived format to the source format. The URL is required to restore encrypted attachments from the Data Vault to the source database.

Maintain Imported Schema Name

Use schema names from the source data imported through the Enterprise Data Manager.

By default, this option is enabled. The Data Vault Loader creates a schema structure in the Data Vault folder that corresponds to the source schema structure imported through the Enterprise Data Manager. It adds the transactional tables to the schemas within the structure. The Data Vault Loader also creates a `dbo` schema and adds the metadata tables to the `dbo` schema.

The imported schema structure is based on the data source. If source connections contain similar structures but use different schema names, you must import the source schema structure for each source connection. For example, you import the schema structure from a development instance. You export metadata from the development instance and import the metadata into the production instance. If the schema names are different in development and production, you must import the schema structure from the production instance. You cannot use the schema structure imported from the development instance.

If this option is not enabled, the Data Vault Loader creates the `dbo` schema in the Data Vault folder. The Data Vault Loader adds all transactional tables for all schemas and all metadata tables to the `dbo` schema.

Archive Store Type

Storage platform for the Data Vault. Select the **Hadoop HDFS** archive store.

HDFS URL

Hostname or IP address for the HDFS server.

HDFS Port

Port number to connect to HDFS. The default HDFS port number is 54310.

Command

Path to the directory for the Data Vault in HDFS. Do not include the HDFS prefix or host name.

Step 4. Run the Create Archive Folder Job

In Data Archive, run the Create Archive Folder job to create the Data Vault folder and the connection to HDFS.

The Create Archive Folder job creates the Data Vault folder and adds a Hadoop connection entry to the ssa.ini file in the Data Vault Service plug-in in Data Archive. The job sets the name of the Data Vault folder and the name of the Hadoop connection based on the folder name property specified in the target connection.

For example, the **Data Vault Archive Folder Name** property in the target connection is set to *HDFS_Sales*. The Create Archive Folder job creates a Data Vault archive folder named *HDFS_Sales* and adds a Hadoop connection named *HDFS_Sales* to the ssa.ini file.

The following example shows an entry for a Hadoop connection named *HDFS_Sales* in the ssa.ini file:

```
[HADOOP_CONNECTION HDFS_Sales]
URL = 10.17.40.25
PORT = 54310
```

Step 5. Copy the Hadoop Connection to Other Data Vault Service Configuration Files

The Hadoop connection definition on the machine that hosts Data Archive must match the Hadoop connection definition on the machines that host other Data Vault components. Copy the Hadoop connection definition from the ssa.ini file on the machine that hosts Data Archive to the ssa.ini files on the machines that host the Data Vault Service and Data Vault Agent.

After you run the Create Archive Folder job, go to the Data Vault Service plug-in directory in Data Archive and find the ssa.ini file. Copy the Hadoop connection definition to the ssa.ini file on the machine that hosts the Data Vault Service. Additionally, if you have installed the Data Vault Agent on another machine, copy the Hadoop connection definition to the ssa.ini file on the machine that hosts the Data Vault Agent. If you have installed the Data Vault Agent on multiple machines, copy the Hadoop connection definition to the ssa.ini file on each machine that hosts a Data Vault Agent.

Step 6. Validate the Connection to HDFS

Verify the connection to HDFS from the Data Vault Service and from the Data Vault Service plug-in in Data Archive.

Use the Data Vault Service *ssadrv* administration command to validate the connection to the HDFS Data Vault store.

On the machine that hosts the Data Vault Service, run the following command:

```
ssadrv -a hdfs://<Connection Name>/<Path to the Data Vault folder in HDFS>
```

For example, run the following command:

```
ssadrv -a hdfs://HDFS_Sales/data/sandqa1/infa_archive
```

In this command, *HDFS_Sales* is the name of the Hadoop connection defined in the ssa.ini file and *data/sandqa1/infa_archive* is the path to the Data Vault archive folder named *infa_archive* in the HDFS Data Vault store.

You can run the same command on the machine that hosts Data Archive to test the connection from Data Archive to HDFS.

Microsoft Azure Storage Configuration

You can use Microsoft Azure Storage as external storage for Data Archive on a Windows Operating System.

Before you configure Microsoft Azure Storage for external storage, perform the following tasks:

- Create a blob storage or V2 storage account in Microsoft Azure Storage.
- Create containers in the Azure Blob Storage or V2 storage account.

Perform the following tasks to set up a Microsoft Azure Storage archive store:

1. Create a Data Vault target connection.
2. Run the Create Archive Folder job.
3. Copy the connection to other Data Vault Service configuration files.
4. Validate the connection to the Microsoft Azure Storage store.

Data Archive supports the following Microsoft Azure storage types:

- Azure Storage general-purpose V2 storage accounts.
- Azure Blob Storage accounts.

Step 1. Create the Data Vault Target Connection

In Data Archive, create a target connection to Microsoft Azure Storage.

The following list describes the properties that you must set to connect to the target store:

Staging Directory

Directory in which the Data Vault Loader temporarily stores data as it completes the archive process. Enter the absolute path for the directory.

The directory must be accessible to the ILM application server.

For SAP application retirement, based on the type of connection between the SAP application server and staging area on the Data Archive server, enter one of the following paths:

- If the connection is through FTP, enter the absolute path for the FTP folder on the Data Archive server.
- If the connections is through an NFS mount point, enter the absolute path of the staging folder on the SAP application server.

Number of Rows Per File

Maximum number of rows that the Data Vault Loader stores in a file in the Data Vault. Default is one million rows.

Data Vault Data Directory

Directory in which the Data Vault Loader creates the archive. Enter the absolute path for the directory. You can set up the directory on a local storage or use Network File System (NFS) to connect to a directory on any of the following types of storage devices:

- Direct-attached storage (DAS)
- Network-attached storage (NAS)
- Storage area network (SAN)

You can specify a different directory for each Data Vault target connection. The directory must be accessible to the ILM application server and the Data Vault Service.

If you select an archive store in the **Archive Store Type** property, the Data Vault Loader archives data to the archive store, not to the location specified in the **Data Vault Data Directory** property. Instead, the Data Vault Loader uses the Data Vault data directory as a staging location when it writes data to the archive store.

Data Vault Archive Folder Name

Name of the folder in the Data Vault in which to store the archived data. The Data Vault folder corresponds to the database in the archive source.

Data Vault Host

Host name or IP address of the machine that hosts the Data Vault Service.

Data Vault Port

Port number used by the ssasql command line program and other clients such as the Data Vault SQL Tool and ODBC applications to connect to the Data Vault. Default is 8500.

Data Vault Administration Port

Port number used by the Data Vault Agent and the Data Vault Administration Tool to connect to the Data Vault. Default is 8600.

Data Vault User

Name of the administrator user account to connect to the Data Vault Service.

You can use the default administrator user account created during the Data Vault installation. The user name for the default administrator user account is *dba*.

Data Vault User Password

Password for the administrator user account.

Confirm Password

Verification of the password for the administrator user account.

Add-On URL

URL for the Data Vault Service for External Attachments component. The Data Vault Service for External Attachments converts external attachments from the archived format to the source format. The URL is required to restore encrypted attachments from the Data Vault to the source database.

Maintain Imported Schema Name

Use schema names from the source data imported through the Enterprise Data Manager.

By default, this option is enabled. The Data Vault Loader creates a schema structure in the Data Vault folder that corresponds to the source schema structure imported through the Enterprise Data Manager. It adds the transactional tables to the schemas within the structure. The Data Vault Loader also creates a `dbo` schema and adds the metadata tables to the `dbo` schema.

The imported schema structure is based on the data source. If source connections contain similar structures but use different schema names, you must import the source schema structure for each source connection. For example, you import the schema structure from a development instance. You export metadata from the development instance and import the metadata into the production instance. If the schema names are different in development and production, you must import the schema structure from the production instance. You cannot use the schema structure imported from the development instance.

If this option is not enabled, the Data Vault Loader creates the `dbo` schema in the Data Vault folder. The Data Vault Loader adds all transactional tables for all schemas and all metadata tables to the `dbo` schema.

Archive Store Type

Storage platform for the Data Vault. Select the **Microsoft Azure Storage** archive store type.

Azure Key

Key for the account that has read and write access to the buckets.

Command

The location of the Microsoft Azure container that hosts the `sct` files.

For example: `https://idv.blob.core.windows.net/idvcontainer1/`

In this example, `https://idv.blob.core.windows.net/` is the URL to the Microsoft Azure Storage account and `idvcontainer1` is the name of the container.

Step 2. Run the Create Archive Folder Job

In Data Archive, run the Create Archive Folder job to create the Data Vault folder and the connection to Microsoft Azure Storage.

The Create Archive Folder job creates the Data Vault folder and adds a Microsoft Azure Storage connection entry to the `ssa.ini` file in the Data Vault Service plug-in in Data Archive. The job sets the name of the Data Vault folder and the Microsoft Azure Storage connection to the name specified in the target connection.

For example, the **Data Vault Archive Folder Name** property in the target connection is set to *AzureSales*. The Create Archive Folder job creates a Data Vault archive folder named *AzureSales* and adds a Microsoft Azure Storage connection named *AzureSales* to the `ssa.ini` file.

The following example shows an entry for a Microsoft Azure Storage connection named *AzureSales* in the `ssa.ini` file:

```
[MAS_CONNECTION Azure_Sales]
AZURE_KEY = xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx
```

Step 3. Copy the Microsoft Azure Storage Connection to Other Data Vault Service Configuration Files

The Microsoft Azure Storage connection definition on the machine that hosts Data Archive must match the Microsoft Azure Storage connection definition on the machines that host other Data Vault Service components. Copy the Microsoft Azure Storage connection definition from the `ssa.ini` file on the machine that hosts Data Archive to the `ssa.ini` files on the machines that host the Data Vault Service and Data Vault Agent.

After you run the Create Archive Folder job, go to the Data Vault Service plug-in directory in Data Archive and find the `ssa.ini` file. Copy the Microsoft Azure Storage connection definition to the `ssa.ini` file on the machine that hosts the Data Vault Service. Additionally, if you have installed the Data Vault Agent on another

machine, copy the Microsoft Azure Storage connection definition to the `ssa.ini` file on the machine that hosts the Data Vault Agent. If you have installed the Data Vault Agent on multiple machines, copy the Microsoft Azure Storage connection definition to the `ssa.ini` file on each machine that hosts a Data Vault Agent.

Step 4. Validate the Connection to Microsoft Azure Storage

Verify the connection to Microsoft Azure Storage from the Data Vault Service and from the Data Vault Service plug-in in Data Archive.

Use the Data Vault Service **ssadrv** administration command to validate the connection to the Microsoft Azure Storage Data Vault store.

On the machine that hosts the Data Vault Service, run the following command:

```
ssadrv –a mas://<Connection Name>/<Path to the Data Vault folder in Microsoft Azure
Storage>/<name of the blob storage bucket>
```

For example, run the following command:

```
ssadrv –a mas://AzureSales/https://idv.blob.core.windows.net/idvcontainer1/
```

In this command:

- `AzureSales` is the name of the Microsoft Azure Storage connection defined in the `ssa.ini` file.
- `https://idv.blob.core.windows.net/` is the URL to the Microsoft Azure Storage account.
- `idvcontainer1` is the name of the container.

You can run the same command on the machine that hosts Data Archive to test the connection from Data Archive to Microsoft Azure Storage.

CHAPTER 7

Datatype Mapping

This chapter includes the following topics:

Datatype Mapping from Source to Archive Overview

Before you can archive data in the Data Vault, each source datatype must be mapped to a Data Vault datatype. The Data Vault Loader job maps most native datatypes to an archive datatype. You must manually map custom datatypes to the appropriate archive datatype.

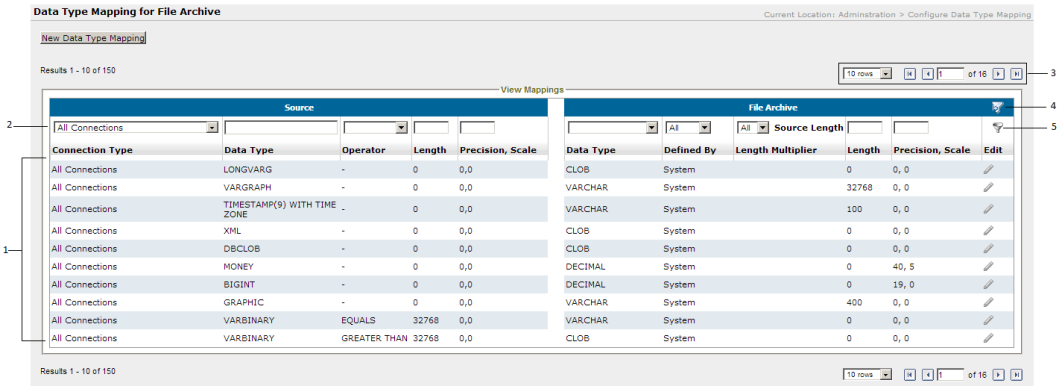
For example, you create a custom datatype named address for alphanumeric values with a length of 100 digits and characters. You must map the address datatype to the varchar datatype in the Data Vault. You can choose to keep the length of this field the same or increase it by a multiple of 100.

Datatype Mapping Interface

You map custom source datatypes to a corresponding archive datatype from the **Data Type Mapping for Data Vault** page.

Access the **Data Type Mapping for Data Vault** from the **Administration** > **Data Type Mappings for Data Vault** menu.

The following image shows the layout of the **Data Type Mapping for Data Vault** page:



1. List of datatypes. Each datatype in the Source section maps to a corresponding data type in the Data Vault section.
2. Filter fields. Enter text or select an option from the drop-down list in one or more fields. Click the filter icon to display the datatypes.
3. Page controls. Set the number of rows to display on a page. Click the arrows to scroll through the list.
4. Clear filter icon. Click this icon to clear any filter settings.
5. Filter icon. After you specify filter criteria in the fields above the column names, click the filter icon to display the datatypes.

The list of datatypes in the Source section has the following columns:

Connection Type

Connection to the source database. Source connections are listed in the drop-down list.

Data Type

The name of the source datatype. Datatypes can be custom or native.

Operator

Specifies if the length of the source data is equal, greater-than, or less-than the number in the **Length** column.

Length

The maximum number of characters or digits in a value.

Precision, Scale

Precision is the maximum number of digits in a value. Scale is the number of digits after the decimal number.

The list of datatypes in the Data Vault section has the following columns:

Data Type

The name of the archive datatype. One or more source datatypes can map to one archive datatype.

Defined By

Specifies if the datatype map was created by the system or the user. Native datatypes mapped during the Data Vault Loader job are listed as system. Custom datatypes that you map are listed as user.

Length Multiplier

You can increase the field length in the Data Vault to a multiple of the field length in the source datatype. For example, if the maximum length of a source field is 25 characters, you can increase the length of this field in Data Vault to 50 by multiplying the source length by 2.

Length

The maximum number of characters or digits in a value.

Precision, Scale

Precision is the maximum number of digits in a value. Scale is the number of digits after the decimal number.

Edit

You can edit a map created by a user.

Data Vault Datatypes

The following table lists the data types available in the Data Vault:

Archive Datatype	Description
BLOB	Use for a large collection of binary data.
CLOB	Use for a large collection of character data.
CHAR	Use for character data with a maximum length of 32768 characters. Use the Length Multiplier options to increase length.
DATE	Use for dates from January 1, 0001 A.D. through December 31, 9999 A.D. Valid range from 0001-01-01 to 9999-21-31
DECIMAL	Use for numbers with decimals. Precision 1 to 120 digits, scale 0-75 digits.
DOUBLE PRECISION	Use for numbers with decimals from -9.99999999999999 E +123 through 9.99999999999999 E+123.
FLOAT	Use for numbers with decimals. For Float(53), values from -9.99999999999999 E+123 through 9.99999999999999 E+123. For Float(21), valid range -9.99999 E+123 through 9.99999 E +123.
INTEGER	Use for whole numbers. Valid range -2,147,483,647 through 2,147,483,647.
REAL	Use for numbers with decimals. Valid range -9.99999 E+123 through 9.99999 E+123.
SMALLINT	Use for whole numbers. Valid range -32,767 through 32,767.
TIME	Use when time data has the following format: HH-MM-SS-NNNN. Valid range 00:00:00 to 23:59:59

Archive Datatype	Description
TIMESTAMP	Use when timestamp data has date and time in the following format: yyyy-mm-dd-HH-MM-SS-NNNN. Valid range 0001-01-01-00.00.00.000000000001 to 9999-12-31-23.59.59.999999999999
VARCHAR	Use for values with number and characters and with a maximum length of 32,768 characters. Use the Length Multiplier options to increase length.

IBM DB2 Source Databases

IBM DB2 databases store character data as binary data. If you want to retire data on an IBM DB2 source database, you must specify a code page in the source connection parameters and create mappings for certain data types.

When you create the IBM DB2 source connection, specify the code page CharsetFor65535=CP950 in the connection parameters. Then create mappings for the VARCHAR() for bit data and LONG VARCHAR for bit data types. If you do not specify the code page in the parameters and create mappings, the file archive loader job retrieves the binary source data as hexadecimal data and the job fails to load data.

You do not need to create a mapping for the CHAR() for bit data type.

Custom Datatypes

If you have a list of the custom datatypes for your source data, then map each custom datatype to an appropriate archive datatype before you run the Data Vault Loader job. If you do not have a list of the custom datatypes, start the Data Vault Loader job. The Data Vault Loader job fails If you have unmapped datatypes. Open the job log to view the list of unmapped datatypes. Create maps for each datatype on the list and restart the Data Vault Loader job.

SAP Datatypes

The File Archive Loader job maps most SAP datatypes to a specific Data Vault datatype.

To view these mappings, click **Adminstration** > **Data Type Mappings for File Archive**. Select **SAP 4.X** from the list of available source connections.

If the SAP application that you want to retire is installed on a Microsoft SQL Server database and includes the LCHR datatype, you must create custom mappings for the LCHR datatype.

The following table lists the datatype mappings to create:

Source Datatype	Source Operator	Source Length	Source Precision, Scale	Data Vault Datatype	Data Vault Length Multiplier	Data Vault Length	Data Vault Precision, Scale
LCHR	Less than	3072		VARCHAR	4x	4x source length	
LCHR	Equal to	3072		VARCHAR	4x	4x source length	
LCHR	Greater than	3072		CLOB			

Unsupported Datatype Mappings

The File Archive Loader job maps most native source data types to a specific Data Vault data type. However, there is no validation to prevent you from creating a custom mapping for a native source data type.

If you map a native source data type to a Data Vault data type other than the data type that the File Archive Loader maps to, the job might truncate or round up the retired data. You might also receive an error if you try to restore the data back to the source database.

The following table describes the results of mapping native source data types to Data Vault data types other than the data types that the File Archive Loader job maps to:

Source Database	Source Data Type	Data Vault Data Type	Result of Mapping
Microsoft SQL Server	BIGINT, MONEY, NUMERIC, or DECIMAL	DOUBLE PRECISION	The File Archive Loader job truncates the source data as it loads the data to Data Vault. If you try to restore the data back to the source database, the restore job fails with an arithmetic overflow error.
Microsoft SQL Server	All data types, if the source data is more than six digits	REAL	The File Archive Loader job truncates the source data past six digits and rounds up the data as it loads the data to Data Vault. If you try to restore the data back to the source database, the restore job fails with an arithmetic overflow error.
Microsoft SQL Server	All data types	FLOAT with scale and precision	If you map to the FLOAT data type, you must set the scale equal to zero. If you do not set the scale to zero, the File Archive Loader job might round up and truncate the source data as it loads to Data Vault. If you try to restore the data back to the source database, the restore job fails with an arithmetic overflow error.

Source Database	Source Data Type	Data Vault Data Type	Result of Mapping
Microsoft SQL Server	TIMESTAMP	N/A	The TIMESTAMP data type in Microsoft SQL Server is used to create a unique binary number for row versioning within Microsoft SQL Server. Microsoft SQL Server does not allow INSERT operations on columns that contain the TIMESTAMP data type. As a result, the TIMESTAMP data type cannot be restored to the source database once it has been archived.
Oracle	NUMBER or REAL	REAL	The File Archive Loader job truncates the source data as it loads the data to Data Vault. If you try to restore the data back to the source database, the restore job fails with an SQL exception.
Oracle	NUMBER or FLOAT	FLOAT	The File Archive Loader job truncates the source data as it loads the data to Data Vault. If you try to restore the data back to the source database, the restore job fails with an SQL exception.

For more information about the data types that the File Archive Loader job maps to, see the *Informatica Data Archive User Guide*.

Mapping a Custom Datatype to an Archive Datatype

1. Click **Administration** > **Data Type Mappings for Data Vault**.

   The **Data Type Mapping for Data Vault** page appears.
2. Click **New Data Type Mapping**.

   The **New Data Type Mapping** page appears.
3. Select a source connection from the **Connection Type** list.
   - Select **All Connections** if you want a custom datatype in more than one connection to map to the same archive datatype.
4. Enter the name of the custom datatype.
5. Optionally, select an option in **Operator** field to indicate if the length is equal, greater-than, or less-than the value in the **Length** field.
6. Enter the maximum number of digits or characters in the **Length** field.
7. Enter the precision and scale in the **Precision, Scale** field for numeric values with a decimal.
8. Select a datatype option from the **Data Type** list in the Data Vault section of the page.

9. Select a multiplier option from the **Length Multiplier** list to increase the number of digits or characters allowed in a field.
10. For numeric values with a decimal, enter the precision and scale in the **Precision, Scale** field.
11. Click **Save**.

Out-of-Range Data Replacement

The Data Vault Loader loads data within the supported Data Vault data type range. If data is not within the supported data type range, the Data Vault Loader job fails.

You can configure the Data Vault Loader to automatically use replacement values for any out-of-range values. Configure the replacement values for the informia.fasLoadProcessingThreads property in the conf.properties file. After you configure the conf.properties file, the Data Vault Loader uses the replacement values for any out-of-range values in future archive jobs.

Configuration for Out-of-Range Data Replacement

Configure out-of-range data replacement in the conf.properties file for the informia.fasLoadProcessingThreads property.

To configure the informia.fasLoadProcessingThreads property, use the following syntax:

```
-k 4 -j 1 -x -lr <data type>:<optype>:<value1>:<value2>,<data type>…
```

Options –k, -j, and -x are default values. Do not change the default values. Use a comma to separate multiple replacement rules.

The following table describes the arguments for the -lr option:

Argument	Description
data type	Data type for which you want to define replacement values. You can specify replacement values for out-of-range values in the following data types: - SMALLINT - INT - REAL - FLOAT - DOUBLE - DECIMAL_<PREC>_<SCALE> For DECIMAL data types, enter the precision and scale. For example, for DECIMAL with precision 32 and scale 16, enter DECIMAL_32_16. - DATE - TIMESTAMP - TIME For the TIME data type, values in the command line must have the colon (:) replaced by a period (.). This is because the -lr notation uses the colon (:) as its delimiter when specifying the data type.
optype	Operation type. Use one of the following values: - R. Replaces a specific out-of-range value with another value. For example, if you enter INT:R:<value 1>:<value 2>, the Data Vault Loader replaces <value 1> with <value 2>. - X. Replaces out-of-range values that are greater than and less than the Data Vault data type range with another value. For example, you enter DOUBLE:X: <value 1>:<value 2>. The Data Vault Loader replaces out-of-range values smaller than the supported data type range with <value 1> and replaces out-of-range values greater than the supported data type range with <value 2>.
value 1	The first replacement value for the operation type. If you specify option R, enter the out-of-range value that you want to replace. If you do not enter a value, then the Data Vault Loader replaces all out-of-range values with the value you configure in value 2. If you specify option X, enter the replacement value for all out-of-range values that are smaller than the Data Vault data type range.
value 2	The second replacement value for the operation type. If you specify option R, enter the value that you want to replace the out-of-range value with. If you specify option X, enter the replacement value for all out-of-range values that are greater than the Data Vault data type range.

Examples for Out-of-Range Data Replacement

You can configure the Data Vault Loader to replace out-of-range values for a specific value, for all values, or for values below or above the supported data type range.

Replacement for specific out-of-range values

You want to archive data that falls outside of the supported Data Vault range for the INTEGER data type. You want to replace out-of-range value 2147483649 with value 100.

Configure the following property:

```
informia.fasLoadProcessingThreads = -k 4 -j 1 -x -lr INT:R:2147483649:100
```

Replacement for all out-of-range values

You want to archive data that falls outside of the supported Data Vault range for the SMALLINT data type. You want to replace all out-of-range values with value 100 for data type SMALLINT.

Configure the following property:

```
informia.fasLoadProcessingThreads = -k 4 -j 1 -x -lr SMALLINT:R:100
```

Replacement for below minimum and above maximum out-of-range values

You want to archive data that falls outside of the supported Data Vault range for the DOUBLE data type. The Data Vault range for the DOUBLE data type is -9.9E +123 to 9.9E+123. You want to replace all out-of-range values that are less than the minimum data type range with 0.0. You want to replace all out-of-range values that are greater than the maximum data type range with 1.0E+120.

Configure the following property:

```
informia.fasLoadProcessingThreads = -k 4 -j 1 -x -lr DOUBLE:X:0.0:1.0E+120
```

CHAPTER 8

Database Optimization

This chapter includes the following topics:

Database Optimization Overview

You can configure Data Archive to use database native utilities for data movement. Use database native utilities to optimize the performance of exporting data from source databases and importing data into target databases. You can use native utilities for data movement for IBM DB2, Oracle, and Teradata databases.

You can use native utilities for data import and export for IBM DB2 and Oracle. You can use native utilities for data export for Teradata. If you do not configure native utilities, then Data Archive uses JDBC for data movement.

IBM DB2 Native Utilities

You can configure Data Archive to use IBM DB2 native utilities for data movement. You may want to use the IBM DB2 native utilities to increase the archive job performance. By default, Data Archive uses JDBC for data movement.

You can use the native IBM DB2 utilities to export data from the source and import data into the target destination. You can use the export utility or the IBM DB2 HPU utility to export data. You can use the load client utility or the import utility to import data.

When you configure Data Archive to use IBM DB2 native utilities, the job log indicates which utility the archive job uses. The job log does not include row count information for the Generate Candidates and Copy to Destination steps. The row count information is not available because the extract and import process is a bulk process. However, the job log includes a confirmation that the rows are successfully inserted.

Utilities for Data Export

You can use the export utility or the IBM DB2 HPU utility to export data. Configure the utility that you want to use for the export based on the installed utilities and the platform that hosts the IBM DB2 database.

You can use the following utilities to export data:

Export Utility

By default, if you configure Data Archive to use IBM DB2 native utilities, Data Archive uses the IBM DB2 export utility to export data from the source. You can use the export utility for all platforms.

High Performance Unload Utility (IBM DB2 HPU)

For IBM DB2 on Linux, UNIX, or Windows, you can use the IBM DB2 HPU utility to export data from the source.

Utilities for Data Import

You can use the load client utility or import utility to import data. The utilities for data import are platform-dependent. The utility that you use for the import depends on the platform that hosts the IBM DB2 database.

You can use the following utilities to import data:

Load Client Utility

For IBM DB2 on AIX, Linux, UNIX, or Windows, Data Archive uses the IBM DB2 load client utility to import all tables that do not contain LOB datatypes to the target. You must configure a temporary LOB directory if you want Data Archive to use the load client utility to import tables with LOB datatypes.

Import Utility

For IBM DB2 on AIX, Linux, UNIX, or Windows, Data Archive uses the IBM DB2 import utility to import tables with LOB datatypes if you do not configure a temporary LOB directory. Note that the import utility may decrease the import performance because the utility generates logs on the IBM DB2 database to undo the import.

For IBM DB2 on AS/400 and z/OS, Data Archive uses the import utility to import all tables.

Export and Import Parameters

Optionally, you can add IBM DB2 parameters to the export or import process. By default, no additional parameters are required to import or export data.

You can add IBM DB2 parameters to the script that the export utility or the import utility uses to import or export data. You can use any IBM DB2 parameter that the import and export utilities support. The parameters that you add depend on your database administration policies for importing or exporting data. For example, you may want to add additional parameters to create error files or to specify no logging.

To add IBM DB2 parameters to the import or export commands, add properties in the `conf.properties` file. You can add the commands in predefined places of the import and export process. For the import command, you can add a parameter after the load file statement, before the insert statement, and after the insert statement. For the export command, you can add a parameter after the unload file statement and after the select statement.

Set Up IBM DB2 Native Utilities

Configure Data Archive to use IBM DB2 native utilities for data movement.

To use IBM DB2 native utilities, IBM DB2 Connect must be installed on the machine that hosts the ILM application server. To use the IBM DB2 HPU utility for data export, the IBM DB2 HPU utility must also be installed on the machine that hosts the ILM application server.

1. Create connections to the source and target databases in IBM DB2 Connect.
2. Create source and target connections in Data Archive.

When you create the connections, add the connection name from IBM DB2 Connect to the IBM DB2 Connect Data Source name.

3. Configure the IBM DB2 native utilities properties in the `conf.properties` file.

Data Archive uses the IBM DB2 native utilities for all archive jobs that archive from IBM DB2 databases.

Teradata Native Utilities

Data Archive uses Teradata native utilities for data movement. Teradata native utilities increase the archive or retirement job performance when the job extracts data from Teradata databases. The job uses Teradata native utilities in the Copy to Destination step.

By default, all archive, retirement, and restore jobs use Teradata JDBC FastExport. Optionally, you can configure Data Archive to use Teradata Parallel Transporter when you archive data to the Data Vault. The job log indicates which utility the job uses to export data.

Utilities for Data Export

You can use Teradata Parallel Transporter or Teradata JDBC FastExport to export data from Teradata databases.

You can use the following utilities to export data:

Teradata Parallel Transporter

For Teradata on Linux, the archive or retirement job can use Teradata Parallel Transporter to export data from Teradata to the Data Vault. You configure the Teradata source connection properties to enable Teradata Parallel Transporter. The job uses Teradata Parallel Transporter to export data from all tables except for tables that include binary or large object datatypes such as BLOB and CLOB.

Teradata JDBC FastExport

By default, all archive, retirement, and restore jobs use Teradata JDBC FastExport to export data from Teradata. If you enable Teradata Parallel Transporter, the job uses Teradata JDBC FastExport to export data from tables that have binary or large object datatypes such as BLOB and CLOB. If the length of all column data types exceeds 8000 characters, the job uses JDBC FastExport.

Data Export Process for Teradata Parallel Transporter

For Teradata sources, the archive or retirement job can use Teradata Parallel Transporter to export data from Teradata to the Data Vault.

When you run an archive or retirement job, the job checks if Teradata Parallel Transporter is enabled. If Teradata Parallel Transporter is enabled, the job may use a combination of utilities to export data. Teradata Parallel Transporter can only process data types that can be serialized into a string. This includes the following supported data types:

- BYTEINT
- SMALLINT
- INTEGER
- BIGINT
- DECIMAL

- NUMERIC
- CHAR
- VARCHAR
- DATE
- TIMESTAMP
- TIMESTAMP WITH TIMEZONE

The job determines the utility to use based on the data types in the source table. The job uses Teradata Parallel Transporter to export data from tables that do not include binary or large object data types, such as BLOB and CLOB. The job uses Teradata JDBC FastExport to export data from tables that include binary or exotic data types.

When you run an archive or retirement job, the job uses a script template to create script files to submit query requests to Teradata Parallel Transporter. Teradata Parallel Transporter rewrites the query requests to optimize the query for maximum parallelism and maximum throughput. Then, Teradata Parallel Transporter writes the query results to BCP files.

Teradata Parallel Transporter saves the BCP files to the BCP file staging directory that is configured in the Data Vault target connection. The Data Vault Loader job moves the BCP files from the BCP file staging directory to the Data Vault.

Teradata Parallel Transporter is not supported in non-bulk mode.

Optionally, Teradata Parallel Transporter can compresses the BCP files. By default, Teradata Parallel Transporter does not compress the BCP files to optimize performance. You may need to compress the BCP files if you have a limited amount of BCP file staging space. You configure BCP compression in the Teradata source connection.

Informatica recommends that you upload Teradata drivers from the source itself so they can be used in the source connection.

Script Template

When you run an archive or retirement job, the job uses a script template to create one script for each table the job extracts data from. The script template includes a scripting language that is supported by Teradata. The script template includes parameters that the job fills in with values at run time.

The script template, TPT_EXTRACT_SCRIPT_TEMPLATE, is packaged as an Enterprise Data Manager API. Optionally, you can copy the script template and customize the scripting language and the parameters. The name of the script template is configured in the Teradata source connection.

The job uses the script template to substitute values for the parameters and creates one script for each table. The job uses `<Job ID>_<Table Name>.bcp` as the naming file convention to create the script. The job saves the scripts in the `<Data Archive installation directory>/webapp/tmp` directory. The scripts are temporary files that Teradata Parallel Transporter uses to process data for individual tables.

The scripts include query requests. Teradata Parallel Transporter optimizes and rewrites the query requests and runs the scripts. Teradata Parallel Transporter creates one BCP file for each table. The BCP file includes the query results. After the Teradata Parallel Transporter creates the BCP files, the job deletes the scripts.

Data Export Process for Teradata JDBC FastExport

By default, archive, retirement, and restore jobs use Teradata JDBC FastExport to export data from Teradata. If you enable Teradata Parallel Transporter for an archive or retirement job, the job uses Teradata JDBC FastExport to export data from tables that have binary or large object datatypes such as BLOB and CLOB.

When you run a job, the job creates an SQL select PreparedStatement for the extract query. The job creates the URL connection string with parameters for Teradata JDBC FastExport.

The job creates the following URL connection string:

```
jdbc:teradata://$HOST/DATABASE=
$SID,TMODE=ANSI,CHARSET=UTF16,TYPE=FASTEXPORT,SPOOLMODE=NOSPOOL,SESSIONS=
$MAX_PARALLEL_SESSIONS
```

The job uses the Teradata source connection properties to populate the $HOST, $SID, and $MAX_PARALLEL_SESSIONS parameters.

If the target is a database, Teradata JDBC FastExport runs the query, streams the results in memory, and writes to the database.

If the target is the Data Vault, Teradata JDBC FastExport runs the query and writes the query results to compressed BCP files. Teradata JDBC FastExport saves the BCP files to the BCP file staging directory that is configured in the Data Vault target connection. The Data Vault Loader job moves the BCP files from the BCP file staging directory to the Data Vault.

Setting Up Teradata Parallel Transporter

Configure Data Archive to use Teradata Parallel Transporter to export data from Teradata to the Data Vault. If you do not set up Teradata Parallel Transporter, all archive and retirement jobs use Teradata JDBC FastExport.

To use Teradata Parallel Transporter for data export, the Teradata Parallel Transporter client must be installed on the machine that hosts the ILM application server.

1. Optionally, customize the script template. By default, no additional configuration is required. You may want to customize the script template if you want to use additional parameters or change the scripting language.

   To customize the script template, copy the default script template, TPT_EXTRACT_SCRIPT_TEMPLATE, and make changes to the copied template. You can find the default script template in the Enterprise Data Manager APIs.
2. Configure the following Teradata Parallel Transporter properties for the Teradata source connection:
   - Maximum Parallel Sessions
   - Use Teradata Parallel Transporter
   - Template Name
   - Compress File

   Enable Teradata Parallel Transporter and configure the maximum amount of parallel sessions and BCP file compression. If you customized the script template, you must update the template name property.

Related Topics:

- "Teradata Source Connections" on page 110

CHAPTER 9

SAP Application Retirement

This chapter includes the following topics:

SAP Application Retirement Overview

You can use Data Archive to retire SAP applications to the Data Vault. When you retire an SAP application, Data Archive retires all standard and custom tables in all of the installed languages and SAP clients. Data Archive also retires all attachments that are stored within the SAP database and in an external file system or storage system.

In SAP applications, you can use the Archive Development Kit (ADK) to archive data. When you archive within SAP, the system creates ADK files and stores the files in an external file system. When you retire an SAP application, you also retire data stored in ADK files.

After you retire the application, you can use the Data Validation Option to validate the retired data. You can use the Data Discovery portal or other third-party query tools to access the retired data.

SAP Application Retirement Architecture Options

The SAP application server cannot write data from attachments or special tables such as cluster and pool tables, directly to an external drive. For this reason, you must provide a connection between the SAP application server and the Data Archive staging area.

You can connect the SAP application server and the Data Archive staging area in one of the following ways:

- Use FTP

- Use NFS

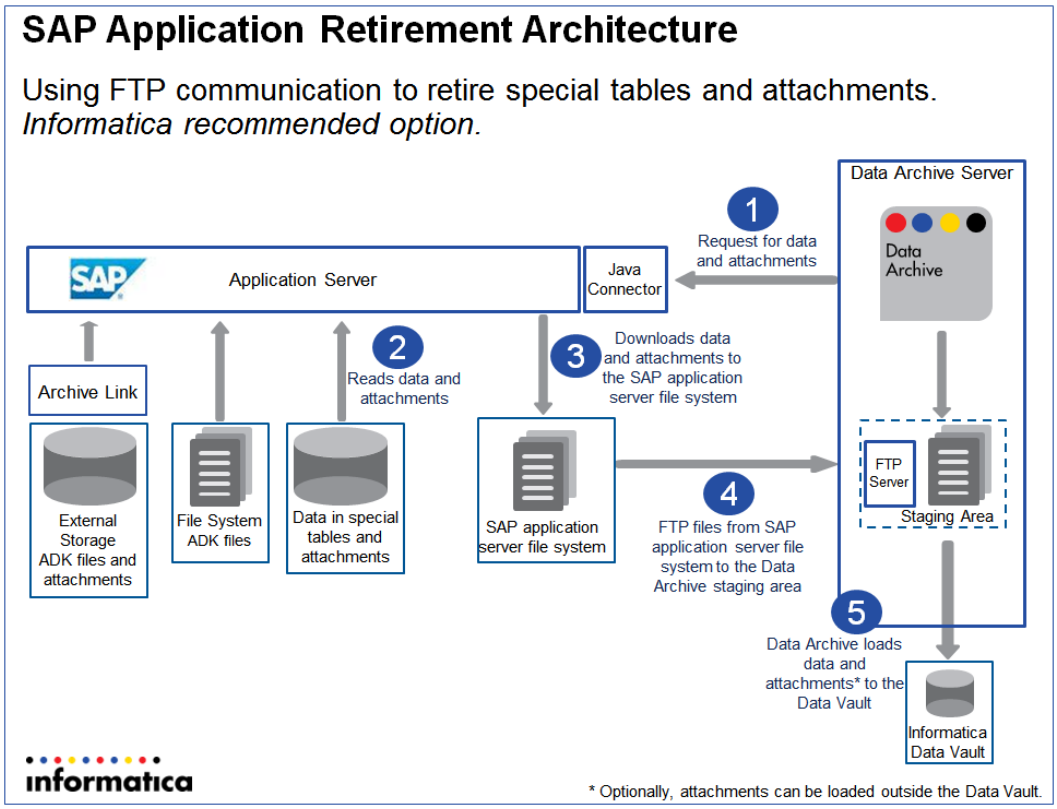
Informatica recommends the FTP option.

FTP Option

To use the FTP option, install the FTP server on the same location as the Data Archive staging area. Specify the SAPFTP properties. Provide the FTP connection details such as location and log-in credentials when you configure the source connection.

When you retire an SAP application, Data Archive first moves the attachments and the data from the special tables to the file system on the SAP application server. Then, Data Archive pushes the files to the FTP folder in the staging area.

The following image shows the SAP application retirement architecture with the FTP option:

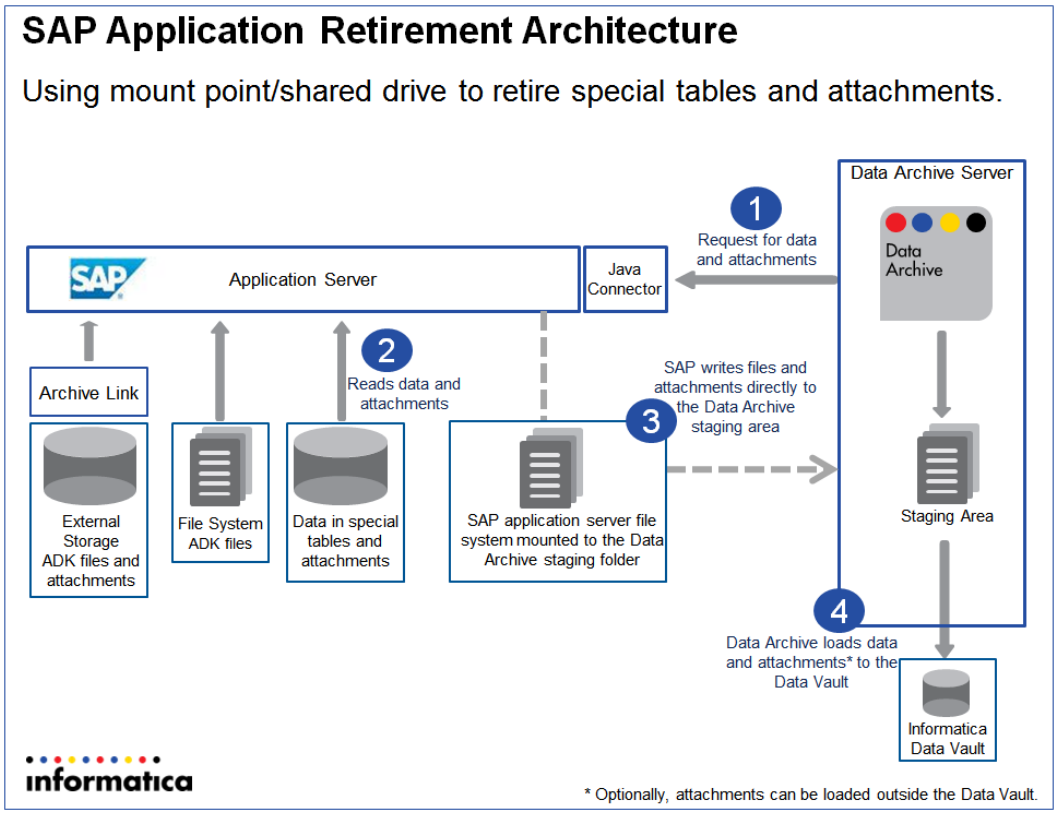


NFS Option

To use the NFS option, create a mount point on the file system of the SAP application server. Specify the mount point location when you configure the source and target connections.

When you retire an SAP application, Data Archive loads the attachments and the data from the special tables to the Data Archive staging area through the shared file system.

The following image shows the SAP application retirement architecture with the FTP option:



Setting Up SAP Application Retirement

You can configure Data Archive to retire data from SAP applications.

Before you begin, verify the following prerequisites:

- You have the required authorizations.
- The size of the staging folder on the SAP application server is at least four times the size of the data that you plan to retire in a single retirement project.
- If you are upgrading from a previous version of the SAP accelerator, delete the "SAP" folder in both the Data Archive Installation directory and in the application (Visualization > Reports and Dashboards) before you upgrade the accelerator. Also delete INFA_REPORTS from the Data Archive installation directory.

Complete the following tasks to setup SAP application retirement:

1. Install the SAP Java Connector.
2. Apply the SAP transports.
3. Assign roles.
4. Configure the `conf.properties` file.
5. Set up the FTP or NFS connection.

Related Topics:

- "SAP Application Retirement Privileges" on page 58

Step 1. Install the SAP Java Connector

Download the SAP Java Connector, version 3.0.12 or later, from the SAP Marketplace and install the SAP Java Connector on the machine that hosts Data Archive. The retirement job uses the SAP Java Connector to log in to the SAP application to read data from special tables, ADK files, and attachments.

1. Use the following URL to access the SAP Marketplace:

   http://service.sap.com/connectors

   Note: SAP requires authentication to the SAP Marketplace. If you do not have the credentials to access SAP Marketplace, contact Informatica Global Customer Service and create a ticket for further guidance. Provide the following information about the server on which Data Archive is installed:

   - Operating system
   - 32- or 64-bit architecture
2. Navigate to the SAP Java Connector download page.
3. Download version 3.0.12 or later. Choose the file for the platform that hosts Data Archive.

   For example, the SAP system is on Windows and Data Archive is on Linux. Choose the Linux version.
4. Install the SAP Java Connector on the machine that hosts Data Archive.

   The installation instructions depend on the operating system. Follow the instructions provided by SAP for the operating system of the machine that hosts Data Archive.

   Tip: You can find the installation instructions after you unzip the downloaded file. The `Readme.txt` files provide instructions on how to access the documentation. Review the `Readme.txt` files in the following directories:

   ```
   <SAP Java Connector>
   <SAP Java Connector>/docs/jco
   ```
5. After you install the SAP Java Connector, copy the `sapjaco.jar` file from the root directory of the SAP Java Connector installation and paste to the following location:

   ```
   <Data Archive installation>/webapp/WEB-INF/lib
   ```
6. Restart the ILM application server.

Step 2. Apply the SAP Transports

To retire SAP applications, you must install SAP transports to the SAP application that you want to retire. After you install the SAP Retirement Accelerator, the transport files are available in the Data Archive installation directory.

The transports include a role and remote function calls (RFCs) to allow Data Archive to run the retirement job on the SAP database. The transports create objects in the ZINFA_RETIREMENT package. The transports do not modify any standard SAP objects.

You can access the transports in the following Data Archive directory:

```
<Data Archive Installation>\optional\SAP_retirement_transports.zip
```

The transports you install depend on the SAP module and version. The zip file includes a readme file that lists the required transports for each SAP module and version.

Import the transports as client-dependent. Import the transports in the client that contains the SAP system user with RFC connection authorizations.

Step 3. Assign Roles

Assign the user that runs the retirement job to a retirement-specific role. You provide the user in the source connection.

The ZINFA_RETIREMENT_PREPARATION role is included in the SAP transports. After you apply the SAP transports, assign the role to the user that will run the retirement project. Assign to any user that has remote function call (RFC) connection authorization. The role includes authorization to call remote function modules to access data from the SAP data dictionary.

Step 4. Configure Conf.Properties

Change the BCP file delimiters in the `conf.properties` file to unique values that are not included in the SAP table data. If you use values that are included in the SAP table data, then the retirement job generates BCP files that are not valid. You can find the `conf.properties` file in the root Data Archive installation directory.

The following table contains examples of the properties and the values to configure:

Property	Value
informia.bcp.columnSeparator	^#^
informia.bcp.rowSeparator	@#@#

Note: If the SAP application is installed on a Microsoft SQL Server database, you must also enable the following property in the `conf.properties` file by setting the value to "Y":
`informia.sqlServerVarBinaryAsVarchar=Y`

Step 5: Set up the FTP or NFS Connection

SAP can write data from special files and attachments only to the SAP application server. To facilitate SAP application retirement, you must create a connection between the SAP application server and the Data Archive staging area.

Use one of the following options to create the connection between the SAP application server and the Data Archive staging folder:

- FTP
- NFS Mount

Creating the FTP Connection

Install the FTP server and user. Specify the connection properties in the source and target connections.

To create the FTP connection, perform the following high-level steps:

1. Install the FTP server in the same location as the Data Archive staging folder.
2. Create the FTP user.
3. Assign the FTP user with read, write, append, and delete permissions to the FTP folder.
4. Create and maintain an entry in the SAPFTP_SERVERS table.
   For instructions on configuring SAPFTP and the list of SAP supported versions, see *SAP Note 1605054*.
5. Set the FTP property **Connection Timeout** to a value between 1,200 and 2,000 seconds.
6. Configure the Data Archive source connection.

   Specify the required source connection properties. In addition, specify the properties specific to SAP application retirement.

   The following table lists the required properties for SAP application retirement with the FTP option:

Property	Description
Source/Staging Attachment Location	The directory path of the staging folder on the SAP application server. SAP downloads the attachments and saves the BCP files to this folder. Enter the full path of the location. For example: `\\10.1.10.10\interfaces\CCO\`. Depending on the OS, the path must end with "/" or "\".
SAP Fetch Size	Number of rows that the retirement job extracts at a time from the SAP cluster and pool tables to write to the BCP file.
SAP Host	Host of the SAP application that you want to retire.
SAP Client	Client in which the user logs in. Note that all clients in the SAP application are retired.
SAP System Number	System number in which the user logs in.
SAP Language	Language in which the user logs in. Note that all languages in the SAP application are retired.

Property	Description
SAP User	User that logs in to the SAP application. The user must be assigned to the ZINFA_RETIREMENT_PREPARATION role and include RFC connection authorizations.
SAP User Password	Password for the SAP user.
FTP User	User name to connect to the FTP server.
FTP Password	Password for the FTP user.
FTP Host	Host name of the FTP server.
FTP Port	Port number of the FTP server. Default port is 21.
FTP Folder Location	Name of the FTP folder on the Data Archive server. For example, `ERP\`. Depending on the OS, the folder name must end with "/" or "\".

7. If you want Data Archive to save the BCP files in compressed file format, enable the check box for the **Compressed** property in the source connection.

   If enabled, Data Archive compresses BCP files and saves them as `.gz` files in the Data Archive staging folder. If not enabled, Data Archive saves BCP files as `.bcp` files.

8. Configure the Data Vault target connection.

   For the **Staging Directory** property, enter the full directory path of the FTP folder on the Data Archive server. For example, `C:\user\FTP\ERP\`.

9. Create the SAP application retirement project.

   For more information, see the chapter "SAP Application Retirement" in the *Informatica Data Archive User Guide*.

Creating the NFS Mount

Use an NFS mount to create a shared file system between the SAP application folder and the Data Archive staging folder. Specify the connection properties in the source and target connections.

To create the NFS mount, perform the following high-level steps:

1. In the SAP application server file system, set up an NFS mount point to the Data Archive staging folder.
2. Set up the user with access to the Data Archive staging folder. If the SAP system is on Windows, that user must have the standard read, write, append, and delete permissions to this folder.
3. Configure the Data Archive source connection.

   Specify the required source connection properties. In addition, specify the properties specific to SAP application retirement.

The following table lists the required properties for SAP application retirement with the FTP option:

Property	Description
Source/Staging Attachment Location	The directory path of the staging folder on the SAP application server. SAP downloads the attachments and saves the BCP files to this folder. Enter the full path of the location. For example, `\\10.1.10.10\interfaces\CCO\`. **Note:** You must specify the same directory path for the **Staging Directory** target connection property.
SAP Fetch Size	Number of rows that the retirement job extracts at a time from the SAP cluster and pool tables to write to the BCP file.
SAP Host	Host of the SAP application that you want to retire.
SAP Client	Client in which the user logs in. Note that all clients in the SAP application are retired.
SAP System Number	System number in which the user logs in.
SAP Language	Language in which the user logs in. Note that all languages in the SAP application are retired.
SAP User	User that logs in to the SAP application. The user must be assigned to the ZINFA_RETIREMENT_PREPARATION role and include RFC connection authorizations.
SAP User Password	Password for the SAP user.

4. If you want Data Archive to save the BCP files in compressed file format, enable the check box for the **Compressed** property in the source connection.

   If enabled, Data Archive compresses BCP files and saves them as `.gz` files in the Data Archive staging folder. If not enabled, Data Archive saves BCP files as `.bcp` files.

5. Configure the Data Vault target connection.

   For the **Staging Directory** property, enter the same directory path that you entered for the **Source/Staging Attachment Location** source connection property.

6. Create the SAP application retirement project.

   For more information, see the chapter "SAP Application Retirement" in the *Informatica Data Archive User Guide*.

CHAPTER 10

z/OS Source Data Retirement

This chapter includes the following topics:

z/OS Source Data Retirement Overview

PowerExchange users can connect to a z/OS data source in order to retire nonrelational data.

Data Archive uses a JDBC-ODBC bridge to connect to the z/OS source data through the ODBC drivers that are included with PowerExchange.

You can retire the following types of nonrelational z/OS data:

- Adabas
- C-ISAM
- Datacom
- IDMS
- IMS
- VSAM
- Complex, sequential flat file

Note: The PowerExchange ODBC interface does not support the complex SQL that is required for live archiving.

Implementation Example

Data Archive can use PowerExchange ODBC drivers to access nonrelational data on z/OS.

One possible implementation might include the following systems and components:

- A z/OS system with the nonrelational data sources, such as IMS and VSAM. You will install PowerExchange 9.1.0 HotFix 4 or later on this system.
- A Linux, UNIX, or Windows system that has the following type of storage and software:
  - Informatica Data Archive engine 6.1.1 or later
  - Data Vault Service
  - NFS or local storage for the encapsulated BCP files

  On this system you will install PowerExchange 9.1.0 HotFix 4 or later. You will also copy the DataDirect Driver Manager and define PowerExchange ODBC data sources.
- A database repository for the ILM and PowerCenter schemas.

Prerequisites

Before you begin a z/OS source data retirement project, you must complete PowerExchange prerequisites.

Complete the following tasks before you begin the retirement project:

1. Install and configure PowerExchange on z/OS.
2. Install and configure PowerExchange on the Data Archive server.
3. Install and configure PowerExchange on Windows.

Install and Configure PowerExchange on z/OS

To install PowerExchange on z/OS, run the MVS Installation Assistant.

After you install PowerExchange, edit the DBMOVER configuration file, configure the PowerExchange Listener JCL, and start the PowerExchange Listener.

Install PowerExchange on z/OS

Install PowerExchange on the z/OS system where the source data for the retirement project is located.

To install PowerExchange on z/OS, run the MVS Installation Assistant. For detailed instructions, see the *PowerExchange Installation Guide*.

Edit the DBMOVER Configuration Member on the z/OS System

Verify that the DBMOVER configuration member includes a LISTENER statement.

You can also include the following statements in the DBMOVER configuration member to minimize data map I/O:

- DMXCACHE_MAX_MEMORY_MB20
- DM_SUBTASK=R

When Data Archive accesses PowerExchange data maps through the PowerExchange Listener, every data map in the DATAMAPS KSDS is opened and processed. The DMXCACHE_MAX_MEMORY_MB and DM_SUBTASK statements configure the PowerExchange Listener on z/OS to use data maps caching and a data maps subtask in read-only mode.

Configure the PowerExchange Listener JCL on z/OS

Configure the PowerExchange Listener JCL before starting the PowerExchange Listener for the first time.

For more information, see the *PowerExchange Bulk Data Movement Guide*.

Install and Configure PowerExchange on the Data Archive Server

Install and configure PowerExchange on the same server as Data Archive.

Install PowerExchange on the Data Archive Server

On the Data Archive server, install PowerExchange 9.1.0 HotFix 4 or later. Use the PowerExchange installation instructions for the operating system type and bit-level.

If the Data Archive server runs on a 64-bit machine, use the PowerExchange 64-bit installation executable or tar file. If the Data Archive server runs on a 32-bit machine, use the PowerExchange 32-bit executable or tar file.

If you are installing on Windows, install the PowerExchange ODBC drivers in the Windows registry:

- On Windows 32-bit systems other than Vista, enter the following command at the command prompt:

  ```
  dtlodbci add
  ```

- On Windows Vista, update the dtlodbci.exe file properties to select **Run this program as an administrator**, and then enter the following command at the command prompt:

  ```
  dtlodbci add
  ```

- On 64-bit Windows, update the registry manually.

For detailed installation instructions, see the *PowerExchange Installation Guide*.

Edit the DBMOVER Configuration File on the Data Archive Server

Edit the DBMOVER configuration file.

Include a NODE statement to define the TCP/IP host name and port that PowerExchange uses to contact the PowerExchange Listener on the machine where the source data for the retirement project is located. Use the following format:

```
NODE=({node_name|node1}
,TCPIP
,host_name
,{port|2480}
[,{send_bufsize|65536}]
[,{receive_bufsize|65536}]
[,{send_size|4096}]
[,receive_timeout]
[,{SSL|ZOSSSL}]
[,service_name]
)
```

For more information, see the *PowerExchange Reference Manual*.

Configure Environment Variables on Linux or Unix

If the Data Archive server is running on Linux or UNIX, edit the path and library path environment variables on the Linux or UNIX system to point to the directories where the files and libraries reside.

You must also set the PWX_HOME environment variable so that PowerExchange can locate various run-time components.

The following table lists the Linux and UNIX environment variables that you need to set:

Variable	Platform	Example
PWX_HOME	All	WX_HOME=/usr/pwxuser/v910
PATH	All	PATH=/usr/pwxuser/v910
LD_LIBRARY_PATH	Sun	LD_LIBRARY_PATH=/usr/pwxuser/v910
LD_LIBRARY_PATH	Linux	LD_LIBRARY_PATH=/usr/pwxuser/v910
SHLIB_PATH	HP-UX 10/11	SHLIB_PATH=/usr/pwxuser/v910

To verify that the environment variables are correct, log off and log in again.

Note: After you create ODBC data sources on UNIX, you must update the library path and ODBCINI environment variables.

Install and Configure PowerExchange on Windows

On the Windows system on which you will run the PowerExchange Navigator, install PowerExchange and edit the DBMOVER configuration file.

Install PowerExchange on Windows

On the Windows machine, run the PowerExchange 32-bit executable file to install PowerExchange. You must use the Windows 32-bit executable to install the PowerExchange Navigator, even on a Windows 64-bit machine.

For detailed installation instructions, see the *PowerExchange Installation Guide*.

Edit the DBMOVER Configuration File on the Windows System

Edit the DBMOVER configuration file.

Include a NODE statement to define the TCP/IP host name and port that PowerExchange uses to contact the PowerExchange Listener on the machine where the source data for the retirement project is located.

For more information, see the *PowerExchange Reference Manual*.

Step 2. Create PowerExchange Data Maps

Create a PowerExchange data map for each data source that you want to retire.

Before you create the data maps, select to use a two-tier naming convention for the data maps. To use a two-tier naming convention, click **Options > Preferences**, and on the Data Maps tab in the **Preferences** dialog box, select the **Use 2-tier names** option.

For more information about creating data maps, see the *PowerExchange Navigator User Guide*.

Step 3. Configure PowerExchange ODBC Data Sources on the Data Archive Server

Configure PowerExchange ODBC data sources on the Linux, UNIX, or Windows system where the Data Archive server and PowerExchange are installed.

Configure ODBC Data Sources on Linux or UNIX

Copy the DataDirect Driver Manager to the server where PowerExchange and the ILM Engine are installed. Then configure environment variables and create the odbc.ini file.

1. Copy the DataDirect Driver Manager from the computer where PowerCenter is installed to the server where PowerExchange and the ILM Engine are installed.
   a. Copy the following directory to the system with the ILM Engine:
   `Powercenter_installation_directory/ODBC6.1`
   b. Tar the entire directory, copy it using FTP, and then untar the directory on the target system.
2. Add the ODBC6.1/lib directory to one of the following environment variables:
   - LD_LIBRARY_PATH on Sun or Linux
   - SHLIB_PATH on HP-UX
3. Create the odbc.ini file and set the ODBCINI environment variable to point to this file. Include the following line for each data source: `Compatibility=ILMBRIDGE`
4. For HP-UX systems, change the driver name in the `odbc.ini` file from `libdtlodbc.so` to `libdtlodbc.sl`.

For more information, see the *PowerExchange Reference Manual*.

The following example shows a section of an `odbci.ini` file for one data source:

```
[nrdb2_mhz8902]
Driver=/pwx_installation_directory/source/libdtlodbc.so
Description=dBase
Database=
DBType=nrdb2
Location=zos_listener_name
DBQual1=
DBQual2=
MaxRows=0
Compress=N
Encrypt=N
ConfirmWrite=N
PaceSize=0
InterpretASRows=N
BulkLoad=N
```

```
DeleteTempFiles=0
LoadOptions=0
JCLTemplate=
CTLTemplate=
ModeType=0
ModeTime=0
Time=0
Space=
pace=0
SecSpace=0
Compatibility=ILMBRIDGE
```

Configure ODBC Data Sources on Windows

Configure ODBC data sources on Windows with the Data Source Administrator Wizard.

1. Copy the DataDirect Driver Manager from the computer where PowerCenter is installed to the server where Power Exchange and the Data Archive engine are installed. Copy the following directory to the system with the Data Archive engine: `PowerCenter_installation_directory/ODBC6.1`
2. To start the 32-bit **ODBC Data Source Administrator** wizard, perform one of the following actions:
   - On a Windows 32-bit system, double-click **Administrative Tools** in the **Control Panel.** In the **Administrative Tools** window, double-click **Data Sources (ODBC).**
   - On a Windows 64-bit system, enter the following command at the command prompt: `%windir%\SysWOW64\odbcad32.exe`

   The ODBC Data Source Administrator wizard appears.
3. On the **System DSN** tab, click **Add**.
4. In the **Create New Data Source** dialog box, select the **Informatica PowerExchange** driver from the list of available drivers and click **Finish**.

   The **PowerExchange Data Source** wizard appears.
5. On the PowerExchange Data Source tab, enter the following information:
   - At **Location**, enter the name of the PowerExchange Listener, as defined in the NODE statement in the DBMOVER configuration file.
   - At **Type**, select **NRDB2.**
   - At **Local Codepage**, select **UTF-8 encoding of Unicode**.
   - Depending on the data source type that you select, enter values for other available properties.
6. On the **General** tab, enter the following values at **Integration Mask**: `ILMBRIDGE,CPOOL`

   ILMBRIDGE sets options that are required by the JDBC-ODBC bridge. CPOOL enables PowerExchange connection pooling.
   For all data source types, you can optionally enter additional information.
7. Click **OK.**

   The ODBC data source appears in the **System Data Sources** list on the **System DSN** tab in the **ODBC Data Source Administrator** wizard.
8. Click **OK**.

For more information, see the *PowerExchange Reference Manual.*

Step 4. Import z/OS Metadata

Use the Enterprise Data Manager (EDM) to import metadata for each z/OS data source.

1. To launch the Enterprise Data Manager, log in to Data Archive and select **Accelerators > Enterprise Data Manager**.
2. To create an application version for the data source, right-click **Custom Apps Only** in the navigator pane, select **New Application Version**, and enter an application name and description.
3. To import PowerExchange metadata for the application version that you created, complete the following tasks:
   a. Highlight the object and click **File > Import Metadata from Database**. The **Connect to Import Metadata from Database** dialog box appears.
   b. If you defined a connection, select it and enter your z/OS user name and password.
   c. If you are defining a conection, enter the data source name and the z/OS user name and password.

   The **Database Mining** wizard appears.
4. Select the schema and list of tables for the data source.
5. At the prompt to mine child tables of the selected tables, click **No**.

   Note: Because the PowerExchange NRDB2 access method does not support foreign key relationships, you cannot mine child tables. By clicking **No**, you ensure that the PowerExchange Listener on z/OS does not spawn multiple subtasks that would return empty result sets. If you enabled the ILMBRIDGE or NOFKEYS Compatibility option for the PowerExchange ODBC connection, the PowerExchange Listener does not spawn multiple empty subtasks, regardless of what you select for this option.

Step 5. Define and Run the Retirement Project

Define and run the retirement archive project. Running the project copies the nonrelational source data to a target database or the Data Vault.

To define a retirement archive project, specify the source and target connections. Then define the retirement project details.

1. When you define the source connection, enter the following information:
   - At **Connection Type**, select **POWER_EXCHANGE**.
   - Enter the **Data Source Name** as defined in the odbc.ini.file.
   - For the **Admin**, **Application**, and **Staging** login names and passwords, enter a valid z/OS user ID and passwowrd.
   - At **JDBC Fetch Size**, enter 0.
   - Clear the **Use Staging** check box.
2. Define the target connection, if you have not already defined it.
3. Create and define the retirement project from the **Workbench > Manage Retirement Projects** menu.
4. Run the retirement project.

CHAPTER 11

Seamless Data Access

This chapter includes the following topics:

Seamless Data Access Overview

The Seamless Data Access Layer provides a combined view of production data and data archived in the history database. You can choose to enable a combined view of production and archive data or a view of just the archive data. A view of the archive data is called the archive view.

Complete the following tasks to make the combined and archive view available to your application users.

1. Create application tables in archive data target, if they do not already exist.
2. Create schemas in your ERP instance to provide access to history data.

Although these steps can be performed before using Informatica Data Archive, archive data target is not accessible until data has been archived by Data Archive.

Seamless Access for IBM DB2

Configure a source application on IBM DB2 for seamless data access. You can configure access to a combined view of the production data and the archive data, a view of the archived data, or both views. Run the Create Seamless Data Access Script job to configure seamless access.

You can configure seamless access on an IBM DB2 for AS/400 database, but Data Archive does not support seamless access to logical files on IBM DB2 for AS/400.

Create Seamless Data Access Script Job

To configure seamless access, you must create views and synonyms in the database. Most IBM DB2 database administrators do not allow external applications to create database objects. Use the Create

Seamless Data Access Script standalone job to create a script that the database administrator can review and then run on the database. The script includes statements to create the required views and synonyms for seamless access.

When you run the job, the job creates a script and stores the script in the location that you specified in the job parameters. The job uses the following naming convention to create the script file:

```
SEAMLESS_ACCESS_<Job ID>.SQL
```

The job uses the parameters that you specify to create the script statements. The script can include statements to create one of the following seamless access views:

- Combined view of the production and archived data.
- Query view of the archived data.

The script generates statements for the schema that you provide in the job parameters. You can enter the combined schema or the query schema. For example, if you provide the combined schema, then the script includes statements to create the views and synonyms for the combined seamless access view.

If you want to create both the combined view and the query view, run the job once for each schema.

Create Seamless Data Access Script Job Parameters

Enter the job parameters to specify how the job creates the script and includes the script statements.

The Create Seamless Data Access Script job includes the following parameters:

Source Repository

Archive source connection name. Choose the source connection for the IBM DB2 database that stores the source or production data.

Destination Repository

Archive target connection name. Choose the target connection for the IBM DB2 database that stores the archived data.

Combined Schema Name

Schema in which the script creates the seamless access view of both the production and the archived data.

Note that the job creates statements for one schema only. Configure either the Combined Schema Name parameter or the Query Schema Name parameter.

Combined Schema Password

Not required for IBM DB2.

Query Schema Name

Schema in which the script creates the view for the archived data.

Note that the job creates statements for one schema only. Configure either the Combined Schema Name parameter or the Query Schema Name parameter.

Query Schema Password

Not applicable for IBM DB2.

Combined/Query Schema Location

Location of the combined and query schemas.

Use one of the following values:

- Source
- Destination

The combined and query schemas may exist on the source database if a low percentage of the source application data is archived and a high percentage of data remains on the source.

The combined and query schema may exist on the target location if a high percentage of the source application data is archived and a low percentage of data remains on the source.

Generate Script

Determines if the job generates a script file. Always choose `Yes` to generate the script.

Database Link

Not required for IBM DB2.

Default is `NONE`. Do not remove the default value. If you remove the default value, then the job might fail.

Script Location

Location in which the job saves the script. Enter any location on the machine that hosts the ILM application server.

Configuring Seamless Access for IBM DB2

Before you configure seamless access for IBM DB2, verify that the combined and archive database schemas are created on the IBM DB2 database.

1. Run a script in the combined database schema to create the XA_OBJ_TEMP_TABLE table.

   The table is a required temporary table for the Create Seamless Data Access Script standalone job. The table stores temporary results of the job, such as which tables the job processes.

   Run the following script to create the temporary table:

   ```
   CREATE TABLE  XA_OBJ_TEMP_TABLE (OBJECT_NAME VARCHAR(2000),OBJECT_TYPE
   VARCHAR(2000), OBJECT_OWNER VARCHAR(2000),DEPENDENCE_LEVEL INTEGER)
   ```

2. Schedule and run the Create Seamless Data Access Script standalone job.

   The job creates a script file that you can provide to the database administrator for review. If you want to create both the combined view and the query view, run the job once for each schema.

3. The database administrator runs the script in the location specified in the script.

   You specify the location as part of the job parameters.

Seamless Access for Oracle

The Oracle seamless access configuration process includes the configuration of users, data groups, and responsibilities in Oracle E-Business Suite. The process enables seamless access from the production database and the archive data destination.

Configuring a Combined User

Configure a combined user to enable seamless access for Oracle.

1. Log in to Oracle E-Business Suite with administrator privileges.
2. Click System Administrator in the Navigator window.
3. Scroll down and select Security: ORACLE.
4. Click Register.

   The Navigator- System Administrator Main window appears and initializes the Oracle Users Sub Window to create a user account.
5. Register a database user for the combined schema.

   The Combined Schema user created previously in the data target or data source must be registered in the Oracle E-Business Suite to enable seamless data access through the Oracle interface.

   A user is registered in Oracle E-business Suite from the menu Security > Oracle > Register from the Navigator- System Administrator window.
6. Create a Data Group.

   A Data Group must be created for including applications that can be later associated with the Combined User.

   A Data Group can be created by navigating to Security > Oracle > Data Group from the Navigator- System Administrator window.
7. Define a responsibility for the Combined Schema user.

Configuring a Query User

A query user can access data from an archive target. The procedure to configure a query user is the same as configuring a combined user, except that you register a database user for the query schema.

Configuring the Query and Combined Users in Oracle E-Business Suite 12.2.5

In Oracle E-Business Suite version 12.2.5, the ability to configure a new user through the **Add** button on the **Oracle Users** screen was disabled for custom applications and schemas. To create the query user and the combined user for seamless access, complete the following steps:

1. Connect to the Linux machine through SQLPlus and run the following package:

```
SQL> sqlplus apps/apps
Connected.
SQL> exec FND_ORACLE_USER_PKG.LOAD_ROW('TEST', 'CUSTOM', 'INVALID', NULL, 'N', 'B');
```

2. To register the user, run the following commands:

```
SQL> conn system/manager
Connected.
SQL>  alter session set current_schema=APPS;
SQL> exec AD_ZD_PREP.ENABLE_CUSTOM_USER('TEST');
```

3. In the **Oracle Users** screen, query the name of the database user you created, in this example "test."

   The new user should be present in the **Oracle Users** screen. You can now edit the user, change the user privilege to enabled, and save the user. You can also change the user password.

Seamless Access for PeopleSoft

The seamless access for Peoplesoft configuration process includes the prerequisites and procedures to enable PeopleSoft seamless access.

Prerequisites

Verify that the following tasks are carried out before creating the PeopleSoft seamless access:

- Create a user for the combined access schema.
- Create a user for the Data Archive access schema.

Running the Data Archive Seamless Access Job

The Data Archive Seamless Data Access job creates the two schemas and populates them with the appropriate synonyms and views.

1. Click **Jobs** > **Schedule a Job**.
2. Select **Standalone Programs** and click **Add Item**.
3. Select **SEAMLESS_DATA_ACCESS** from the list.

   A list of job parameters appear. The following table describes each job parameter:

Parameter	Description
Source	Database with production data.
Target	Database with archived data.
Combined Schema Name	Schema name for seamless access across the two databases.
Combined Schema Password	Schema password for seamless access across the two databases.
Archive Schema Name	Schema name for allowing access to archived data.
Archive Schema Password	Schema password for allowing access to archived data.
Combined / Archive Schema Location	Specify whether schemas were created at the data source or data target.
Database Link Name	Database link name from data source to data target.

4. Select options from the list of values or enter the value for the appropriate parameters.
5. Schedule the Create Seamless Data Access job.

PeopleSoft Seamless Access Script

You must run a PeopleSoft seamless access script against your PeopleSoft instance. The scripts are available on the Informatica Product Download Center website. The script should be run as the SYSTEM user. You need to run the script using COMBINED as the "Combined Access ID" and once using ARCHIVE.

This script configures PeopleSoft-specific security objects needed for the two Seamless Access schemas, the most important of which are the PSDBOWNER and PSSTATUS tables. The data in the PSDBOWNER and PSSTATUS tables must be correct to enable seamless access. After running the script twice, these tables should resemble the following:

```
PS.PSDBOWNER
DBNAME   OWNERID
CRM890   SYSADM
COMBINED COMBINED
ARCHIVE  ARCHIVE

CRM890.PSSTATUS
VERSION     OWNERID  TOOLSREL  LASTREFRESHDTTM  LASTCHANGEDDTTM  OWNERACCT
UNICODE_ENABLED  DBID
157102 SYSADM 8.48 12-JUN-08 12-JUN-08 1 SERVER

COMBINED.PSSTATUS
VERSION     OWNERID  TOOLSREL  LASTREFRESHDTTM  LASTCHANGEDDTTM  OWNERACCT
UNICODE_ENABLED  DBID
157102 COMBINED 8.48 12-JUN-08 12-JUN-08 1 SERVER

ARCHIVE.PSSTATUS
VERSION OWNERID  TOOLSREL  LASTREFRESHDTTM  LASTCHANGEDDTTM  OWNERACCT  UNICODE_ENABLED
DBID
157102 ARCHIVE  8.48 12-JUN-08 12-JUN-08 1 SERVER
```

As the archived data is by project read-only, the seamless access script makes all PeopleSoft pages in the COMBINED and ARCHIVE schemas "display only" by default.

"Display-only" is a PeopleSoft notion that disables the action buttons and prevents the user from modifying data on a given page.

The seamless access script also unsets the "display only" flag on pages that allow you to click the button to run certain reports. This is a generic script and modifications may need to be made to adapt it to your particular environment.

Seamless Access for PeopleSoft on IBM DB2

To configure seamless data access for PeopleSoft on an IBM DB2 database, you must complete the following tasks:

1. Create combined and archive schemas in the database.
2. Create operating system users on the source database server.
3. Grant permissions to the combined and archive users.
4. Create application tables for each schema in the history database.
5. Create indexes for each schema in the history database.
6. Populate seamless data access schemas with views and synonyms.
7. Add the combine and archive schemas in the ODBC drivers on the PeopleSoft database server.
8. Customize and run the PeopleSoft Seamless Access Application Script.
9. Create and start an application server domain for the combined and archive schemas.
10. Create and start a web server for the combined and archive schemas.
11. Validate the data from the combined and archive schemas. Ensure the data is retrievable from the production and history databases.

Step 1. Create Combined and Archive Schemas

Create the combined and archive schemas on the IBM DB2 database.

Step 2. Create Operating System User Accounts on the Source Database Server

Create operating system users for the combined and archive schemas in the Windows or Linux operating system. Each account will permit the application user to access the appropriate view. Typical user names are ILMCOMB for combined access and ILMARCHIVE for archive-only access.

Complete the following steps to create users on a Windows environment:

1. Right-click **My Computer**.

   Click **Manage** > **Configuration**.
2. Open **Local Users and Groups**.
3. Open **Users**.
4. Add a user for the combined schema.
5. Add a user for the archive schema.

Step 3. Grant Permissions to Combined and Archive Users

Modify the sample script to grant the combined and archive users access to a seamless view of production and archive data.

Sample Script

The following script is a sample script that you can use to grant seamless access privileges to the combined user and archive users. Replace `USERNAME` with the correct combined or archive user name and run the script.

```
GRANT ALTER SESSION TO &&USERNAME;
GRANT CREATE DATABASE LINK TO &&USERNAME;
GRANT CREATE PROCEDURE TO &&USERNAME;
GRANT CREATE SEQUENCE TO &&USERNAME;
GRANT CREATE SESSION TO &&USERNAME;
GRANT CREATE SYNONYM TO &&USERNAME;
GRANT CREATE TABLE TO &&USERNAME;
GRANT CREATE TRIGGER TO &&USERNAME;
GRANT CREATE VIEW TO &&USERNAME;
GRANT EXECUTE ANY PROCEDURE TO &&USERNAME;
GRANT SELECT ANY DICTIONARY TO &&USERNAME;
GRANT SELECT ANY SEQUENCE TO &&USERNAME;
GRANT SELECT ANY TABLE TO &&USERNAME;
GRANT SELECT_CATALOG_ROLE TO &&USERNAME;
```

Step 4. Create Application Tables

Use the Create Tables job to create the managed tables in the archive schema on the destination repository.

1. Log on to the Informatica Data Archive user interface.
2. Click **Jobs** > **Schedule a Job**.
3. Click **Schedule Jobs**.
4. Select the **Create Tables** job from the list of values, and click **Select**.
5. Select the source and destination repositories.
6. Click **Schedule** to run the Create Tables job.

For every table in an entity defined in the application version of your source repository, the Create Tables job creates a table in the history schema of the destination repository.

7. You can navigate to **Jobs** > **View Current Jobs** to monitor the progress of this job. Wait until this job finishes before proceeding to the next step.

Step 5. Create Indexes

Use the Create Indexes job to create the indexes for the seamless access schemas on the destination repository.

1. Click **Jobs** > **Schedule a Job**.
2. Click **Schedule Jobs**.
3. Select the **Create Indexes** job from the list of values and click **Select**.
4. Choose the source and destination repository from the drop-down list.
5. Click **Schedule** to run the Create Indexes job.

   For every table in an entity defined in the application version of your source repository, the Create Indexes job creates indexes in the history schema of the destination repository.
6. You can navigate to **Jobs** > **View Current Jobs** to monitor the progress of this job. Wait until this job finishes before proceeding to the next step.

Step 6. Populate Seamless Data Access Schemas with Views and Synonyms

The Create Seamless Data Access Script standalone job generates a script. The script contains the statements necessary to create the required views and synonyms for seamless access. After you review the script and ensure its accuracy, run the script to create the views and synonyms, and populate the schemas with the views and synonyms.

Run the Create Seamless Data Access Script standalone job once for the archive schema and once for the combined schema.

1. Click **Jobs** > **Schedule a Job**.
2. Click **Schedule Jobs**.
3. Select the **Create Seamless Data Access Script** job from the list of values, and click **Select**.
4. Enter a value for the job parameters.
5. Click **Schedule** to run the job.
6. You can navigate to **Jobs** > **View Current Jobs** to monitor the progress of this job. Wait until this job finishes before proceeding to the next step.
7. Run the script generated by the **Create Seamless Data Access Script** job. The script runs in the location you specify in **Create Seamless Data Access Script** job parameters.

Step 7. Add the Schemas to the ODBC Drivers

Add the combined schema and archive schema to the ODBC drivers on the PeopleSoft database server.

Complete the following steps on a Windows environment:

1. Open the Windows Control Panel.
2. Click **Administrative Tools** > **Data Sources (ODBC)**.

3. Open **System DSN**.
4. Click **Add**.
5. Enter the schema name and details, and click **OK**.

Step 8. Run the PeopleSoft Seamless Access Application Script

The PeopleSoft Seamless Access Application script creates PeopleSoft application related tables and privileges. Customize the sample script in this section. Then run the script once for the archive schema and once for the combined schema.

Sample Script

The following script is a sample of the PeopleSoft Seamless Access Application script.

```
insert into PS.PSDBOWNER values(upper('COMB'),upper('COMB'))
/
grant select on PS.PSDBOWNER to COMB;
/
grant select on PSADM.PSOPRDEFN      to COMB;
grant select on PSADM.PSACCESSPRFL   to COMB;
grant select on PSADM.PSSTATUS       to COMB WITH GRANT OPTION;
grant select on PSADM.PSAUTHITEM     to COMB;
grant select on PSADM.PSPNLGROUP     to COMB;
grant select on PSADM.PSMENUITEM     to COMB;
grant select on PSADM.PSPNLGROUP     to COMB;
grant select on PSADM.PS_PRCSDEFNPNL to COMB;
grant select on PSADM.PSPNLGRPDEFN   to COMB WITH GRANT OPTION;
grant update on PSADM.PSLOCK         to COMB;
grant update on PSADM.psversion      to COMB;
grant update on PSADM.psserverstat   to COMB;
grant insert on PSADM.ps_serveractvty to COMB;
grant insert,update on PSADM.PSAcceSSLOG to COMB;
/
grant insert,update,delete on PSADM.PS_AERUNCONTROL to COMB;
grant insert,update,delete on PSADM.PS_AETEMPTBLMGR to COMB;
grant update,insert,delete on PSADM.PSTREESELECT10  to COMB;
/
grant update,insert,delete on PSADM.PSBATCHAUTH to COMB;
grant insert,update,delete on PSADM.PS_AERUNCONTROLPC to COMB;
grant insert,update,delete on PSADM.PS_QUERY_RUN_PARM to COMB;
grant insert,update,delete on PSADM.PS_PRCSRUNCNTLDIST to COMB;
/
grant select on PSADM.PS_SCRTY_QUERY     to COMB;
/
grant update,insert,delete on PSADM.PSIBSUBSLAVE    to COMB;
grant update,insert,delete on PSADM.PSIBBRKSLAVE    to COMB;
grant update,insert,delete on PSADM.PSAPMSGDOMSTAT  to COMB;
grant update,insert,delete on PSADM.PSWEBPROFHIST   to COMB;
grant update,insert,delete on PSADM.PSAPMSGQUEUESET to COMB;
grant update,insert,delete on PSADM.psrenclus_owner to COMB;
grant update,insert,delete on PSADM.psmcfrenurlid   to COMB;
grant update,insert,delete on PSADM.psrencluster    to COMB;
grant update,insert,delete on PSADM.PSAPMSGDSPSTAT  to COMB;
grant update,insert,delete on PSADM.PSIBPUBSLAVE    to COMB;
grant update,insert,delete on PSADM.PSAPMSGSUBCON   to COMB;
grant update,insert,delete on PSADM.PSIBFOLOCK      to COMB;
grant update,insert,delete on PSADM.PSAPMSGPUBHDR   to COMB;
grant delete on PSADM.PSANALYTICREG                 to COMB;
/
drop synonym COMB.PSSTATUS;
create or replace view COMB.PSSTATUS
as
select
    VERSION,
    'COMB' OWNERID,
    TOOLSREL,
    LASTREFRESHDTTM,
```

```
    LASTCHANGEDTTM,
    OWNERACCT,
    UNICODE_ENABLED,
    DBID,
    DATABASE_OPTIONS   --- Newly added field in 9.1
from PSADM.PSSTATUS;
/
SELECT * FROM PSADM.PSACCESSPRFL
/

drop synonym COMB.PSACCESSPRFL;
CREATE TABLE COMB.PSACCESSPRFL LIKE PSADM.PSACCESSPRFL
create unique index COMB.PS_PSACCESSPRFL on COMB.PSACCESSPRFL (symbolicid);
INSERT INTO COMB.PSACCESSPRFL (SYMBOLICID,VERSION,ACCESSID,ACCESSPSWD,ENCRYPTED) values
('psadm1',0,'COMB=','COMB=',0)
--update COMB.PSACCESSPRFL set ACCESSID = 'COMB', ACCESSPSWD = 'COMB', VERSION = 0,
ENCRYPTED = 0;
/

grant select on COMB.PSSTATUS to people;
grant select on COMB.PSACCESSPRFL to people;
/
drop synonym COMB.PSAUTHITEM;
/

create or replace view COMB.PSAUTHITEM
as
select
    a1.CLASSID,
    a1.MENUNAME,
    a1.BARNAME,
    a1.BARITEMNAME,
    a1.PNLITEMNAME,
    1 DISPLAYONLY,
    a1.AUTHORIZEDACTIONS
   from PSADM.PSAUTHITEM a1
   WHERE a1.BARITEMNAME NOT IN ('PROCESSMONITOR','QUERY_MANAGER','QUERYVIEWER',
   'PO_INQUIRY',  -- This Allows the user to click on buttons in PO Inquiry
    'DRILLDOWNREGISTER','IC_REPORTBOOK','IC_RUN_DRILL','NVS_REPORT_REQUEST','NVS_SCOPE',
-- This allows NVISION access
    'PYMNT_INQ_SRCH',  -- This alllows Payment Inquiry access
    'CONTENT_LIST',  -- This allows REPORT MANAGER access
    'SCHEDQUERY2',  -- This allows Scheduling a Query
    'SCHEDQUERY',  -- This allows Scheduling a Query
    'VOUCHER_ACCOUNTING_ENTRIES',  -- This allows Voucher Accounting Entries Inquiry page
    'PYCYCL_DATA_INQ',  -- This allows Voucher Payment Inquiry Detail page
    'PYCYCL_DATA_SUM',  -- This allows Voucher Payment Inquiry Summary page
    'JOURNAL_STATUS',  -- This allows Jounrnal Status Inquiry Page - bpisack 06/16/2008
    'INQ_LED_CMP_PNL',  -- This allows the Ledger Inqury Page - bpisack 06/16/2008
    'PV_MANAGE',  -- This allows the Manager Requisition Page - bpisack 06/16/2008
    'AP_VCHR_INQ',  -- This allows AP Voucher Inquiry page - bpisack 06/16/2008
    'GL_CASH_VOIDS',
    'HCR_HCRRP026_PGRP',
    'HCR_HIST_RPT',
    'HCR_JRNL_ENTRIES',
    'HCR_LEGAL_SERV',
    'HCR_MONTHLYCASH',
    'HCR_NETPAYROLL',
    'HCR_RC_REIMB',
    'HCR_RUN_PAY002H',
    'HCR_RUN_TAX910AU',
    'HCR_RUNCTL_REIM',
    'HCR_TAX001H_REPORT',
    'HCR_VAC_SICK_HOL',
    'HCR_WAGE_PASS',
    'HCRRB026_PG',
    'HCRRP001',
    'HCRRP002',
    'HCRRP013',
    'HCRRP015_PNLGRP',
    'SMSPCONTRIBUTIONS',        -- wholowac 30/07/2008
```

```
    'HCR_401KCONTRIB_PG',        -- wholowac 30/07/2008
    'HCRRPTX2')        -- wholowac 30/07/2008
   and a1.BARNAME NOT IN ('REPORT','RPT_N-Z','RPT_A-M')
union
select
    a2.CLASSID,
    a2.MENUNAME,
    a2.BARNAME,
    a2.BARITEMNAME,
    a2.PNLITEMNAME,
    0 DISPLAYONLY,
    a2.AUTHORIZEDACTIONS
   from PSADM.PSAUTHITEM a2
   WHERE a2.BARITEMNAME IN ('PROCESSMONITOR','QUERY_MANAGER','QUERYVIEWER',
      'PO_INQUIRY',  -- This Allows the user to click on buttons in PO Inquiry
    'DRILLDOWNREGISTER','IC_REPORTBOOK','IC_RUN_DRILL','NVS_REPORT_REQUEST','NVS_SCOPE',
-- This allows NVISION access
    'PYMNT_INQ_SRCH',  -- This alllows Payment Inquiry access
    'CONTENT_LIST',  -- This allows REPORT MANAGER access
    'SCHEDQUERY2',  -- This allows Scheduling a Query
    'SCHEDQUERY',  -- This allows Scheduling a Query
    'VOUCHER_ACCOUNTING_ENTRIES',  -- This allows Voucher Accounting Entries Inquiry page
    'PYCYCL_DATA_INQ',  -- This allows Voucher Payment Inquiry Detail page
    'PYCYCL_DATA_SUM',  -- This allows Voucher Payment Inquiry Summary page
    'JOURNAL_STATUS',  -- This allows Jounrnal Status Inquiry Page - bpisack 06/16/2008
    'INQ_LED_CMP_PNL',  -- This allows the Ledger Inqury Page - bpisack 06/16/2008
    'PV_MANAGE',  -- This allows the Manager Requisition Page - bpisack 06/16/2008
    'AP_VCHR_INQ',  -- This allows AP Voucher Inquiry page - bpisack 06/16/2008
    'GL_CASH_VOIDS',
    'HCR_HCRRP026_PGRP',
    'HCR_HIST_RPT',
    'HCR_JRNL_ENTRIES',
    'HCR_LEGAL_SERV',
    'HCR_MONTHLYCASH',
    'HCR_NETPAYROLL',
    'HCR_RC_REIMB',
    'HCR_RUN_PAY002H',
    'HCR_RUN_TAX910AU',
    'HCR_RUNCTL_REIM',
    'HCR_TAX001H_REPORT',
    'HCR_VAC_SICK_HOL',
    'HCR_WAGE_PASS',
    'HCRRB026_PG',
    'HCRRP001',
    'HCRRP002',
    'HCRRP013',
    'HCRRP015_PNLGRP',
    'SMSPCONTRIBUTIONS',        -- wholowac 30/07/2008
    'HCR_401KCONTRIB_PG',        -- wholowac 30/07/2008
    'HCRRPTX2')        -- wholowac 30/07/2008
   and a2.BARNAME NOT IN ('REPORT','RPT_N-Z','RPT_A-M')
union
select
    a4.CLASSID,
    a4.MENUNAME,
    a4.BARNAME,
    a4.BARITEMNAME,
    a4.PNLITEMNAME,
    0 DISPLAYONLY,
    a4.AUTHORIZEDACTIONS
   from PSADM.PSAUTHITEM a4
   WHERE a4.BARNAME IN ('REPORT','RPT_N-Z','RPT_A-M')
union
select
    a5.CLASSID,
    a5.MENUNAME,
    a5.BARNAME,
    a5.BARITEMNAME,
    a5.PNLITEMNAME,
    0 DISPLAYONLY,
    a5.AUTHORIZEDACTIONS
```

```
   from PSADM.PSAUTHITEM a5
   WHERE a5.MENUNAME IN ('QUERY_MANAGER')
union
select
    a6.CLASSID,
    a6.MENUNAME,
    a6.BARNAME,
    a6.BARITEMNAME,
    a6.PNLITEMNAME,
    0 DISPLAYONLY,
    a6.AUTHORIZEDACTIONS
   from PSADM.PSAUTHITEM a6
   WHERE a6.BARNAME = 'INQUIRE'
   and a6.MENUNAME = 'CREATE_PAYMENTS';
 /
 drop synonym COMB.PS_SCRTY_QUERY;
create or replace view COMB.PS_SCRTY_QUERY
as
select
   CLASSID,
   VERSION,
--  'N' QRY_RUN_ONLY,      Use this line to allow edit of queries and comment out field
below
   'Y' QRY_RUN_ONLY,
   QRY_CREATE_PUBLIC,
   QRY_CREATE_WFLOW,
   QRY_MAX_FETCH,
   QRY_MAX_RUN,
   QRY_ADV_DISTINCT,
   QRY_ADV_ANY_JOIN,
   QRY_ADV_SUBQUERY,
   QRY_ADV_UNION,
   QRY_ADV_EXPR,
   QRY_MAX_JOINS,
   QRY_MAX_IN_TREE,
 --  'N' QRY_OUT_LISTBOX,  Use this line to turn off Query to HTML and Comment out field
below
   QRY_OUT_LISTBOX,
 --  'N' QRY_OUT_NVISION,  Use this line to turn off Query to Excel and Comment out
field below
   QRY_OUT_NVISION,
   QRY_OUT_CRYSTAL,
   QRY_ADM_AUTOPUBLIC,
   QRY_ADM_AUTOPRIV,
   QRY_ADM_LIMUNAPPRV,
   QRY_ADM_UNAPP_ROWS
   from PSADM.PS_SCRTY_QUERY;

 /
 DROP SYNONYM COMB.PSPNLGRPDEFN;
/
CREATE OR REPLACE VIEW COMB.PSPNLGRPDEFN
  (PNLGRPNAME, MARKET, VERSION, ACTIONS, DESCR,
   ADDSRCHRECNAME, SEARCHRECNAME, SEARCHPNLNAME, LOADLOC, SAVELOC,
   DISABLESAVE, OBJECTOWNERID, LASTUPDDTTM, LASTUPDOPRID, PRIMARYACTION,
   DFLTACTION, DFLTSRCHTYPE, DEFERPROC, EXPENTRYPROC, REQSECURESSL,
   INCLNAVIGATION, FORCESEARCH, ALLOWACTMODESEL, PNLNAVFLAGS, TBARBTNS,
   SHOWTBAR, ADDLINKMSGSET,  ADDLINKMSGNUM, SRCHLINKMSGSET,SRCHLINKMSGNUM,
   SRCHTEXTMSGSET, SRCHTEXTMSGNUM, WSRPCOMPLIANT, DESCRLONG)
AS
SELECT PNLGRPNAME,MARKET,VERSION,ACTIONS - MOD(ACTIONS,
2) ,DESCR,ADDSRCHRECNAME,SEARCHRECNAME,SEARCHPNLNAME,LOADLOC,SAVELOC
,DISABLESAVE,OBJECTOWNERID,LASTUPDDTTM,LASTUPDOPRID,PRIMARYACTION,DFLTACTION,DFLTSRCHTYPE
,DEFERPROC,EXPENTRYPROC,REQSECURESSL
,INCLNAVIGATION,FORCESEARCH,ALLOWACTMODESEL,PNLNAVFLAGS,TBARBTNS,SHOWTBAR,ADDLINKMSGSET,A
DDLINKMSGNUM,SRCHLINKMSGSET,SRCHLINKMSGNUM
,SRCHTEXTMSGSET,SRCHTEXTMSGNUM,WSRPCOMPLIANT,DESCRLONG FROM PSADM.PSPNLGRPDEFN;
```

Step 9. Create and Start an Application Server Domain

Use the PeopleSoft Administrator's Utility to create and start an application server domain for the combined schema and archive schema.

Complete the following steps on a Windows environment:

1. Open the PeopleSoft Application Designer application.

   The **Signon** window appears.
2. Log in to the application. Use the seamless data access schema name in the **Database Name** field.
3. Navigate to the PeopleSoft home directory, for example `C:\PT8.52`.
4. Open the `appserv` folder and double-click `psadmin.exe`.

   The PeopleSoft Administrator's Utility opens.
5. Enter `1` at the prompt to indicate you want to configure the application server.
6. Enter `2` at the prompt to indicate you want to create a domain on the application server.
7. Enter a name for the domain. Use a name that indicates that the domain is for the archive schema or combined schema. For example, enter a name such as `DEV_COMBS` if you are creating a domain for the combined schema.
8. Enter a number to indicate the configuration template. For example, enter `2` for a large template.
9. Enter `Y` to indicate you want to configure the domain.
10. Enter `8` to disable the Event Notification feature.
11. Enter `18` to update the user ID.
12. Enter the new user ID.
13. Enter `19` to update the user password.
14. Enter the new password.
15. Enter `23` to update the password for the Connect user ID.
16. Enter the password for the Connect user ID.
17. Enter `13` to load the new configuration.

    The following message appears: Domain configuration complete.
18. Enter `1` to start the domain.
19. Enter `1` to start a serial boot.

    The list of servers appear as they start up.

Step 10. Create and Start a Web Server

Create and start a web server for the combined schema and archive schema.

1. Go to the PeopleSoft home directory, for example `C:\PT8.52`.
2. Go to `setup/PsMpPIAInstall` folder and double-click `setup`.

   The PeopleSoft Pure Internet Architecture installer application appears.
3. Select the PeopleSoft home directory to install the PeopleSoft Pure Internet Architecture application.
4. Select **Oracle WebLogic Server**.
5. Select the Oracle folder.
6. Enter the administrator login and password to the WebLogic domain.

7. Select **Create New WebLogic Domain**.
8. Enter a domain name such as `DEV_COMBS` if you are creating the web server for the combined schema.
9. Select **Single Server Domain**.
10. Enter a website name such as `dev_combs` if you are creating the web server for the combined schema.
11. Enter the application server name and port numbers.
12. Select the default options for the remaining steps.

    Wait for the installation to complete before going to the next step.
13. Navigate to the following folder: `/<PeopleSoft home directory>/webserv/<domain name>/bin`
14. Double-click `StartPIA`.

    The WebLogic Server starts.
15. Open a browser and enter the URL in this format: `http://<host name>/psp/<website name>/?cmd=login`

Step 11. Validate Data from the Combined and Archive Schemas

Verify you have access to the combined and archive views.

1. Log in to the PeopleSoft Enterprise application.
2. Click **Main Menu** and select a module.
3. Verify that the data displays.

CHAPTER 12

Data Discovery Portal

This chapter includes the following topics:

Data Discovery Portal Overview

Use the Data Discovery portal to search for transactions that are archived to the Data Vault. You can search across applications or within an entity. You can also apply retention policies, legal holds, or tags to archive data through the Data Discovery portal.

You can access a transaction in Data Vault from an external application. You create a customized URL that contains information about the entity and table the transaction belongs to. You use the URL to access data from the Data Vault.

Enable the following options in Data Discovery:

Search Data Vault

Use Search Data Vault to search for records across applications in Data Vault. You must first specify columns that you want to use in Search Data Vault. Then, create a search index of these columns. Only values from indexed columns can trigger a search. For example, if the user wants to search for records containing the word Emily, the search results display records if the word Emily is in an indexed column.

Search Within an Entity

Use Search Within an Entity to access records within an entity in Data Vault. You can define search options for each entity in an application version. When you define search options, you specify the columns to include as search parameters and the columns to display and sort on in the search results. If you do not define search options, you can use any column as a search parameter and all columns appear in the results. Users might not need all columns to search with or view in the results. For example, if you do not want to include unused columns in the search or search results, select columns with data to search with and display. Or, if the original application permitted you to search for data by using the column Vendor Name but not Vendor ID, then to avoid confusion, present the same search elements for archived data. Choose the column Vendor Name as a search option.

Mask Sensitive Data

You can mask or block sensitive information in Data Vault from appearing in search results or in data visualization reports. When you configure Dynamic Data Masking, queries sent to Data Vault go through Dynamic Data Masking. Dynamic Data Masking applies masking rules based on the Data Vault access role assigned to the user. Fields that a user does not have privileges to appear masked in the search results.

You must install and configure Informatica Dynamic Data Masking. For more information on masking data in Data Vault, see the *Informatica Dynamic Data Masking Data Archive Accelerator Guide*.

Integrate with E-Discovery Solutions

You can bridge the gap between a legal hold notification and the point of data preservation by integrating Data Archive with Exterro Fusion, an e-discovery solution. Users can search and preserve data in Data Vault directly from the Exterro Fusion user interface.

Search Data Vault

Use keywords to search for records across entities and applications in Data Vault.

To enable Search Data Vault, you must create a search index for each table you want to include in the search. Each search index contains a list of columns that you specify in Enterprise Data Manager. After you specify columns for the search index, you create the search index by running the Create Indexes in Data Vault job.

When you search Data Vault with a keyword or term, the search engine looks for the keyword or term in indexed columns. If the keyword or term is in an indexed column, the corresponding record appears in the search results. If the keyword or term is not in an indexed column, the record does not appear in the results even if the record contains the keyword in another column.

You must maintain the search index to keep it in sync with Data Vault.

Setting up Search Data Vault

To enable Search Data Vault, perform the following tasks:

1. Specify columns to include in the search index.
2. Configure the `conf.properties` file.
3. Assign roles.
4. Run the Create Indexes in Data Vault job.

Step 1. Specify Columns for the Search Index

In Enterprise Data Manager, specify columns that you want to include in the search index. Optionally, specify the columns for the record header in the search results.

1. Click **Accelerators** > **Enterprise Data Manager**.

   The Enterprise Data Manager interface appears.
2. Select an application version and a table.
3. Click **View** > **Constraints**.
4. Click the **Columns** tab.

A list of all the columns in the table appears.

5. Enable the **Index for Search** field for each column you want to add to the search index.
6. Click the **Constraints** tab.
7. Click the add icon to add a constraint.

   A row appears with the name `NEW CONSTRAINT`.
8. Double-click the **Name** field and enter a name.
9. Optionally, to specify a column name as the record heading on the results page, click the **Type** field and select **Search Header**.

   If you do not specify a column for the record heading, Data Archive uses the column name of the primary key. In the absence of a primary key, Data Archive uses the row ID for the record heading.
10. Enable the **Enabled** field.
11. Click the add icon in between the Child Table Columns and Parent Table Columns panels of the Constraints tab.

    A row appears in the Child Table Columns section.
12. Click the **Type** field and select **Column**.
13. Click the **Name** field and select a column.
14. To add more columns to the record header, repeat steps 11 to 13
15. Click **File** > **Save**.

Specifying Large Number of Columns at a Time

To add many columns to the search indexes, update the index metadata file in Enterprise Data Manager.

1. Click **Accelerators** > **Enterprise Data Manager**.

   The Enterprise Data Manager interface appears.
2. Click **View** > **Constraints**.
3. Right-click an application version and select **Export Index Metadata**.

   A csv file appears with a list of columns for all the tables in the application version.
4. Enter **Y** in the **Quick_Search_Index** column for the columns you want to include in the search index.
5. Save the csv file.
6. Right-click the application version and select **Import Index Metadata**.

Step 2. Configure Conf.Properties

Update the `conf.properties` file with the file path to the search indexes. Optionally, specify the number of search indexes that Data Archive can create in parallel. You can find the `conf.properties` file in the root Data Archive installation directory.

Configure the following parameters in the `conf.properties` file:

informia.keywordSearchIndexDir

Required. Enter the file path of the folder that contains the search indexes.

The folder must be in the same root directory as the Data Archive installation. The ILM Engine must have read and write privileges to the folder.

The Create Indexes on Data Vault and Delete Indexes on Data Vault jobs use the file path entered in this property to find the search indexes.

Based on your operating system, use one of the following formats for the file path:

- On a Windows operating system, use a double back slash between folder names. For example, `C:\\KeywordSearchIndex`
- On all other operating systems, use a single forward slash between folder names. For example, `/Data/KeywordSearchIndex`

informia.maxActiveIndexThreads

Optional. Enter the number of search indexes that Data Archive can create in parallel.

Minimum value is 1. Default is one less than the number of cores on the machine hosting the ILM Engine.

Step 3. Assign Roles

Assign the system-defined role and access roles to the user authorized to search Data Vault.

1. Assign the Discovery User role to the user.
2. Assign Data Vault access roles to the user.

   Users can access entities that have the same Data Vault access role.

Step 4. Create the Search Index

To create the search indexes required for Search Data Vault, run the Create Indexes in Data Vault job.

Create Indexes in Data Vault Job Parameters

The following list describes the job parameters for the Create Indexes in Data Vault job:

Destination Repository

Required. Archive folder that contains the archive data. Choose the archive folder from the list of values. The target connection name appears before the archive folder name.

Entity

Optional. The name of the entity for the table whose columns you want to add or remove from the search index.

Table

Optional. The name of the table whose columns you want to add or remove from the search index.

If you do not select an entity or table, Data Archive creates search indexes for all the tables in the archive folder.

Scheduling the Create Indexes in Data Vault Job

1. Click **Jobs** > **Schedule a Job**.

   The **Schedule Job** page appears.
2. Select **Standalone Programs** and click **Add Item**.

   The **Select Definitions** window appears with a list of available jobs.
3. Select the **Create Indexes in Data Vault** job and click **Select**.

   The job parameters appear on the **Schedule Job** page.

4. Enter values for the job parameters.
5. Schedule the job to run immediately.
6. Enter an email address to receive notification when the job completes, terminates, or returns an error.
7. Click **Schedule**.

Maintaining Search Indexes

To keep the search indexes in sync with Data Vault, you must update the list of index columns. Then run jobs to delete and create the search indexes for the updates to apply.

Update the search indexes after the following scenarios:

You archive an application to Data Vault.

To include the application in the search, perform the following steps:

1. Specify columns for the search index in Enterprise Data Manager.
2. Run the Delete Indexes in Data Vault job.
3. Run the Create Indexes in Data Vault job.

You purge data from Data Vault.

To exclude the purged data from the search, perform the following steps:

1. Run the Delete Indexes in Data Vault job.
2. Run the Create Indexes in Data Vault job.

You discover that users predominantly enter keywords from a column that is not indexed.

To add a column to the search index, perform the following steps:

1. Specify the column for the search index in Enterprise Data Manager.
2. Run the Delete Indexes in Data Vault job.
3. Run the Create Indexes in Data Vault job.

Search Within an Entity in Data Vault

You can customize the Search Within an Entity in Data Vault feature.

You can select the types of columns to include in a search query and specify the display format for the data. You can also specify the number of records you want Data Archive to display in the search results.

To customize Seach Within an Entity in Data Vault, perform the following steps:

1. Define search options for each entity in the application.
2. Specify the maximum number of records to display in the search results.

Step 1. Define Search Options

Define search options to search for records within an entity in Data Vault.

You define search options for each entity archived to Data Vault. When you define search options, the system displays a list of all columns from the entity driving table. You can use the columns from the driving table or add columns from another table such as a child or reference table.

Use search options to specify the columns that you can use to search Data Vault. You can also restrict columns from appearing in the search results.

You can specify the following options for each column:

Display

Displays the column data in the search results. Only columns marked as display are included when you export search results from the Data Vault search.

Data Format

The display format for numeric and datetime values in the search results. For example, you might set the data format for the ITEM_PRICE column to `$####.##`. If you do not enter a data format for datetime values, Data Archive uses the default discovery result date format from the system profile. If you do not enter a data format for numeric values, Data Archive displays the number as it is stored in the database.

This field is available for numeric and datetime datatypes.

Data Conversion For Display

The display format for dates in Data Vault search. If dates are in the Julian format, you can select the Julian to Gregorian Date option. This enables you to search and view data in the Gregorian format in Data Vault search.

Julian to Gregorian conversion is only supported for columns with Julian dates that contain five or six digits.

The following types of Julian dates are supported:

1. Julian dates with the century included. If the Julian date includes the century, enter "`C`" in the Data Format field to indicate that the date includes the century. In this case, "15001" would convert to 1-Jan-1915, and "115001" would convert to 1-Jan-2015.
2. Julian dates without the century included. If the Julian date does not contain the century, specify the limiting year in the `conf.properties` file with the following property: informia.julianLimiterYear=

   For example: informia.julianLimiterYear=1975

   The 5-digit Julian date "15001" could be either 1915 or 2015. If you set the limiting year to 1975, then the dates "05001," "15001," and "65001" are converted as 2005, 2015, and 2065. The Julian dates "75001," "85001," and "95001" are converted as 1975, 1985, and 1995.

Search

Uses the column as a search parameter. You will not be able to select exotic or customized datatype columns because these types of columns cannot be used to search the Data Vault.

Sort

Allows you to use the column to sort data in the search results. If you select this option, you must also select the **Display** option for the column.

Description

The column heading in the search results. For example, you might set the description for the ORDID column to "Order ID."

Note: If you do not select any columns for display, search, or sort, then all columns are available to display, search, and sort on.

Editing Search Options

Edit search options to specify columns that you want to display, search, and sort in Data Vault search. You can also choose to display column headings with custom text and column values with a different format on Data Vault search.

1. Click **Data Discovery** > **Search Options**.

   The **View Search Options** page appears.
2. Select an **Application Version**.
3. Select an **Application Module**.
4. Select an **Entity**.
5. Click **Edit**.

   The **Edit Search Options** page appears with a list of the columns in each table in the entity.
6. Click **Add Entity Table** if you want to include columns from a reference or child table.

   A list of child and reference tables appear.
7. Select a table and click **Search**.

   The columns from the table are displayed on the **Edit Search Options** page.
8. Specify search options for each table column.
9. Click **Save**.

Step 2. Specify Number of Records in Results

You can limit the number of records in the search results by specifying the maximum number of records you want Data Archive to display in the results.

When you search through Search Within an Entity in Data Vault, by default, Data Archive displays a maximum of 1000 records in the search results.

You can specify a different default value according to your business requirements. The default value you specify applies to the search results of all queries by all users.

For example, an entity that you archived to the Data Vault contains 50,000 records. You want to view records that contain the value Amsterdam in the City column. Two thousand of the records in the entity match the search criteria. The default number of records in the search results is set to 1000. As a result, Data Archive displays the first 1000 records that match the search criteria.

A user can specify a different number of records for Data Archive to display in the results. The user specifies the number on the Search Within an Entity in Data Vault page. The user-specified number applies to the current search.

To set the default maximum number of records in the search results, specify a value from one to 10,000. Do not use a comma, period, or space in the value.

Note: When you export results, Data Archive exports all records that meet the search criteria regardless of the value you set to specify the default maximum number of records in the results.

Editing the Default Maximum Records in Results Value

To specify the maximum number of records you want to view in a search result, edit the value for the Default Maximum Records in Results parameter.

1. Click **Administration** > **System Profile**.

   The **Configuration Settings** page appears.

2. Go to the **Data Discovery Portal** tab.
3. Enter a value in the **Default Maximum Number of Records in Results** parameter.

   The value you enter must meet the following requirements:

   - Enter a number from one to 10,000.
   - Do not use a comma, space, or period in the value.
4. Click **Save**.

   The **Default Maximum Number of Records in Results** value applies to search queries made through Search Within an Entity in Data Vault.

Masking Sensitive Information in Data Vault

You can prevent sensitive information in Data Vault from appearing in search results and in data visualization reports.

You can selectively mask or block sensitive information based on user privileges. For example, a user has the privilege to access an entity containing employee information through Data Discovery. The entity includes fields for salary and birth date. If the user does not have the privilege to view salaries and birth dates, then the salary and birth date fields appear as Xs instead of the real value.

To prevent a user or application from accessing sensitive data, you must install and configure Informatica Dynamic Data Masking. You define rules in Dynamic Data Masking to specify the data to mask and who to display masked data to.

You can specify how to display masked data. For example, you can choose to replace values partially or completely with randomly generated characters or predefined characters. Or, you can choose to block sensitive fields from appearing in search results or in reports.

When a user queries the Data Vault, the masking rules apply to the search results presented to the user.

After you enable masking, Data Archive protects sensitive information about the user querying Data Vault. Data Archive encrypts the user ID and user access roles in the Data Archive logs and in the query sent to Dynamic Data Masking.

Enabling Dynamic Data Masking

To implement Dynamic Data Masking for Data Vault, perform the following tasks:

1. Before you configure Dynamic Data Masking, create the Data Vault archive folder in the Data Archive target connection.
2. Install the Dynamic Data Masking Server and Management Console.

   For more information, see the *Informatica Dynamic Data Masking Installation and Upgrade Guide*.
3. In the Dynamic Data Masking Management Console, configure Dynamic Data Masking for the Data Vault.
   a. Add the Data Vault service in the Management Console.
   b. Define the listener port for the Data Vault service.
   c. Define the Data Vault connection in the Management Console.
   d. Create a security rule set and define security rules.
   e. Create connection rules.

For more information about Dynamic Data Masking services and connection management, see the *Informatica Dynamic Data Masking Administrator Guide*. For more information about Dynamic Data Masking connection rules and security rule sets, see the *Informatica Dynamic Data Masking User Guide*.

4. Enable Dynamic Data Masking in Data Archive.
5. Define the Data Vault host and port parameters in Data Archive.
6. Create Data Archive access roles.
7. Assign access roles to users.
8. Update the target connection with the Dynamic Data Masking host name and port.

For more information about how to set up and use the Dynamic Data Masking Accelerator for Data Archive, see the *Informatica Dynamic Data Masking Data Archive Accelerator Guide*.

Integration with E-Discovery Solutions

Electronic discovery (e-discovery) refers to the initial stage in civil litigation and regulatory inquiries. The e-discovery process includes tasks to identify, preserve, process, produce, and present potentially relevant data. E-discovery solutions help the legal team manage legal and compliance obligations required during the e-discovery phase.

When data relevant to e-discovery is in the Data Vault, a legal team might depend on a technical team to search and preserve data through the Data Archive interface. This dependency might lead to delays between the request and response or to incomplete responses. To reduce the risk of e-discovery fines due to late or incomplete responses, enable a legal team to directly search and preserve data in Data Vault through a single, central application, the e-discovery solution.

To integrate Data Archive with an e-discovery solution, you must have both applications installed and configured to communicate with each other.

In-Place Preservation for Data Vault Through Exterro Fusion

Data preservation is a crucial step in the e-discovery process. Preservation refers to securing potentially relevant data as quickly as possible to prevent it from being altered or deleted. In-place preservation is the preferred type of preservation where data is preserved at its current location. You apply a legal hold to preserve data. A legal hold is like a virtual lock on data. When you apply a legal hold, the data continues to be retained even if the retention policy indicates that the data has expired. When you remove a legal hold, data is subject to the scheduled purge and retention policies.

Exterro Fusion is an e-discovery solution that includes the workflow for in-place preservation. Previously, to find and preserve relevant data, a legal team would need to use two applications. The team would use Data Archive to search through structured data archived to the Data Vault. The team would use Exterro Fusion to search through unstructured data on other sources. To use Data Archive, the legal team might need the help of a technical team. The need to use different applications and the dependency on another team to search and preserve data could present a delay in preserving data. To bridge the gap between the point of a legal hold notification and the point of data preservation, integrate Data Archive with Exterro Fusion. With the integration, a non-technical user can search and manage data preservation across structured and unstructured data from a single application, Exterro Fusion.

Legal Hold Specifications

To preserve data, you apply a legal hold at an entity or an application level. Within an entity, you can specify a date range to select a set of records to preserve. The date range applies across all date fields in the driving

table of the entity. For example, if you have four date columns in the driving table of the entity, the legal hold applies to records with a date that falls within the date range, in at least one of the four columns.

When you apply a legal hold to an application, you apply it at the target connection or the archive folder level. An application-level legal hold applies to all entities and tables in the application. You can apply different legal holds at a time to an application or entity.

For more information, see the Exterro Fusion documentation.

Integrating Data Archive and Exterro Fusion

To integrate Data Archive and Exterro Fusion, configure both applications.

Configure Exterro Fusion to access Data Vault. For more information, see the Exterro Fusion documentation.

Configure Data Archive to communicate with Exterro Fusion. You must specify connection details to the Exterro Fusion machines that require access to the ILM Engine. You must also configure connection parameters for Data Vault.

1. Open the Data Archive `conf.properties` file.

   You can find the `conf.properties` file in the root Data Archive installation directory.
2. Find the ValidHosts property.
3. Enter the IP address of each Exterro Fusion machine that requires access to the ILM Engine.
4. Save and close the `conf.properties` file.
5. From the Data Archive interface, click **Administration** > **Manage Connections**.

   The **Manage Connections** page appears.
6. Click the **Target** tab.
7. Click the Data Vault target connection.

   The Data Vault connection properties appear.
8. Enter the user name of the data custodian in the **Application Owner** property.
9. Enter the email address of the data custodian in the **Application Owner Email ID** property.

   Exterro Fusion users can use this email to set up legal hold notifications.
10. Click **Save**.

Troubleshooting Legal Holds Applied from the Exterro Fusion Interface

If you encounter one of the following error messages during a legal hold application, perform the steps in the respective section to resolve the issue.

The <IP address> IP address is not authorized to complete the legal hold task. Contact the Data Archive administrator.

1. Stop the ILM Engine.
2. Open the `conf.properties` file.
3. Find the ValidHosts property.
4. Add the IP address of the machine on which the legal hold request originated.
5. Save the `conf.properties` file.
6. Restart the ILM Engine.

The <user ID> user does not have the privileges to access NCDS <NCDS name>. Contact the Data Archive administrator.

Note: Non-custodial data sources (NCDS) is an e-discovery term that refers to an entity or application.

1. Determine if the NCDS is an entity or application.
   Check the globally unique identifier (GUID) of the NCDS.
   - An entity-level NCDS has the following format:

     `E<entity_ID>_A<rep_ID>`

     where `E` refers to entity and `A` refers to application.
   - An application-level NCDS has the following format:

     `A<application_ID>_A<rep_ID>`

     where `A` refers to application.
2. From the Data Archive interface, go to **Administration** > **Manage Users**.
3. If the NCDS is an entity, verify if the user has the same Data Vault access role as the entity.
4. If the NCDS is an application, verify if the user has the Data Vault access role of each entity in the application.
5. Assign any missing Data Vault access roles that the user requires to complete the legal hold request.

The <legal hold name> legal hold does not exist. Provide the name of an existing legal hold.

1. From the Data Archive interface, go to **Data Discovery** > **Manage Legal Hold Groups**.
2. Verify the name of the legal hold.

The <user ID> user does not have the privileges to the NCDS associated with the <legal hold name> legal hold. Contact the Data Archive administrator.

1. Determine if the NCDS is an entity or application.
2. From the Data Archive interface, go to **Administration** > **Manage Users**.
3. If the NCDS is an entity, verify if the user has the same Data Vault access role as the entity.
4. If the NCDS is an application, verify if the user has the Data Vault access role of each entity in the application.
5. Assign any missing Data Vault access roles that the user requires to complete the legal hold request.

404 Not found.

The ILM Engine might be down. Restart the ILM Engine.

Accessing Archive Data from an External Application

You can access data in the Data Vault from an external application. You use an API in an external application to run a Data Discovery search and to return the search results in the application view. When you search from an external application, you do not need to use the Data Discovery portal.

To search Data Discovery from an external application, you must first add the IP address of the machine that hosts the external application to the `conf.properties` file. Then form a URL and enter the Data Discovery search parameters. Add the URL to the external application code. When the code calls the URL, the URL runs

the Data Discovery search and returns the search results in the application view. The entity that you search from must have an associated XSL stylesheet to view the data in the application view.

Perform the following steps to search Data Discovery from external applications:

1. Configure security. Specify the machines that have access to call the API.
2. Form the URL that calls the API. In the URL, specify the parameters for the Data Discovery search.
3. Add the URL to the external application code. When the external application calls the API, Data Archive uses the parameters in the URL to run the Data Discovery search.

Step 1. Configure Security

The security to run Data Discovery searches is enforced based on the machine that calls the API. The authentication is not based on users.

In the `conf.properties` file, specify the machines that have access to call the API. By default, no machines have access.

1. Access `conf.properties` from the Web container folder of the Data Archive installation.
2. Configure the validHosts property.

   Enter the IP address or the host name of the machines that can access the API from the external application. Use a comma to separate multiple values.

   The following text is an example of the property:

   ```
   validHosts=host1, 10.11.22.33, hyw172967.abc.com, dev.*
   ```
3. Remove the comment tag before and after the informia.appViewGetAllowed property in `conf.properties` to access the Data Discovery portal application view from an external application.
4. Save the file.

Step 2. Form the URL that Runs the Data Discovery Search

To run the Data Discovery search, form a URL with the required parameters. The URL parameters determine on which Data Archive instance the search runs and the search criteria that Data Discovery uses to narrow the search results.

Form a URL to call the API that allows external applications to run Data Discovery queries.

The URL that calls the Data Discovery API has the following syntax:

```
http://<Data Archive Host>:<Data Archive Port>/api/appview?<Parameter String>
```

The URL syntax includes parameters that you have to replace with values when you form the URL. The parameters are in separate brackets. The URL syntax includes the following parameters:

Host

Host of the Data Archive instance on which you want to run the Data Discovery search.

Port

Port of the Data Archive instance on which you want to run the Data Discovery search.

Parameter String

Data Discovery search parameters. The parameter string includes the parameters that Data Discovery uses to narrow the search results. You can configure the parameter string to search the Data Vault by repository ID or repository name. You can also specify the entity ID, column names, and column values to include in your search.

Use the following syntax for the parameter string to search on the repository ID:

```
repid=<Repository ID>&entityid=<Entity ID>&column=<Column Name 1,Column Name
2>&value=<Value for Column 1|Value for Column 2>
```

Use the following syntax for the parameter string to search on the repository name:

```
repname=<Repository Name>&entityid=<Entity ID>&column=<Column Name 1,Column Name
2>&value=<Value for Column 1|Value for Column 2>
```

The following table describes the parameters for the parameter strings:

Parameter	Description
Repository ID	ID of the archive target connection.
Repository Name	Name of the archive target connection.
Entity ID	ID of the entity that you want to search.
Column	Columns that you want to search on. Use a comma to separate multiple values.
Value	Values for the column. Enter a value for each column that you include in the search. Use a pipe symbol (|) to separate values for multiple columns. Enter the values in the same order as you entered the columns.

URL Syntax Example

The syntax of the parameter string you use depends on whether you search by repository ID or repository name.

Search by Repository ID

You want to run a Data Discovery search based on the repository ID.

The following table lists the parameter values that you want to use for the search:

Parameter	Value
Data Archive Host	10.12.12.12
Data Archive Port	8080
Repository ID	4
Entity ID	3
Column 1	ID
Value for Column 1	1

Formulate the following URL to include the parameter values:

```
http://10.12.12.12:8080/appview?repid=4&entityid=3&column=ID&value=1
```

Search by Repository Name

You want to run a Data Discovery search based on the repository name.

The following table lists the parameter values that you want to use for the search:

Parameter	Value
Data Archive Host	10.12.12.12
Data Archive Port	8080
Repository Name	ILM_AUDIT_LOGS
Entity ID	1027
Column 1	HEADER_ID
Value for Column 1	15

Formulate the following URL to include the parameter values:

```
http://10.12.12.12:8080/appview?
repname=ILM_AUDIT_LOGS&entityid=1027&column=HEADER_ID&value=15
```

Step 3. Add the URL to the External Application Code

Add the URL to the external application code to run the Data Discovery search. When the code calls the URL, the URL runs the Data Discovery search and returns the search results in the App View.

CHAPTER 13

Security

This chapter includes the following topics:

Security Overview

Security in Data Archive is based on a system of users and the roles assigned to each user. Different roles grant a user different privileges to perform tasks in Data Archive. Users may have multiple roles.

There are two categories of predefined roles that you can assign to a user: system-defined roles and Data Vault access roles. System-defined roles grant privileges for tasks within Data Archive, such as creating archiving projects or accessing the Data Discovery portal. Data Vault access roles determine the data that users can access in Data Discovery searches.

After you have created users and assigned roles, you can create security groups. Security groups determine which users can access a source connection, and what data the assigned users can access on the source connection.

You can add a layer of security to data in Data Vault by enabling Dynamic Data Masking. With data masking, users can access the entities that they have access to but not the sensitive information in the entity. For example, a user has the same access role as an entity that contains credit card numbers. However, the user does not have permissions to view credit card numbers. When the user accesses the entity through the Data Visualization, Data Browse, or Search Data Vault options, the credit card numbers either do not display or display with different characters.

Users

To log in and perform tasks in Data Archive, you must have a user account. The tasks that you can perform depend on the roles assigned to your user account.

The default administrator user created during the installation process includes all system-defined roles and all privileges for Data Archive. Use the default administrator user to log in to Data Archive and to create other user accounts.

Typically, the default administrator user creates and manages user accounts and assigns roles to user accounts. However, any user with the administrator role can create user accounts and assign roles.

You can manage user accounts in Data Archive. Or, you can configure LDAP user authentication to synchronize users from the LDAP directory service to Data Archive. If you enable LDAP user authentication, you must maintain user details in the LDAP directory service.

User Properties

When you create or edit users, you configure properties for general information, login information, login options, and roles.

General Information

The following table describes general information that you configure for a user:

Field	Description
Full name	Full name of the user. This can consist of letters, spaces, apostrophes ('), and dashes (-).
Email	The user receives notifications at this email address.
Department	Department or organization of the user. This can consist of letters, spaces, apostrophes ('), and dashes (-).

Login Information

The following table describes login information that you configure for a user:

Field	Description
Login ID	Login ID for the user account. The login ID can include letters and digits.
Password	Initial password for user account. You cannot include the following characters in a password: ! | / = $ & @ " ' ` , The password can include up to a maximum of 40 characters.
Valid From	Start date of the user account validity period.
Valid Until	End date of the user account validity period.

Login Options

The following table describes login options that you configure for a user:

Field	Description
Login is (Enabled / Disabled / Locked)	User account status.
Must change password at next login	Forces the user to choose a new password at the next login.
Password never expires	Overrides the global setting to prompt for password change after a specified period.

Roles

The following table describes properties that you configure for a role:

Field	Description
Role	List of system-defined roles and Data Vault access roles.
Product	Product for which the role is applicable.
Valid from	Start date of the user account validity period.
Valid until	End date of the user account validity period.
Status	Allows the administrator to enable, disable, or lock the role.

Password Management

The installation process creates a default administrator user account. Both the login ID and the password are `amadmin`.

Change the default password the first time that you log in.

User Management

You can create and edit users. You can also generate role reports for specific users.

Editing Users

After you create a user, you can edit the user properties and role assignments.

1. Click **Administration** > **Manage Users**.
2. Click the **Edit** icon next to the desired user.

   The **Edit User** page appears.
3. Edit the user properties and click **Save**.

Creating Users

When you create a user, you provide properties for contact and login information. You can also assign, disable, or remove roles.

If you enabled LDAP user authentication, you must create and edit users in the LDAP directory service. You can view user details and edit the default administrator user account in Data Archive.

1. Click **Administration** > **Manage Users**.
2. Click **New User**.

   The **New User** page appears.
3. Enter the user properties.
4. To add a role assignment, click **Add Role** and select the role. Then enter the validity dates for the role.
5. Click **Save**.

Role Reports for Users

You can generate a report that lists the roles that are assigned to a user.

To view a role report for a user, click the **View Roles** link on the **Manage Users** page.

The report shows the following information:

General Information

General details about the user, such as the user login ID, name, department, email address, status, and dates that the role is valid.

Report Generation Time

Time that the report was generated. The report uses the time zone of the ILM application server location.

System-Defined Roles

System-defined roles that are assigned to the user and the period of time that the role is valid.

Data Vault Access Roles

Data Vault access roles that are assigned to the user and the period of time that the role is valid.

System-Defined Roles

A system-defined role is a collection of predefined privileges delivered with the product. Privileges determine specific tasks a user can perform in Data Archive and in Enterprise Data Manager. Any user with the system-defined role can perform all of the privileges included in the role.

System-defined roles restrict the menu paths that are available in Data Archive. For example, only users with the discovery user role have access to the **Data Discovery** menu. All system-defined roles include access to the **Home** and **Help** menus.

System-defined roles also determine the tasks that are available within the menus, such as display and edit privileges. For example, the administrator role can create and edit connections. The operator role can only view connections.

The default administrator user, AMADMIN, includes all system-defined roles except for the healthcare metadata administrator role. Use the default administrator user to create users and to assign roles to users.

You can assign one or more system-defined roles to a user. By default, users do not include role assignments. You must manually assign roles when you create users.

You can add or change role assignments to users. Role changes take effect the next time the user logs in.

System-Defined Roles and Privileges

Each role has a set of predefined privileges. Some roles include privileges of other roles. For example, the administrator role includes all privileges of the operator role.

The following table lists the system-defined roles and corresponding privileges:

System-Defined Role	Privileges
Administrator	- Create, edit, enable, and disable users. - Create, edit, and delete security groups. - Edit the system profile. - Edit the user profile. - View, create, and edit archive source and target connections. - Map datatypes from source connection to Data Vault. - View the dashboard. - Configure audit logs. - Upload JDBC driver. - All operator role privileges. - All audit log viewer role privileges. - All archive developer role privileges. - Encryption user privileges. - Select the status of a deviated table when you review the results of the integrated validation process.
Archive Security Administrator	- Assign access roles to data in the Data Vault. - Create and edit access roles. - Assign access roles to users and groups. - View security reports. - Map datatypes from source connection to Data Vault. - All discovery user role privileges.
Archive User	- Create or modify archive projects. - Restore from a database or Data Vault. - All operator role privileges.
Archive Developer	- Create and edit metadata in the Enterprise Data Manager. - View and edit Data Discovery search options. - All operator role privileges.
Audit Log Viewer	- View, search, print, and export audit logs.
Discovery Technical User	- View Data Discovery search results in the application view. - View Data Discovery search results in the technical view. The role includes privileges to access the technical view when a XSL style sheet is available. If an entity has a configured style sheet and the user does not have this role, the user can only view the data through the style sheet. - All discovery user role privileges.

System-Defined Role	Privileges
Discovery User	- Search the Data Vault and legal holds. - Add or remove legal hold tags. - Browse catalogs. - Browse data. - Export data to PDF files and schedule export jobs. - Preview search conditions in Data Discovery. The following roles include all of the discovery user role privileges, in addition to the individual role privileges: - Discovery technical user - Export administrator - Legal hold user - Retention administrator - Archive security administrator - Tag administrator - Tag viewer
Encryption User	- Schedule the encrypt data in Data Vault job. - Must be coupled with a role that can schedule jobs.
Export Administrator	- Export data from Data Discovery searches to supported file types, such as XML, PDF, and comma separated text files. Users can export search results from Data Discovery when you browse data or search the Data Vault. - Export data from audit logs to supported file types, such as XML, PDF, and comma separated text files. - All discovery user role privileges.
Healthcare Information Management User	- Access the release of information form in the Application Retirement for Healthcare accelerator. - Submit the release of information form.
Healthcare Metadata Administrator	- View, edit, and create materialized views in the Patient Archives. Applicable to users of the Application Retirement for Healthcare accelerator. The healthcare metadata administrator role is not given to any user by default, including the AMADMIN user.
Import Metadata	- Import metadata from a database in the Enterprise Data Manager. - Import metadata, either traditionally or using enhanced import, in the Enterprise Data Manager.
Legal Hold User	- Manage legal hold groups. - Apply or remove legal hold tags for records in the Data Vault. - Schedule retention-related legal hold jobs. - All discovery user role privileges.
Migration Administrator	- Run and monitor and progress of the standalone jobs required for the retirement migration process, such as the Export Informatica Data Vault Metadata job, Migrate Data Archive Metadata job, and Import Data Vault Metadata job.

System-Defined Role	Privileges
Operator	- Monitor jobs. - View archive and retirement project definitions. - View metadata in the Enterprise Data Manager. - View source and target connections. - Map datatypes from source connection to the Data Vault Service. - View the system profile. - View the job log. - Review the results of the integrated validation process. The following roles include all of the operator role privileges, in addition to the individual role privileges: - Administrator - Archive developer - Archive user - Retention administrator - Scheduler
Report Admin	If you installed the Data Visualization component, this role allows the user to create, run, copy, and delete reports on data stored in the Data Vault. This role also allows the user to grant other users and access roles the run, copy, delete, and grant permissions.
Report Designer	If you installed the Data Visualization component, this role allows the user to create reports on data stored in the Data Vault. For the user to be able to run, copy, or delete a report, this role must be granted those permissions specifically by the Report Admin user or another user with the privilege to grant report permissions. Users can create and run reports only if they have the same Data Vault access role as the entity used to create the report. If you upgraded from a previous version of Data Archive, this role retains the same permissions it formerly had on any pre-existing reports.
Report Viewer	If you installed the Data Visualization component, this role allows the user to view and run reports on data stored in the Data Vault. In order to run reports, this role must be granted the run permission by the Report Admin or another user with the privilege to grant report permissions. Users can run reports only if they have the same Data Vault access role as the entity used to create the report. If you upgraded from a previous version of Data Archive, this role retains the same permissions it formerly had on any pre-existing reports.
Retention Administrator	- Create and edit retention policies. - Assign retention policies to records in the Data Vault. - Schedule retention-related jobs. - Perform data discovery searches and Data Vault restore searches based on retention policy and retention expiration date. - Manage retention jobs. - All discovery user role privileges. - All operator role privileges.
Retention Viewer	- View retention policy details (retention policy and expiration date). - View table data. In order to export data from a search, this role must be coupled with the export administrator role.

System-Defined Role	Privileges
SAP Portal User	If you installed the Data Visualization component, this user can access the SAP Archives under the Data Visualization menu. When you assign this role to a user, Data Archive launches the SAP Archives when the user logs in.
Scheduler	- Schedule, resume, and terminate archive and retirement projects. - Delete, manage, schedule, search, resume, terminate, and view jobs. - All operator role privileges.
Tag Administrator	- Create, edit, or delete tags. - Add or remove tags from archived records. - Schedule tag-related jobs. - All discovery user role privileges.
Tag Viewer	- View tags that are associated with archived records. - All discovery user role privileges.

User Reports for System-Defined Roles

You can generate a report that shows a list of users that are assigned to a role. You generate the report when you view the list of system-defined roles from the Manage Roles transaction.

The report is available as link next to the system-defined role. When you click the link, Data Archive generates the report and opens it in PDF format.

The report shows the following information:

General Information

General details about the system-defined role, such as the name and the validity period of the role.

Report Generation Time

Time that the report was generated. The report uses the time zone of the ILM application server location.

Users

List of all users that are assigned to the role including user login ID, name, department, email address, and the user status.

Data Vault Access Roles

Data Vault access roles determine the data that users can access in Data Discovery searches. Data Vault access roles also determine the data that users can access when they create and view custom reports. You can restrict data access at the entity level and at the archive or retirement project level.

Data Vault access roles also restrict which users can view and download files that contain exported search results. Users can only download the exported files if the user is assigned to the same access role as the entity. System-defined roles determine which files users can export to.

When you create a Data Vault access role, you specify the date that the role becomes valid and the date that the role expires. Access privileges expire with the role. Changes to the period of time that a role is valid are dynamic.

If you installed the Data Visualization component, Data Vault access roles determine the entities that users can access to create reports. To create and manage reports, users must have the same role assignment as the entities that contain the data that they want to include in the reports.

You can create multiple Data Vault access roles. The number of Data Vault access roles that you create depends on the type of security that you want to implement. For example, you might want to create one Data Vault access role for all entities within an application.

Data Vault Access Role Properties

You define the Data Vault access role properties when you create or edit a Data Vault access role.

Data Vault access roles include the following properties:

Role Name

Unique name for the role. Note that after you create the role, you cannot edit the role name. You cannot use the following special characters in the role name: < >

Description

Description of the role.

Valid From

Start date of the period of time that the role is valid.

Valid Until

End date of the period of time that the role is valid. The date is optional. By default, roles do not have an end date unless you specify one. Roles without an end date are valid indefinitely.

Data Vault Access Role Assignments

The Data Vault access role assignments determine the data that users can see when they run Data Discovery searches. The role assignments also determine the data that users can see when they view and create custom reports and dashboards.

After you create the Data Vault access roles, you assign the Data Vault access roles to one or more of the following objects:

Users

A Data Vault access role assignment is required. Users can only access data for entities or projects that have the same Data Vault access role assignment as the user.

Entities

A Data Vault access role assignment is optional. If you assign a Data Vault access role to an entity, access is restricted to all archived data from the entity, regardless of the project that includes the entity. Only users that have the Data Discovery role and the access role that is assigned to the entity can access the archived data from the corresponding entity.

If you view or create custom reports and dashboards, you must assign a Data Vault access role to each entity on which you want to base a report. Users that create or view a report must have the same role assignment as the entity on which the report is based.

Archive or Retirement Projects

A Data Vault access role assignment is optional for projects that have the Data Vault as the target connection. If you assign an access role to an archive or retirement project, access is restricted to data that is archived from the project. The assignment at the project level overrides the assignment at the entity level. If the project include entities that have role assignments, Data Discovery ignores the entity

level assignment. Only users that have the Data Discovery role and the Data Vault access role that is assigned to the project can access the archived data.

You must assign a Data Vault access role to an entity or to a project. If you do not assign a Data Vault access role to an entity or to a project, no users have access to the archived data.

When you assign Data Vault access roles to users or entities, you specify a validity period of the role assignment. You determine when the role assignment begins. For example, you can assign a Data Vault access role to a user, but make the role assignment effective in two months. You can also determine when the role assignment ends. For example, you may want to assign a Data Vault access role to a user for a limited amount of time, such as a few months. By default, all role assignments do not have an end date unless you specify one.

Data Vault Access Roles Management

You can create, edit, and assign Data Vault access roles to restrict access to archived data.

For auditing purposes, you cannot delete a Data Vault access role. To make a role obsolete, you can give the role a Valid Until date that is in the past.

Creating Data Vault Access Roles

After you determine the level of access that you want users to have to archived data, create the Data Vault access roles. When you create a Data Vault access role, you define the name, description, and a period of time that the role is valid.

1. Click **Administration** > **Manage Roles**.
2. Click **New Role**.

   The role name cannot contain special characters.
3. Enter the Data Vault access role properties.
4. Click **Save**.

Assigning Data Vault Access Roles to Entities

After you create the Data Vault access roles, assign the roles to entities, archive projects, or retirement projects. You assign roles to archive or retirement projects when you create or edit the projects. You assign roles to entities when you manage roles.

1. Click **Administration** > **Manage Roles**.
2. Click **Assign Role to Entity**.
3. Choose the entity that you want to assign a Data Vault access role to.
4. Click **Add Role**.

   If the entity has role assignments, a list of the assigned roles appears.
5. Click **Add Role**.

   A new line appears in the list of roles.
6. Select the role that you want to assign and enter the validity dates of the role assignment.
7. Click **Save**.

Assigning Data Vault Access Roles to Users

After you assign Data Vault access roles to entities, archive projects, or retirement projects, assign the roles to users.

1. Click **Administration** > **Manage Users**.

   A list of users appears.
2. Click **Edit** next to the user you want to assign the role to.

   The user details appears.
3. Click **Add Role**.

   A new line appears in the list of roles.
4. Select the role that you want to assign and enter the validity dates of the role assignment.
5. Click **Save.**

Editing Data Vault Access Roles

After you create a role, you can edit the role name, description, or validity dates.

1. Click **Administration** > **Manage Roles.**
2. Click the **Edit** icon.
3. Edit the role properties.
4. Click **Save**.

Disabling or Removing User Assignments

After you assign a Data Vault access role to a user, you can disable or remove the user assignment.

1. Click **Administration** > **Manage Users.**
2. Click the **Edit** icon next to the user assigned to the role you want to disable.

   The **Edit User** page appears.
3. Disable or remove the role for the user.
   - To disable or lock the role for the user, select **Disabled** or **Locked** from the drop-down **Status** menu.
   - To remove the role for the user, click the **Delete** icon.
4. Click **Save**.

Disabling or Removing Entity Assignments

To disable or remove an entity assignment from a Data Vault access role, you must know the entities that the role is assigned to. Generate an associated entities report for the role to view all associated entities.

1. Click **Administration** > **Manage Roles**.
2. Click **Assign Access Role to Entity**.
3. Select the application version and module for the entity you want to disable or remove. Then select the entity.

   The roles assigned to the entity appear.
4. Disable or remove the entity assignment for the role.
   - To disable the entity assignment, enter a **Valid Until** date that is in the past.

- To remove the entity assignment, click the **Delete** icon.

5. Click **Save**.

Data Vault Access Role, User, Entity Relationships

You can view the Associated Users and the Associated Entities reports for the list of users and entities associated with a Data Vault access role. To view a list of Data Vault access roles that can access a certain entity, select the entity and view the list of access roles on the **Manage Roles** page.

Associated Users Report for Data Vault Access Roles

The Associated Users report contains the list of users assigned to a Data Vault access role.

To access the Associated Users report, go to **Administration** > **Manage Roles** and click the **Access Roles** tab. Click the PDF icon in the Associated Users column for the row of the file access role you are interested in.

The Associated Users report contains the following information:

General Information

General details about the Data Vault access role, such as the name, description, and validity period of the role.

Report Generation Time

Time that the report was generated. The report uses the time zone of the ILM application server location.

Users

List of all users that are assigned to the role including user login ID, name, department, email address, and the user status.

Associated Entities Report for Data Vault Access Roles

The Associated Entities report contains the list of entities a Data Vault access role can access.

To access the Associated Entities report, go to **Administration** > **Manage Roles** and click the **Access Roles** tab. Click the PDF icon in the Associated Entities column, on the row of the file access role you are interested in.

The Associated Entities report contains the following information:

General Information

General details about the Data Vault access role, such as the name, description, and validity period of the role.

Report Generation Time

Time that the report was generated. The report uses the time zone of the ILM application server location.

Entities

List of all entities that this role has access to. For each entity listed, this section includes the application version, name of the application module, and name of the entity.

Viewing Data Vault Access Roles Associated With an Entity

You can view the list of Data Vault access roles with access to an entity via the **Assign Role to Entity** transaction on the **Manage Roles** page.

1. Go to **Administration** > **Manage Roles**.
2. Click **Assign Role to Entity**.
3. Select the appropriate **Application Version**, **Application**, and **Entity Name** from the list of values.

   A list of Data Vault access roles with access to the selected entity appears.

Security Groups

A security group is a collection of permissions that determine what data users can access from a source database when you run a database archive project.

You define security groups for source connections. The permissions in the security group determine the entities that a user can select when addomg entities to archive and retirement projects.

When you create a security group, you specify the connection, define the permissions, and add users. If a user does not belong to a security group, the user cannot select entities in an archive or retirement project.

Security Group Properties

When you create or edit security groups, you configure properties for general information, permissions, and users.

General Information

The following table describes general information that you configure for a security group:

Property	Description
Security Group Name	Unique name for the security group.
Connection	Source connection that you want to configure the security group for.
Description	Description of the security group.
Enabled	Activates the security group. Clear the field to disable the security group.

Permissions

The following table describes permissions properties that you configure for a security group:

Property	Description
Scope	Scope of the data that the permission applies to. You can select an application, group of applications such as custom applications, or entities.
Values	The value of the scope provided. For example, if you select custom applications, you can specify a particular application. If you leave this field blank, users of the security group will have access to all applications or entities defined in the scope field.
Enabled (check box)	Enables the permission you configured. Clear the field to disable the permission.

Users

The following table describes user properties that you configure for a security group:

Property	Description
Full Name	Name of the user to add to the security group.
Valid From	Start date of the user as a security group member.
Valid Until	End date of the user as a security group member.
Enabled	Enables the user as a member of the security group. Clear the check box to disable the user as a member of the security group.

Creating Security Groups

When you create security groups, you configure the security group properties. You must add and configure at least one permission.

You must define source connections and users before you create security groups.

1. Click **Administration** > **Manage Security Groups**.
2. Click **New Security Group**.

   The **Create or Edit a Security Group** page appears.
3. Enter the general information properties.
4. Click **Add Permission**.

   A window appears with a list of defined permissions.
5. Select the permission and click **Select**.

   The permission appears in the list of permissions.
6. Enter the permissions properties.
7. Click **Add User**.

   A window appears with a list of users.
8. Select a user and click **Select**.

   The user appears in the list of users.

9. Enter the user properties.
10. Click **Save**.

CHAPTER 14

SSL Communication with Data Vault

This chapter includes the following topics:

SSL Communication with Data Vault Overview

Data Vault uses TLS1.2 technology to secure communication between the Data Vault service and clients such as ODBC, JDBC, SSASQL, Data Archive, and more.

To configure Data Vault for SSL communication, see the *Data Vault Administrator Guide*.

When SSL is enabled in Data Vault, you must also configure Data Archive to communicate with the Data Vault service. When the Data Vault service is SSL-enabled, any clients must have a client-side certificate provided by the Data Vault service.

After you have the required client-side certificate for Data Archive, you must configure the `informia.idv.ssl.certificatePath` parameter in the `conf.properties` file. Then, add the file path and password for the trustStore to the `startApplimation.sh` or `startApplimation.bat` file.

If you use the Enterprise Data Manager to connect to Data Vault with SSL enabled, then you must provide SSL=1 in the JDBC URL and add the client certificate in the default Java security trustStore.

If you use the JReport Designer to connect to Data Vault with SSL enabled, you must configure the trustStore in the JReport catalog and append SSL=1 in the catalog connection name when the catalog is created from JReport Designer.

Creating a Certificate and TrustStore

Create a client-side certificate and a trustStore.

Create a client-side certificate and add it to a trustStore using the Java keytool:

```
keytool -import -alias test -file <certificate path> -keystore <xxxx>.jks
```

Provide the password for the trustStore and trust the certificate.

For more information about certificates and trustStores, see the *Data Vault Administrator Guide*.

Updating Conf.Properties

To enable SSL communication with Data Vault from Data Archive, first update the `conf.properties` file.

Uncomment and update the parameter "informia.idv.ssl.certificatePath" in the `conf.properties` file. This parameter defines the file path to the client-side certificate.

For example:

```
informia.idv.ssl.certificatePath=/home/ora12c/jenkins/demossl.pem
```

Modifying the StartApplimation File

Modify the `startApplimation.bat` or `startApplimation.sh` file to include the file path and password to the trustStore.

At the end of the startApplimation.bat or startApplimation.sh file, locate the Start Applimation section and add the following arguments to the file:

```
-Djavax.net.ssl.trustStore=<path of trustStore file>
-Djavax.net.ssl.trustStorePassword=<trustStore password>
```

```
if [ $VERSION_NUMBER -le $REQUIRED_VERSION_NUMBER ]; then
        $JAVA_HOME/bin/java -Dfile.encoding=UTF-8 -Djavax.net.ssl.trustStore=/home/ora12c/jenkins/cacerts.jks -Djavax.net.ssl.trustStorePassword=password -Duser.language=en -Dapp.home=
        $SEARCHDIR -Djava.awt.headless=true -Xms512m -Xmx2048m -XX:MaxPermSize=256m -jar ./lib/informia-server-6.4.3-HF1.jar start $@
```

Depending on the operating system/platform, add the trustStore details in the following locations:

SPARC Platform

```
if [ "$PLATFORM" = "sparc" ]; then
    if [ ! -f $JAVA_HOME/bin/sparcv9/java ];
    then
        echo "JAVA_HOME needs to have a 64bit JVM installed on SPARC"
        echo "Please install 64bit JAVA to proceed with running Data Archive"
        exit 1
    fi
    if [ $VERSION_NUMBER -le $REQUIRED_VERSION_NUMBER ];  then
        LD_PRELOAD=$JAVA_HOME/jre/lib/sparcv9/libjsig.so $JAVA_HOME/bin/sparcv9/java -Djavax.net.ssl.trustStore= /home/trsustore.jks -Djavax.net.ssl.trustStorePassword=password -
        Djavac_path=$JAVA_HOME/bin/sparcv9/javac -Dfile.encoding=UTF-8 -Dapp.home=$SEARCHDIR -Djava.awt.headless=true -Xms512m -Xmx2048m -XX:MaxPermSize=256m -jar ./lib/
        informia-server-6.4.3-HF1.jar start $@
    else
        LD_PRELOAD=$JAVA_HOME/jre/lib/sparcv9/libjsig.so $JAVA_HOME/bin/sparcv9/java -Djavac_path=$JAVA_HOME/bin/sparcv9/javac -Djavax.net.ssl.trustStore= /home/trsustore.jks -
        Djavax.net.ssl.trustStorePassword=password -Dfile.encoding=UTF-8 -Dapp.home=$SEARCHDIR -Djava.awt.headless=true -Xms512m -Xmx2048m -jar ./lib/informia-server-6.4.3-HF1.jar
        start $@
    fi
```

i386 Platform

```
elif [ "$PLATFORM" = "i386" ]; then
    if [ ! -f $JAVA_HOME/bin/amd64/java ];
    then
        echo "JAVA_HOME needs to have a 64bit JVM installed on SPARC"
        echo "Please install 64bit JAVA to proceed with running Data Archive"
        exit 1
    fi
    if [ $VERSION_NUMBER -le $REQUIRED_VERSION_NUMBER ];  then
        LD_PRELOAD=$JAVA_HOME/jre/lib/amd64/libjsig.so $JAVA_HOME/bin/amd64/java -Dfile.encoding=UTF-8 -Dapp.home=$SEARCHDIR -Djavax.net.ssl.trustStore= /home/trsustore.jks -
        Djavax.net.ssl.trustStorePassword=password -Djava.awt.headless=true -Xms512m -Xmx2048m -XX:PermSize=32m -XX:MaxPermSize=128m -Xss2m -XX:+UseConcMarkSweepGC -XX:+
        CMSClassUnloadingEnabled -XX:+CMSPermGenSweepingEnabled  -jar ./lib/informia-server-6.4.3-HF1.jar start $@
    else
        LD_PRELOAD=$JAVA_HOME/jre/lib/amd64/libjsig.so $JAVA_HOME/bin/amd64/java -Dfile.encoding=UTF-8 -Djavax.net.ssl.trustStore= /home/trsustore.jks -
        Djavax.net.ssl.trustStorePassword=password -Dapp.home=$SEARCHDIR -Djava.awt.headless=true -Xms512m -Xmx2048m -XX:PermSize=32m -Xss2m -XX:+UseConcMarkSweepGC -XX:+
        CMSClassUnloadingEnabled -XX:+CMSPermGenSweepingEnabled  -jar ./lib/informia-server-6.4.3-HF1.jar start $@
    fi
```

Linux Operating System

```
else
    if [ $VERSION_NUMBER -le $REQUIRED_VERSION_NUMBER ]; then
            $JAVA_HOME/bin/java -Dfile.encoding=UTF-8 -Duser.language=en -Dapp.home=$SEARCHDIR -Djavax.net.ssl.trustStore= /home/trsustore.jks -Djavax.net.ssl.trustStorePassword=password -
            Djava.awt.headless=true -Xms512m -Xmx2048m -XX:MaxPermSize=256m -jar ./lib/informia-server-6.4.3-HF1.jar start $@
    else
            $JAVA_HOME/bin/java -Dfile.encoding=UTF-8 -Duser.language=en -Dapp.home=$SEARCHDIR -Djavax.net.ssl.trustStore= /home/trsustore.jks -Djavax.net.ssl.trustStorePassword=password -
            Djava.awt.headless=true -Xms512m -Xmx2048m -jar ./lib/informia-server-6.4.3-HF1.jar start $@
    fi
```

Microsoft Windows Operating System

```
if %JAVA_VERSION_NUMBER% leq %REQUIRED_JAVA_VERSION% (
"%JAVA_HOME%\bin\java.exe" -Dfile.encoding=UTF-8 -Djavax.net.ssl.trustStore=c:\\ssl\\tester.jks -Djavax.net.ssl.trustStorePassword=password -Xdebug -Xrunjdwp:transport=dt_socket,address=
5666,server=y,suspend=n -Duser.language=en -Duser.language=en -Dapp.home="%SEARCHDIR%" -Djava.awt.headless=true -Xms512m -Xmx1024m -XX:MaxPermSize=256m -jar "%WORKLIBDIR%
/informia-server-6.4.3-HF1.jar" start %CMD_LINE_ARGS%

) else (
"%JAVA_HOME%\bin\java.exe" -Dfile.encoding=UTF-8 -Djavax.net.ssl.trustStore=c:\\ssl\\tester.jks -Djavax.net.ssl.trustStorePassword=password -Xdebug -Xrunjdwp:transport=dt_socket,address=
5666,server=y,suspend=n -Duser.language=en -Duser.language=en -Dapp.home="%SEARCHDIR%" -Djava.awt.headless=true -Xms512m -Xmx1024m -jar "%WORKLIBDIR%/informia-server-6.4.3-HF1.jar" start
%CMD_LINE_ARGS%
)
:end
endlocal
```

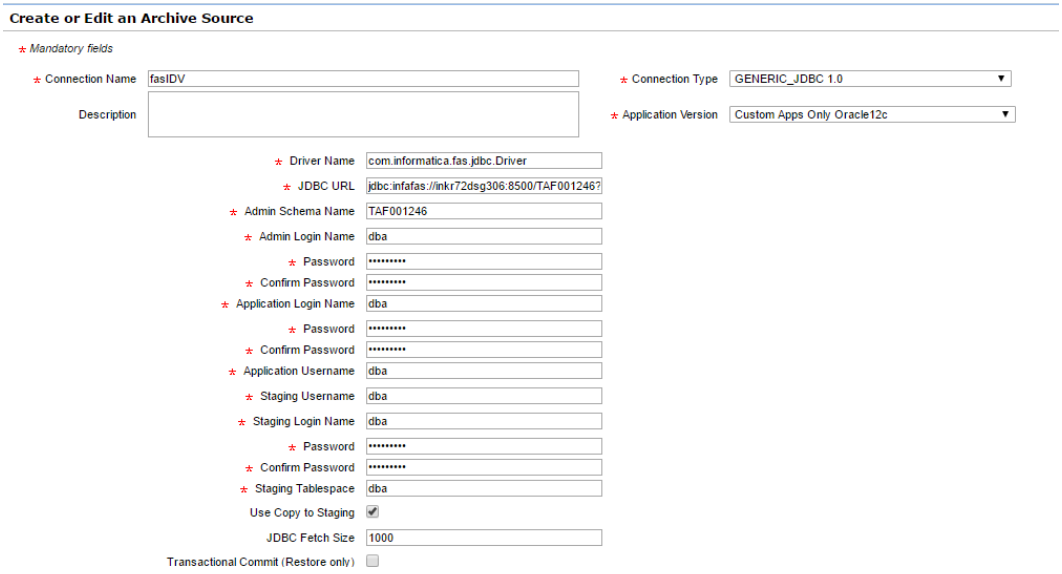
Creating a Generic JDBC Source Connection

Use a generic JDBC source connection to the Data Vault when SSL is enabled.

The JDBC URL should contain SSL=1. For example:

`jdbc:infafas://ivw12ilm07:8564/test?SSL=1`

The following image is an example of the generic JDBC source connection details:



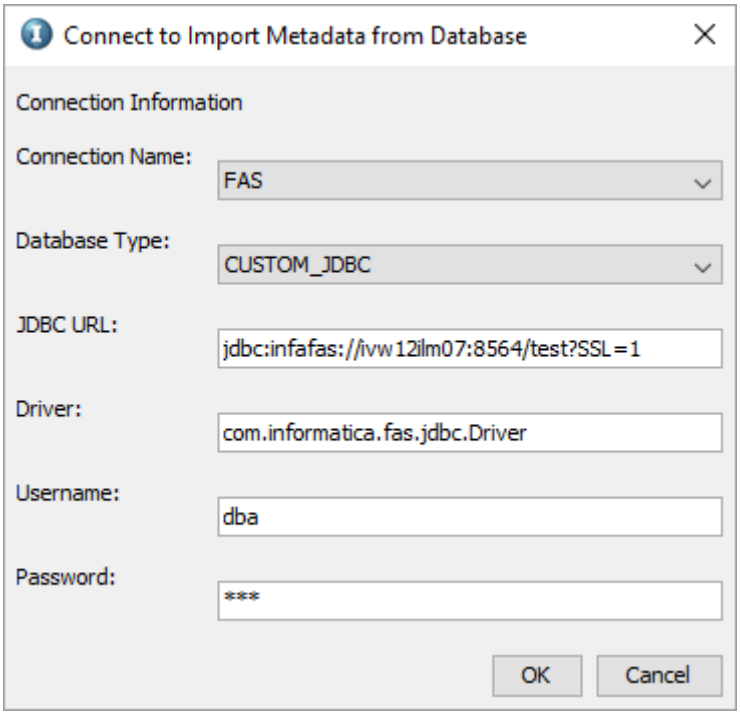
Importing Data Vault Table Metadata with SSL Enabled

To use Enterprise Data Manager to import Data Vault Table metadata with SSL enabled, add the client-side certificate to Java's default trustStore under `jre/lib/security` in the client machine where the Enterprise Data Manager runs.

For example:

```
keytool -keystore cacerts -importcert -alias password -file c:\ssl\demo.pem
```

When you connect to Data Vault to import metadata through the Enterprise Data Manager, verify that the JDBC URL contains "SSL=1," as shown in the following image:



Updating JReport Designer

To use JReport Designer with SSL-enabled Data Vault, you must update JReport Designer with the trustStore details and update the JReport catalog connection JDBC URL with SSL=1, if the URL does not already contain it.

1. To connect Data Vault to JReport Designer, update `jReport.bat` under <Designer\bin> with the trustStore details.

   Djavax.net.ssl.trustStore=<path of trustStore file>

   Djavax.net.ssl.trustStorePassword=<trustStore password>

```
"%JAVAHOME%\bin\java.exe" -Xms40m -Xmx1024m "-Dinstall.root=%REPORTHOME%" -classpath "%CLASSPATH%" -Dreporthome="%REPORTHOME%" -Djreport.url.encoding="UTF-8" -Djavax.net.ssl.trustStore=
"C:\SSL\caserts.jks" -Djavax.net.ssl.trustStorePassword="password" com.jinfonet.designer.JReport %1 %2 %3 %4 %5 %6 %7 %8 %9
```

2. To connect Data Vault to Jinfonet Server, update `jrsever.bat` under < designer\server\bin > with the trustStore details.

   Djavax.net.ssl.trustStore=<path of trustStore file>

   Djavax.net.ssl.trustStorePassword=<trustStore password>

```
"%JAVAHOME%\bin\java.exe" -Xms40m -Xmx1024m "-Dinstall.root=%REPORTHOME%" -classpath "%CLASSPATH%" -Dreporthome="%REPORTHOME%" -Djreport.url.encoding="UTF-8" -Djavax.net.ssl.trustStore=
"C:\SSL\caserts.jks" -Djavax.net.ssl.trustStorePassword="password" com.jinfonet.designer.JReport %1 %2 %3 %4 %5 %6 %7 %8 %9
```

3. To modify the JReport catalog for SSL, open the catalog through Designer.
4. Verify that the catalog is editable. Select **File** > **Options** > **Catalog**.

5. Uncheck **Forbid editing data object properties**.
6. Change the existing catalog connection name by appending "SSL=1" to the end of the Name URL. For example:

   jdbc:infafas://dba@inkr79dsg184.informatica.com:8500/TAF133903?SSL=1

CHAPTER 15

LDAP User Authentication

This chapter includes the following topics:

LDAP User Authentication Overview

If you have user accounts in an enterprise LDAP directory service, you can configure Data Archive to use LDAP authentication. When you enable LDAP user authentication, you can synchronize users and role assignments. Note that corporate passwords are not stored in any product databases.

Create and manage LDAP users and groups in the LDAP directory service. Optionally, you can assign security groups to users and groups in the LDAP directory service and import the assignments to Data Archive. LDAP security groups are equivalent to Data Archive roles.

You can use the following LDAP directory services for LDAP authentication:

- Microsoft Active Directory Service
- Sun Java System Directory Service

If you manage users in Data Archive and want to implement LDAP authentication, you must create all users in the LDAP directory service. After you enable LDAP authentication, only users that are maintained in the LDAP directory service can log in to Data Archive.

User Synchronization

Data Archive uses a combination of synchronization methods to synchronize users with the LDAP directory service. You run a standalone job to synchronize users after you create users in the LDAP directory service. Data Archive automatically synchronizes any additional changes to users.

When you set up LDAP user authentication and when you create additional users in the LDAP directory service, you run the Sync with LDAP Server job. The job synchronizes users between the LDAP directory service and Data Archive.

After you run the job to synchronize users, Data Archive automatically synchronizes any additional changes to the users when users log in to Data Archive. When users log in, Data Archive silently runs the Sync with LDAP Server job in the background. Data Archive only runs the job for the user that logs in. If there are any changes to users in the LDAP directory service, the job updates the corresponding user account in Data Archive.

Nested Group Synchronization for Users

You can maintain users in nested groups in the LDAP directory service. You may want to use nested groups for organizational purposes and to group similar types of users. Attributes that you maintain at the group level apply to all users within the group.

You may want to group users by the types of privileges they should have in Data Archive. Then, you can assign roles to the group. By default, all users in the group inherit the role assignment. Assigning roles to groups saves time as you avoid individually assigning roles to every user. Note that LDAP security groups are equivalent to Data Archive roles.

Data Archive synchronizes users based on the group that you provide as a parameter when you run the Sync with LDAP Server job. The job reviews all users under the group, including users in nested groups. There is no limit to the number of nested groups that you can have in one group.

For example, you create the following structure in the LDAP directory service:

```
ILM Users (Parent LDAP Group)
---User 1 (LDAP Member)
---User 2 (LDAP Member)
---ILM Administrator Users (Nested LDAP Group)
------User 3 (LDAP Member)
------User 4 (LDAP Member)
```

The ILM Users group includes User 1, User 2, and the ILM Administrator Users nested group. The ILM Administrator Users group includes User 3 and User 4.

When you synchronize the ILM Users group, Data Archive synchronizes User 1, User 2, User 3, and User 4.

Data Archive Roles and LDAP Security Groups

When you configure LDAP authentication, you can enable role assignment synchronization. However, LDAP directory services do not have a role concept. Use security groups to represent Data Archive roles.

If you enable role assignment synchronization, create a security group for each system-defined role and Data Vault access role that exists in Data Archive. After you create security groups, you can assign the security groups to users and to groups of users. When Data Archive synchronizes users, Data Archive reviews the security groups that are assigned to users and groups. If a security group name matches the technical name of a Data Archive role, then Data Archive assigns the role to the user.

Use the following rules and guidelines when you create security groups in the LDAP directory service:

- Create one security group for each Data Archive system-defined role and Data Vault access role.
- The name of the LDAP security group for a Data Vault access role must match the unique name of that Data Vault access role.
  To get the names of Data Vault access roles, go to Data Archive, click **Administration** > **Manage Roles** and select the **Access Roles** tab. You will see the list of Data Vault access roles.
- The name of the LDAP security group for a system-defined role must match the technical name of that system-defined role.
  To get the technical names of system-defined roles, view the AM_ROLES table in the ILM repository or see the reference table below.
- LDAP security group names are not case sensitive.
- LDAP security groups are not managed as groups in Data Archive. You must enter the appropriate value in the *User Filter* field of the *Sync with LDAP Server* job to enable Data Archive to identify individual users in an LDAP security group.

Technical Names for System-Defined Roles

When you create security groups for Data Archive system-defined roles, use the technical name of the role as the security group name. Data Archive only assigns roles to users if the security group name is the same as the technical role name.

The following table lists the system-defined roles and corresponding technical names:

System-Defined Role	Technical Name
Administrator	SYSADMIN
Audit Log Viewer	AUDIT_LOG_VIEWER
Data Privacy User	DATA_PRIVACY_USER
Developer	DEVELOPER
Discovery Technical User	TECHVIEW_USER
Discovery User	SEARCH_USER
Encryption User	ENCRYPTION_USER
Export Administrator	EXPORT_ADMINISTRATOR
Healthcare Account Administrator	HEALTHCARE_ACCOUNT_ADMIN
Healthcare Information Management User	HIM_USER
Healthcare Portal User	HEALTHCARE_PORTAL_USER
Import Metadata	IMPORT_METADATA
Legal Hold User	Legal_Hold_User
Migration Administrator	Migration_Administrator

System-Defined Role	Technical Name
Operator	OPERATOR
Report Administrator	REPORT_ADMIN
Report Designer	REPORT_DESIGNER
Report Viewer	REPORT_VIEWER
Retention Administrator	RETENTION_ADMINISTRATOR
Retention Viewer	RETENTION_VIEWER
SAP Portal User	SAP_PORTAL_USER
Scheduler	SCHEDULER
Security Administrator	SECURITY_ADMINISTRATOR
Tag Administrator	Tag_Administrator
Tag Viewer	Tag_Viewer
User	USER

Role Assignments

To perform tasks and to access data in Data Archive, you must assign users to roles. Roles determine the tasks that users can perform or the data that users can access. The way you assign roles to users depends on if you enable role assignment synchronization. You enable role assignment synchronization in the `conf.properties` file.

When you run the Sync with LDAP Server job, the job assigns the User system-defined role to the user account. You need the User role to log in to Data Archive. You need to assign any relevant additional roles to the user. You can assign roles to users in the LDAP directory service, or you can assign roles in Data Archive after you synchronize users.

If you enable role assignment synchronization, then you maintain role assignments in the LDAP directory service. You add roles to users and maintain subsequent role changes in the LDAP directory service. The role assignments are synchronized when users log in to Data Archive. You can view the role assignments when you view the user account in Data Archive.

If you do not enable role assignment synchronization, then you maintain role assignments in Data Archive after you synchronize users. You add roles directly to the user account after the initial user synchronization. You maintain any subsequent role changes in Data Archive.

Role Assignment Synchronization

If you enable role assignment synchronization, then you manage role assignments for users in the LDAP directory service. Role assignments are synchronized to the user account in Data Archive when users log in to Data Archive. Role assignments are only synchronized for users that exist in Data Archive.

When users log in, Data Archive connects to the LDAP directory service and reviews all of the role assignments for the user. Data Archive synchronizes the user account to match role assignments in the LDAP directory service. If there are any changes, such as if you add or delete a role assignment in the LDAP directory service, Data Archive updates the roles for the corresponding user account.

Data Archive only synchronizes role assignments from the LDAP directory service. The synchronization does not include roles or security groups.

If you change role assignments while a user is logged in to Data Archive, the changes are synchronized the next time the user logs in to Data Archive.

Nested Group Synchronization for Role Assignments

You can maintain roles in nested groups in the LDAP directory service. You may want to use nested groups for organizational purposes and to group similar types of roles. Attributes that you maintain at the group level apply to all roles within the group.

You may want to group roles by the type of roles that you commonly assign to users. For example, you may have several roles that you typically assign to users that perform administrator tasks. Then, you can assign users or groups of users to the group. By default, the user gets assigned to all of the roles that are within the group and any nested groups. Assigning roles to groups saves time as you avoid individually assigning roles to every user.

When you run the Sync with LDAP Server job, the job reviews all role assignments for users, including roles in nested groups. There is no limit to the amount of nested groups that you can have in one group.

For example, you can create the following structure in the LDAP directory service:

```
ILM Administrator Roles (Parent LDAP Security Group)
---Administrator role (ILM Role)
---Export Administrator role (ILM Role)
---ILM Tag Roles (Nested LDAP Security Group)
------Tag Administrator role (ILM Role)
------Tag Viewer role (ILM Role)
```

The ILM Administrator Roles group includes the Administrator role, the Export Administrator role, and the ILM Tag Roles nested group. The ILM Tag Roles group includes the Tag Administrator role and the Tag Viewer role.

You assign a user to the ILM Administrator Roles security group. The next time the user logs in to Data Archive, Data Archive adds the Administrator, Export Administrator, Tag Administrator, and Tag Viewer roles to the user account.

Sync with LDAP Server Job

The Sync with LDAP Server job synchronizes users between the LDAP directory service and Data Archive. Use the job to create users in Data Archive. Run the job when you initially set up LDAP authentication and after you create additional users in the LDAP directory service.

If you enable LDAP authentication, you must create and maintain users in the LDAP directory service and use the job to create user accounts in Data Archive. Run the job once for each group base that you want to synchronize.

When you run the job, the job uses the LDAP properties that are configured in the `conf.properties` file to connect to the LDAP directory service. If you specify the group base and the group filter in the job parameters, the job finds all of the users within the group and any nested groups. The job compares the users to users in Data Archive. If a user is in the group, but not in Data Archive, then the job creates a user account in Data Archive.

If you enabled role assignment synchronization, the job checks the security groups that the user is assigned to, including nested groups. The job matches the security group names to the names of the system-defined or Data Vault access role names. If the names are the same, the job adds the role to the user account in Data Archive. Data Archive automatically synchronizes any subsequent changes to security group assignments when users log in to Data Archive.

After the job creates users in Data Archive, any additional changes to users in the LDAP directory service are automatically synchronized when users log in to Data Archive. For example, if you change user properties, such as email addresses, or role assignments.

Sync with LDAP Server Job Parameters

When you configure the job parameters for the Sync with LDAP Server job, you specify how Data Archive synchronizes users from the LDAP directory service. You configure the connection properties to connect to the LDAP directory service and filter criteria to determine which users you want to synchronize.

The Sync with LDAP Server job includes the following parameters:

LDAP System

Type of LDAP directory service.

Use one of the following options:

- Active Directory
- Sun LDAP

Host of LDAP Server

The IP address or DNS name of the machine that hosts the LDAP directory service.

For example, `ldap.mycompany.com`.

Port of LDAP Server

The port on the machine where the LDAP directory service runs.

For example, `389`.

User

User that logs in to the LDAP directory service. You can use the administrator user. Or, you can use any user that has privileges to access and read all of the LDAP directories and privileges to complete basic filtering.

For example, `corpid@domain.com`.

Password

Password for the user.

Search Base

The search base where the LDAP definition starts before running the filter.

For example, `dc=mycompany,dc=com`

User Filter

A simple or complex filter that enables Data Archive to identify individual users in the LDAP security group.

For example, you might use one of the following filters:

- `objectClass=inetOrgPerson`
- `objectClass=Person`
- `objectClass=*` where `*` indicates that all entries in the LDAP security group should be treated as individual users.

Group Base

Optional. Sets the base entry in the LDAP tree where you can select which groups you want to use to filter users from the user filter.

If you do not specify a group base, then the job synchronizes all users in the LDAP directory service.

For example, `OU=Application Access,OU=Groups,DC=mycompany,DC=com`.

Group Filter

Optional. Determines which groups are selected. After the user filter returns the result set to the application, those users are compared to users in the selected groups only. Then, only true matches are added to Data Archive.

For example, `cn=ILM`.

Parameter Usage

Data Archive uses the following parameter combinations to fetch the list of users from the LDAP server:

Search Base and User Filter

or

Group Base and Group Filter

If you include values for both parameters, the Group Base and Group Filter combination takes precedence. If one of the parameters is empty, the Search Base and User Filter combination takes precedence.

Troubleshooting the Sync with LDAP Server Job

If you encounter the following error message during the sync with LDAP server job, perform the steps below to resolve the issue.

The sync with LDAP server job fails with the following error: **"Unable to find valid certification path to requested target."**

This issues occurs when a valid certificate from the LDAP server is not installed on the Data Archive server.

To resolve this issue, perform the following steps to install a valid certificate from the LDAP server to the Data Archive server:

1. Import the LDAP server certificate into the Java that you use for Data Archive.

   You must use the JDK bundled with the Data Archive installer.
2. Restart the Data Archive server and run a new sync with LDAP server job.

Setting Up LDAP User Authentication

Perform the set up activities when you initially set up LDAP user authentication and for any users that you create subsequent to the initial setup. After you set up LDAP user authentication, users can log in to Data Archive. Any changes for users are synchronized when users log in to Data Archive.

1. Set up the LDAP directory service.

   Create users in the LDAP directory service. If you enable role assignment synchronization, create security groups for Data Archive system-defined roles and Data Vault access roles. Then, assign the security groups to users.
2. Configure the LDAP user authentication properties in the `conf.properties` file.
3. Run the Sync with LDAP Server standalone job to synchronize users.
4. Assign roles to users if you did not enable role assignment synchronization.

LDAP User Authentication Maintenance

After you set up LDAP user authentication, use the LDAP directory service to manage subsequent changes, such as creating or deleting users or changing role assignments.

Use the following rules and guidelines for maintaining LDAP user authentication:

Creating users

When you create users in the LDAP directory service, run the Sync with LDAP Server job to synchronize users to Data Archive. If you enabled role assignment synchronization, then Data Archive automatically synchronizes the role assignments the next time users log in to Data Archive.

Deleting users

When you delete users in the LDAP directory service, users are not automatically deleted in Data Archive. However, users that are deleted in the LDAP directory service cannot log in to Data Archive. The role assignments are deleted from the user account the next time the user attempts to log in to Data Archive.

If you do not want a user to have authorization to log in to Data Archive, you can remove all of the roles from the user in the LDAP directory service or you can delete the user in the LDAP directory service.

Adding or removing role assignments

If you enabled role assignment synchronization, any changes that you make to role assignments in the LDAP directory service are automatically synchronized to the Data Archive user account when users log in to Data Archive. If you remove all role assignments for users, then users cannot log in to Data Archive.

Single Sign-On

Single sign-on is an authentication service that allows a user to use one set of credentials to access multiple applications or services. Single sign-on reduces the number of usernames and passwords across an enterprise. This makes identity management simpler and increases security standards.

You can use a single sign-on service to log into Data Archive. To use Data Archive's single sign-on feature, you must also use an identity provider that supports the SAML (security assertion markup language) standard. SAML consists of multiple components that together permit the exchange of identity, authentication, and authorization information between different services, in this case the identity provider and Data Archive.

To enable Data Archive's single sign-on feature, you must first create a KeyStore and an encryption certificate. Next you must configure the chosen identity provider for use with Data Archive. Last, you must configure Data Archive to enable single sign-on.

Step 1. Create the KeyStore and Encryption Certificate

The KeyStore and encryption certificate are used by both the identity provider and Data Archive. Run the following commands to generate the KeyStore and certificate.

1. Run the KeyStore command.

   ```
   " keytool -genkeypair -alias testkey01 -dname cn=INVR7ILM90.informatica.com -
   validity 3650 -keyalg RSA -keysize 1024 -keypass testkey01  -keystore c:/
   generatedKeys -storepass testkey01 "
   ```

   In this example:

   - "testkey01" is the KeyStore alias name.
   - "INVR7ILM90.informatica.com" is the Data Archive host machine.
   - "c:/generatedkeys" is the location where the KeyStore is generated with the name "generatedkeys." You can use any name and choose any file path.
   - "testkey01" is the KeyStore password.

2. Run the certificate command.

   ```
   " keytool -export -keystore c:/generatedKeys -alias testkey01 -file c:/cert.cer "
   ```

   In this example, "cert.cer" is the certificate name and can by anything.

   The command prompts you for a password.

3. Enter the KeyStore password (storepass) used in the KeyStore command.

Step 2. Configure the Identity Provider for Data Archive

Configure the chosen identity provider to host Data Archive. Refer to the documentation for the identity provider to configure the general settings for Data Archive.

1. In the identity provider's application settings for Data Archive, enter the following details:
   a. Single sign-on URL: http://<ilmhost:port>/sso.htm
   b. Entity ID: http://<ilmhost:port>/sso.htm
   c. Assertion encryption: Encrypted
2. Upload the encryption certificate (.cert) created in Step 1.
3. Download the identity provider metadata file and copy it to the machine where Data Archive is installed.

   The metadata file is typically available for download in the application settings.

4. Add users manually in the identity provider, or if you want to add users from an LDAP directory, integrate the LDAP directory with the identity provider.
5. Add the Data Archive application to the users that you want to grant Data Archive access to.

   Note: Data Archive does not support single sign-out, so you do not need to configure logout URL's.

Step 3. Configure Data Archive for Single Sign-On

To configure Data Archive for single sign-on, update the `conf.properties` file.

1. In the ILM installation directory, open the `conf.properties` file.
2. Locate the "#Properties to be provided for enabling Single sign on" section in the `conf.properties` file. Uncomment the statements below and enter the following details:
   a. Enable the property **informia.sso.enable = Y.**
   b. Provide the path of the identity provider metadata file, which you copied to the Data Archive machine from the identity provider in Step 2. Example: **informia.idp.metedata.file = c:\\metadata**
   c. Update the property **informia.key.path** with the KeyStore path created in Step 1. Example: **informia.key.path = c:\\generatedKeys**
   d. Update the property **informia.key.alias.name**, the alias name for the KeyStore. This is the name used when you created the KeyStore file in Step 1. Example: **informia.key.alias.name = testkey01**
   e. Update the property **informia.key.password**, the password for the KeyStore, which you set when you created the KeyStore. Before you update the property, this password must be encrypted using the encrypt password utility (`encryptPassword.bat` or `encryptPassword.sh`) provided in Data Archive. Run the utility using the commands below to encrypt the password from the ILM directory:

      `encryptPassword.bat` for Microsoft Windows

      `encryptPassword.sh` for Unix

      For example, if you created the KeyStore password "testkey01" in Step 1., run the following command to return the encrypted password: `encryptPassword.bat testkey01`

      Example of an encrypted password: **informia.key.password = D5YgPl814QpCtSgoYHbsCg==**
   f. Add the property "**informia.idp.home.url**" with the value as the identity provider home URL. This proprety is not required for all identity providers (Okta, Onelogin). Example: **informia.idp.home.url= https://desktop.pingone.com**
3. Save and close the `conf.properties` file.

   The following screenshot is an example of the single sign-on properties in the `conf.properties` file.

```
#Properties to be provided for enabling Single sign on.
informia.sso.enable = Y
#Path of meta-data file generated after configuring the application to IDP.
informia.idp.metedata.file = c:\\metadata
#Path of the key which will be used to encrypt the authRequest send to IDP.
informia.key.path = c:\\generatedKeys
#Alias name used to generate the key mentioned above.
informia.key.alias.name = amarkey01
#Password used to generate the key mentioned above in encrypted format.
informia.key.password = D5YgPl814QpCtSgoYHbsCg==
```

4. If you integrated an LDAP directory with the identity provider, you must run the sync with LDAP server standalone job in Data Archive to sync the users and roles to the ILM repository (AMHOME) before you restart the Data Archive server. When you run the LDAP sync job to sync the LDAP users to Data Archive,

your Data Archive user name will be same as the LDAP user name. To use single sign-on, you must configure that same user name between the identity provider and Data Archive, so that the user matches in AMHOME. For more information on the sync with LDAP server standalone job, see Chapter 4 of the *Informatica Data Archive User Guide*.

5. Restart the Data Archive server.

After you enable single sign-on, you can access Data Archive through either the identity provider or through the Data Archive environment URL. In both cases, Data Archive will open to the page appropriate for your user role. You will not be asked to log into Data Archive. If you are logged out of your identity provider or do not have an authenticated session, opening the Data Archive URL redirects you to the login page for your identity provider. However, if the identity provider session expires while you are still working in Data Archive, the Data Archive session will not automatically expire.

CHAPTER 16

Auditing

This chapter includes the following topics:

Auditing Overview

Use audit logs to monitor user actions in Data Archive. You can configure, archive, view, and purge audit logs. When you configure audit logs, you configure information such as audit level, retention period, and details of the Data Vault host to which you want to archive audit logs.

Audit Levels

The audit level you choose determines the user actions recorded in the audit logs.

The following table lists the audit levels and the corresponding user actions:

Audit Level	User Action
None	- Disable audit logs.
Audit Data Deletion	- Apply or remove legal holds. - Update expiration dates for archived records. - Delete and purge expired archived records.

Audit Level	User Action
Audit Metadata Modifications	- Add, edit, or remove tags from archived records. - View, create, edit, or delete Data Vault access roles. - Assign or delete Data Vault access role assignments to entities. - View or edit retention policies. - View, create, edit, or delete security groups. - View, create, or edit users. - Run a job. - Delete a data visualization report. - All Audit Data Deletion audit level actions.
Audit Data Access	- Log in to and out of Data Archive. - Open the application view or the technical view from the Data Vault search results. - Run a browse catalog search in Data Discovery. - Run a browse data search in Data Discovery. - Run a Data Vault search in Data Discovery. - Enter search criteria in Search Data Vault in Data Discovery. - View records and primary keys in the Search Data Vault search results. - View, create, or edit Data Discovery search options. - Create, edit, move, copy, or run a data visualization report. - All Audit Data Deletion audit level actions. - All Audit Metadata Modifications audit level actions.

Configuring Audit Logs

Define the type of user actions you want to record. The audit log stores all the recorded actions in the AM_AUDIT_HEADER and AM_AUDIT_DETAIL database tables. To archive audit logs to the Data Vault, define a Data Vault host.

1. Click **Administration** > **Audit Logs**.
2. Click **Settings**.

   The **Configuration** tab appears.
3. From the **Audit Level** list, select an audit level.
4. To retain audit logs for a certain time period, enter the **Retention** period in months or years.

   Note: When you choose indefinite retention, you do not need to schedule a purge job.
5. Enter a valid path for the **Staging Directory**.
6. Enter a valid path for the **Data Vault Data Directory**.
7. Enter the IP address of the **Data Vault Host**.
8. Enter the port number of the Data Vault host.
9. Enter the administration port number of the Data Vault host.
10. Enter the user name and password for the Data Vault host.
11. Enter the name of the **Data Vault Folder**.
12. Click **Save**.

Archiving Audit Logs

After you complete the configuration, schedule an audit log archive. When you schedule the audit log archive, Data Archive creates audit logs in the AM_AUDIT_HEADER and AM_AUDIT_DETAIL database tables. Then, Data Archive runs the Audit Log Archive Loader job to move the audit logs from the database tables to the Data Vault. After you schedule the audit log archive, you can view the audit log content.

Optionally, you can choose to receive email alerts when the archive job is complete, terminated, or exits with an error.

1. Click **Administration** > **Audit Logs**.
2. On the **Settings** tab, click **Archive Schedule**.
3. Select how often you want to run the archive job.
4. Optionally, choose **Completed**, **Terminated**, or **Error** notifications. Then, enter the email address where you want to receive alerts about the archive job.
5. Click **Schedule Archive** to schedule the audit log archive.

Viewing Audit Logs

You can view the user actions performed in Data Archive from the audit logs page.

Note: You must archive the audit logs before you can view the audit logs.

1. To verify that the audit log is enabled, perform any action and navigate to the audit log page.
2. On the **View** tab, choose the search criteria to filter the list of audit logs.

   The audit logs appear.

Exporting Audit Logs

When you view audit logs, you can export all the audit logs. When you export the logs, you are prompted to schedule a standalone job. Then, you can access the file when you monitor the job status. You can export the data to PDF, XML, or CSV.

1. Click **Administration** > **Audit Log**.

   The **Audit Logs** page appears.
2. Click **Export All**.

   The **Export Audit Log** page appears.
3. Select a format for the export file.
4. Click **Fetch Data**.

   The **Schedule Job** page appears.
5. Specify a time for the job to run, and click **Schedule**.

   The Export All Records standalone job is created.
6. Access **Jobs** > **Monitor Jobs.**

The **Monitor Jobs** screen appears if you schedule the job to run immediately.

7. Once the job completes, expand the job ID.
8. Expand the **Export All Records** job.
9. Click **View Exported Data**.

   The **File Download** dialog box appears.
10. Open or save the file.

Purging Audit Logs

To keep the audit logs from growing in size in the Data Vault, schedule a purge job. The purge job runs periodically to purge expired audit logs from the Data Vault. Optionally, you can choose to receive email alerts when the purge job is complete, terminated, or exits with an error.

1. In the **Settings** tab, click **Purge Schedule**.

   The **Purge Schedule** page appears.
2. From the list, select how often you want to run the purge job.
3. Optionally, choose **Completed**, **Terminated**, or **Error** notifications then enter the email address where you want to receive alerts about the purge job.

   The purge job deletes all the expired audit logs from the Data Vault.
4. Click **Schedule Purge** to save the settings.

CHAPTER 17

Running Jobs from External Applications

This chapter includes the following topics:

Running Jobs from an External Application Overview

You can use external applications to run standalone jobs without logging in to Data Archive. The external application, such as a third-party scheduler, uses a Java API to run the standalone jobs.

You can use triggers from a third-party scheduler or a command line program to start, resume, or terminate jobs in Data Archive.

When you use a third-party scheduler, you formulate URLs that call Data Archive JavaServer Pages (JSPs) and insert the URLs in the external application. The URLs include job-specific parameters. When the external

application runs the URL, the URL calls the Data Archive JSP. Data Archive uses the parameters from the formatted URL and immediately runs the job.

You can do the following tasks from outside the Data Archive:

- Start, resume, or terminate most standalone jobs through the Job Handler JSP (`JobHandler.jsp`).
- Update the retention policy for records in the Data Vault through the Run Update Retention JSP (`RunUpdateRetention.jsp`).
- Archive data to the Data Vault or history database through the Run Definition JSP (`RunDefinition.jsp`). If you archive data to the Data Vault and you require a staging schema, you must also call the Job Handler JSP to run the Data Vault Loader job.

Job Handler JSP

The Job Handler JSP, when called, immediately starts, resumes, or terminates a standalone job in Data Archive. Use a specially formatted URL to call the `JobHandler.jsp` from an external application to run a standalone job in Data Archive without logging in to the ILM application. The JSP uses the parameters that are specified in the URL to run the job.

If you use a third-party scheduler to schedule jobs in other applications, you can schedule and coordinate jobs across multiple applications at the same time. For example, you may want to run standalone jobs in Data Archive after you run monthly maintenance routines in other applications.

You can call `JobHandler.jsp` for the following standalone jobs:

- Archive Structured Digital Records
- Copy Application Version for Retirement
- Create Archive Folder
- Create Cycle Index
- Create Indexes
- Create Indexes on Data Vault
- Create Seamless Data Access
- Create Seamless Data Access Script
- Create Tables
- DB2 Bind Package
- DGA Data Collection
- Data Vault Loader
- Delete Indexes on Data Vault
- Load External Attachments
- Move External Attachments
- Purge Expired Records
- Restore External Attachments from Archive Folder
- Sync with LDAP Server
- Test Email Server Configuration
- Test JDBC Connectivity

Perform the following steps to start, resume, or terminate standalone jobs from external applications:

1. Configure security. Specify the machines that have access to call the JSP.
2. Form the URL that calls `JobHandler.jsp`. In the URL, specify the parameters for the job.
3. Add the URL to the external application code. When the external application calls the JSP, Data Archive uses the parameters in the URL to run the standalone job.

Step 1. Configure Security

You can run standalone jobs from an external application after you specify the IP address of the machine that hosts the external application in the `conf.properties` file.

In the `conf.properties` file, specify the machines that have access to call the JSP or API. By default, no machines have access.

1. Access `conf.properties` from the Web container folder of the Data Archive installation.
2. Configure the validHosts property.

   Enter the full IP address or the host name of the machines that can access the JSP or API from the external application. Use a comma to separate multiple values.

   The following text is an example of the property:

   ```
   validHosts=10.11.22.33, 192.168.20
   ```
3. Save the file.

Step 2. Form the URL that Calls JobHandler.jsp

To call `JobHandler.jsp`, form a URL that includes the required parameters. The parameters determine on which Data Archive instance the jobs runs, the job name, the job execution mode, and the required and optional job-specific parameters.

The URL that calls `JobHandler.jsp` has the following general syntax:

```
http://<host>:<port>/jsp/JobHandler.jsp?&action=<job action>&ProcessName=<process
name>&<parameter
 1 name>=<parameter 1 value>&<parameter 2 name>=<parameter 2 value>
```

The URL syntax includes general and job-specific parameters. The parameters are included in brackets. Some parameters are specific for all standalone jobs, such as the host, port, job action, and process name. The rest of the parameters depend on the standalone job. When you formulate the URL, you insert values for all of the parameters.

The URL syntax includes the following parameters:

Host

Host of the Data Archive instance on which you want to run the standalone job.

Port

Port of the Data Archive instance on which you want to run the standalone job.

Job Action

Execution mode for the standalone job. `JobHandler.jsp` supports the following values for the job action:

- START
- RESUME
- TERMINATE

Process Name

Technical name of the standalone job.

The following table lists the process names for each standalone job:

Standalone Job Name	Process Name
Archive Structured Digital Records	ARC_STRUCTURED_DIGITAL_RECS
Copy Application Version for Retirement	COPY_PFV
Create Archive Folder	CREATE_ARCHIVE_FOLDER
Create Cycle Index	CREATE_CYCLE_INDEX
Create Indexes	CREATE_INDEX
Create Indexes on Data Vault	CREATE_QUICKSEARCH_INDEX_ON_FAS
Create Seamless Data Access	SEAMLESS_DATA_ACCESS
Create Seamless Data Access Script	SEAMLESS_ACCESS_SCRIPT
Create Tables	CREATE_TABLE
Data Vault Loader	ARCHIVE_CRAWLER
DB2 Bind Package	DB2_BIND_PACKAGE
Delete Indexes on Data Vault	DELETE_QUICKSEARCH_INDEX_ON_FAS
DGA Data Collection	DGA_DATA_COLLECTION
Load External Attachments	LOAD_EXTERNAL_ATTACHMENTS
Move External Attachments	MOVE_EXTERNAL_ATTACHMENTS
Purge Expired Records	PURGE_EXPIRED_RECORDS
Restore External Attachments from Archive Folder	RESTORE_EXTERNAL_ATTACHMENTS
Sync with LDAP Server	LDAP_SCRIPT
Test Email Server Configuration	EMAIL_SERVER_TEST
Test JDBC Connectivity	JDBC_TEST

Parameter Name

Job-specific parameters.

Each standalone job requires program-specific parameters as input. Not all of the parameters are required. When you formulate the URL, you must include at least the required parameters.

For more information about the required and optional job parameters, see "Job Handler JSP Job Parameters" on page 253.

Parameter Value

Value for the job parameter name.

Forming the URL that Calls JobHandler.jsp

To call `JobHandler.jsp`, form a URL that includes the required parameters.

1. Use the sample syntax for the standalone job to form the URL base.

   The sample syntax includes the process name and all of the required and optional job parameters. For more information, see the specific job.

   Note: Except for the parameter values, the URL syntax is case-sensitive.
2. Replace the parameters with values.

   Enter values for all of the required parameters and for the optional parameters that you want to use. For example, replace the values for the host, port, job action, and parameter values.
3. Remove the optional parameters that you do not want to use.

JobHandler.jsp Example

You want to use `JobHandler.jsp` to start the Archive Structured Digital Records Job.

The following table lists the parameter values that you want to use for the job:

Parameter	Value
Host	10.12.12.12
Port	8080
Action	START
DIRECTORY	/u01/stage
MD_FILE	/u01/mtadata/structure.xml
FILE_ARCHIVE_REP_ID	2
PURGE_AFTER_LOAD	Y

Formulate the following URL to include the parameter values:

```
http://10.12.12.12:8080/jsp/JobHandler.jsp?
&action=START&ProcessName=ARC_STRUCTURED_DIGITAL_RECS&DIRECTORY=/u01/stage&MD_FILE=/u01/
mtadata/structure.xml&FILE_ARCHIVE_REP_ID=2&PURGE_AFTER_LOAD=Y
```

Step 3. Add the URL to the External Application Code

Add the URL to the external application code to call `JobHandler.jsp`. When the external application calls `JobHandler.jsp`, Data Archive uses the parameters in the URL to run the standalone job. Data Archive runs the job immediately. The output that the URL returns varies on the standalone job.

Job Handler JSP Job Parameters

Each standalone job requires program-specific parameters as input. Not all of the parameters are required. When you formulate the URL, you must include at least the required parameters.

Note: The JSP parameter name is the technical name of the parameter that corresponds to the user interface parameter name that you see if you run the standalone job in Data Archive. The user interface parameter name is provided for informational purposes only.

Archive Structured Digital Records Job

Use the Archive Structured Digital Records standalone job to archive structured digital records, such as call detail records. When you run the job, the job populates the AMHOME tables and creates an entity in EDM. You provide the entity creation details in a metadata file, such as the application name, application version, application module, and entity name. The job generates a metadata .xml file and calls the Data Vault Loader to load the files into the Data Vault. The Data Vault Loader loads the files from the root staging directory that is configured in the Data Vault target connection.

The date field in the digital record must be in the format `yyyy-MM-dd-HH.mm.ss.SSS`. For example: `2012-12-27-12.11.10.123`.

The following table describes the parameters for the Archive Structured Digital Records job:

JSP Parameter Name	User Interface Parameter Name	Required
MD_FILE	Metadata File	Yes
FILE_ARCHIVE_REP_ID	Target Archive Store	Yes
PURGE_AFTER_LOAD	Purge After Load	Yes Use one of the following values: - Y = Yes - N = No

Use the following syntax to formulate the URL that calls the `JobHandler.jsp`:

```
http://<host>:<port>/jsp/JobHandler.jsp?&action=<job
action>&ProcessName=ARC_STRUCTURED_DIGITAL_RECS&MD_FILE=<parameter
value>&FILE_ARCHIVE_REP_ID=<parameter value>&PURGE_AFTER_LOAD=<parameter value>
```

Copy Application Version for Retirement

When you run the Copy Application Version for Retirement job, the job creates a customer-defined application version. The job copies all of the metadata to the customer-defined application version. The job copies the table column and constraint information. After you run the job, customize the customer-defined application version in the Enterprise Data Manager. You can create entities or copy pre-packaged entities and modify the entities in the customer-defined application version. If you upgrade at a later time, the upgrade does not change customer-defined application versions.

The following table describes the parameters for the Copy Application Version for Retirement job:

JSP Parameter Name	User Interface Parameter Name	Required
PRODUCT_FAMILY_VERSION_ID	Product Family Version	Yes
NEW_PFV_NAME	New Product Family Version Name	Yes

Use the following syntax to formulate the URL that calls the `JobHandler.jsp`:

```
http://<host>:<port>/jsp/JobHandler.jsp?&action=<job
action>&ProcessName=COPY_PFV&PRODUCT_FAMILY_VERSION_ID=<parameter
value>&NEW_PFV_NAME=<parameter value>
```

Create Archive Folder

The Create Archive Folder job creates an archive folder for each retired application in the Data Vault. Run this job if you are retiring an application for the first time.

The following table describes the parameters for the Create Archive Folder job:

JSP Parameter Name	User Interface Parameter Name	Required
FILE_ARCHIVE_REP_ID	Destination Repository	Yes

Use the following syntax to formulate the URL that calls the `JobHandler.jsp`:

```
http://<host>:<port>/jsp/JobHandler.jsp?&action=<job
action>&ProcessName=CREATE_ARCHIVE_FOLDER&FILE_ARCHIVE_REP_ID=<parameter value>
```

Create Cycle Index

The Create Archive Cycle Index job creates database indexes based on job IDs for archive job executions. Run this job to optimize the restore performance. Perform the restore operation for a database archive after this job completes.

The following table describes the parameters for the Create Cycle Index job:

JSP Parameter Name	User Interface Parameter Name	Required
DEST_REP_ID	Destination Repository	Yes
ENTITY_ID	Entity ID	Yes
OUTPUT_LOCATION	Output Location of Generated Script File	No

Use the following syntax to formulate the URL that calls the `JobHandler.jsp`:

```
http://<host>:<port>/jsp/JobHandler.jsp?&action=<job
action>&ProcessName=CREATE_CYCLE_INDEX&DEST_REP_ID=<parameter
value>&ENTITY_ID=<parameter value>&OUTPUT_LOCATION=<parameter value>
```

Create Indexes Job

The Create Indexes job is a prerequisite to viewing history data when you upgrade to a newer version of PeopleSoft or an Oracle ERP application. Run the Create Indexes job to create indexes for new managed tables and to create indexes that did not exist in the original application version.

The following table describes the parameters for the Create Indexes job:

JSP Parameter Name	User Interface Parameter Name	Required
SRC_REP_ID	Source Repository	Yes
DEST_REP_ID	Destination Repository	No
OUTPUT_LOCATION	Output Location of Generated Script File	No

Use the following syntax to formulate the URL that calls the `JobHandler.jsp`:

```
http://<host>:<port>/jsp/JobHandler.jsp?&action=<job
action>&ProcessName=CREATE_INDEX&SRC_REP_ID=<parameter value>&DEST_REP_ID=<parameter
value>&OUTPUT_LOCATION=<parameter value>
```

Create Indexes on Data Vault

The Create Indexes on Data Vault job is a prerequisite to using Search Data Vault. Run the Create Indexes on Data Vault job to create a search index of columns for Search Data Vault. If you add or remove columns from the search index, run the Delete Indexes on Data Vault job first. Then run the Create Indexes on Data Vault job. This sequence of jobs deletes the existing search index and creates a new search index with the updated list of columns.

The following table describes the parameters for the Create Indexes on Data Vault job:

JSP Parameter Name	User Interface Parameter Name	Required
FILE_ARCHIVE_REP_ID	Destination Repository	Yes
ENTITY_ID	Entity	No
TABLE_ID	Table	No

Use the following syntax to formulate the URL that calls the `JobHandler.jsp`:

```
http://<host>:<port>/jsp/JobHandler.jsp?&action=<job
action>&ProcessName=CREATE_QUICKSEARCH_INDEX_ON_FAS&FILE_ARCHIVE_REP_ID=<parameter
value>&ENTITY_ID=<parameter value>&TABLE_ID=<parameter value>
```

Create Seamless Data Access Job

The Create Seamless Data Access job creates and runs a script that enables seamless data access to a database archive.

The following table describes the parameters for the Create Seamless Data Access job:

JSP Parameter Name	User Interface Parameter Name	Required
SRC_REP_ID	Source	Yes
DEST_REP_ID	Target	Yes
COMBINED_USERNAME	Combined Schema Name	No
COMBINED_PASSWORD	Combined Schema Password	No
QUERY_USERNAME	Archive Schema Name	No
QUERY_PASSWORD	Archive Schema Password	No
WHERE_IS_USER	Combined/Archive Schema Location	Yes
DBLINK	Database Link Name	Yes

Use the following syntax to formulate the URL that calls the `JobHandler.jsp`:

```
http://<host>:<port>/jsp/JobHandler.jsp?&action=<job
action>&ProcessName=SEAMLESS_DATA_ACCESS&SRC_REP_ID =<parameter
value>&DEST_REP_ID=<parameter value>&COMBINED_USERNAME=<parameter
value>&COMBINED_PASSWORD=<parameter value>&QUERY_USERNAME=<parameter
value>&QUERY_PASSWORD=<parameter value>&WHERE_IS_USER=<parameter
value>&DBLINK=<parameter value>
```

Create Seamless Data Access Script Job

The Create Seamless Data Access Script job creates a script that the database administrator can review and then run on the database to create objects necessary for Seamless Data Access. The script must be run manually which allows the administrator to validate the SQL statements that Data Archive will run. You must create a Seamless Data Access script if your IBM DB2 administrator does not allow external applications to create database objects.

The following table describes the parameters for the Create Seamless Data Access Script job:

JSP Parameter Name	User Interface Parameter Name	Required
SRC_REP_ID	Source Repository	Yes
DEST_REP_ID	Destination Repository	Yes
COMBINED_USERNAME	Combined Schema Name	No
COMBINED_PASSWORD	Combined Schema Password	No
QUERY_USERNAME	Query Schema Name	No
QUERY_PASSWORD	Query Schema Password	No
WHERE_IS_USER	Combined / Query Schema Location	Yes
DBLINK	Database Link	Yes

JSP Parameter Name	User Interface Parameter Name	Required
GENERATE_SCRIPT	Generate Script	No
SCRIPT_LOCATION	Script Location	No

Use the following syntax to formulate the URL that calls the `JobHandler.jsp`:

```
http://<host>:<port>/jsp/JobHandler.jsp?&action=<job action>
  &ProcessName=SEAMLESS_ACCESS_SCRIPT&SRC_REP_ID=<parameter value>
  &DEST_REP_ID=<parameter value>&COMBINED_USERNAME=<parameter value>
  &COMBINED_PASSWORD=<parameter value>&QUERY_USERNAME=<parameter value>
  &QUERY_PASSWORD=<parameter value>&WHERE_IS_USER=<parameter value>
  &GENERATE_SCRIPT=<parameter value>&DBLINK=<parameter value>&SCRIPT_LOCATION=<parameter
value>
```

Create Tables Job

The Create Tables Job generates a script file that you can use to manually create history and interim tables in a database. You will have to manually create history and interim tables if your IBM DB2 policies do not allow external applications to create objects in your IBM DB2 databases.

The following table describes the parameters for the Create Tables job:

JSP Parameter Name	User Interface Parameter Name	Required
SRC_REP_ID	Source Repository	Yes
OUTPUT_LOCATION	Output Location of Generated Script File	No
DEST_REP_ID	Destination Repository	No

Use the following syntax to formulate the URL that calls the `JobHandler.jsp`:

```
http://<host>:<port>/jsp/JobHandler.jsp?&action=<job
action>&ProcessName=CREATE_TABLE&SRC_REP_ID=<parameter value>&OUTPUT_LOCATION=<parameter
value>&DEST_REP_ID=<parameter value>
```

DB2 Bind Package Job

Use the DB2 Bind Package standalone job to bind packages. You must bind packages on the source before you can use DataDirect JDBC drivers to connect to IBM DB2 sources.

The following table describes the parameters for the IBM DB2 Bind Package job:

JSP Parameter Name	User Interface Parameter Name	Required
DB2_HOST	DB2 Host Address	Yes
DB2_PORT	DB2 Port Number	Yes
DB2_DATABASE	DB2 Location/Database Name	Yes

JSP Parameter Name	User Interface Parameter Name	Required
DB2_USER	User ID	Yes
DB2_PASSWORD	User Password	Yes

Use the following syntax to formulate the URL that calls the `JobHandler.jsp`:

```
http://<host>:<port>/jsp/JobHandler.jsp?&action=<job
action>&ProcessName=DB2_BIND_PACKAGE&DB2_HOST=<parameter value>&DB2_PORT=<parameter
value>&DB2_DATABASE=<parameter value>&DB2_USER=<parameter value>&DB2_PASSWORD=<parameter
value>
```

Delete Indexes on Data Vault

The Delete Indexes on Data Vault job deletes the search indexes that Search Data Vault uses. Run the Delete Indexes on Data Vault job when you add or remove columns to the search indexes. Then run the Create Indexes on Data Vault job to create the updated search indexes.

The Delete Indexes on Data Vault job does not have job parameters.

Use the following syntax to formulate the URL that calls the `JobHandler.jsp`:

```
http://<host>:<port>/jsp/JobHandler.jsp?&action=<job
action>&ProcessName=DELETE_QUICKSEARCH_INDEX_ON_FAS
```

DGA Data Collection Job

Run the DGA Data Collection Job to view information about a specific repository in the Dashboard. The job retrieves information about the selected repository including the graphical representation of Modules, Tablespaces, and Modules Trend.

The following table describes the parameters for the DGA Data Collection job:

JSP Parameter Name	User Interface Parameter Name	Required
SRC_REP_ID	Source Repository	Yes

Use the following syntax to formulate the URL that calls the `JobHandler.jsp`:

```
http://<host>:<port>/jsp/JobHandler.jsp?&action=<job
action>&ProcessName=DGA_DATA_COLLECTION&SRC_REP_ID=<parameter value>
```

Data Vault Loader Job

When you publish an archive project to archive data to the Data Vault, Data Archive moves the data to the staging directory during the Copy to Destination step. You must run the Data Vault Loader standalone job to complete the move to the Data Vault. The Data Vault Loader job executes the following tasks:

- Creates tables in the Data Vault.
- Registers the Data Vault tables.
- Copies data from the staging directory to the Data Vault files.
- Registers the Data Vault files.

- Collects row counts if you enabled the row count report for the Copy to Destination step in the archive project.
- Deletes the staging files.
- Generates a row count report and a row count summary report accessible under the Job ID on the **Monitor Jobs** page. The row count report lists each entity loaded to the Data Vault by archive job ID, along with the tables for each entity and row counts in the source database and the Data Vault archive folder. The report also lists any discrepancies. The row count summary report is a summary of the Data Vault Loader job, and lists each table loaded to the Data Vault along with their source row count, Data Vault archive folder row count, and any discrepancies.

The following table describes the parameters for the Data Vault Loader job:

JSP Parameter Name	User Interface Parameter Name	Required
ARCHIVE_JOB_ID	Archive Job Id	No

Use the following syntax to formulate the URL that calls the `JobHandler.jsp`:

```
http://<host>:<port>/jsp/JobHandler.jsp?&action=<job
action>&ProcessName=ARCHIVE_CRAWLER&ARCHIVE_JOB_ID=<parameter value>
```

Load External Attachments Job

Use the Load External Attachments standalone job to archive external attachments. When you run the job, the job reads the attachments from the directory you specify in the job parameters. The job archives records from the directory and all subdirectories within the directory. The job creates BCP files and loads the attachments to the AM_ATTACHMENTS table in the Data Vault.

The following table describes the parameters for the Load External Attachments job:

JSP Parameter Name	User Interface Parameter Name	Required
ATT_ENTITY_NAME	Attachment Entity Name	Yes
DIRECTORY	Directory	Yes
FILE_ARCHIVE_REP_ID	Target Archive Store	Yes
PURGE_AFTER_LOAD	Purge After Load	Yes Use one of the following values: - Y = Yes - N = No

Use the following syntax to formulate the URL that calls the `JobHandler.jsp`:

```
http://<host>:<port>/jsp/JobHandler.jsp?&action=<job
action>&ProcessName=LOAD_EXTERNAL_ATTACHMENTS&ATT_ENTITY_NAME=<parameter
value>&DIRECTORY=<parameter value>&FILE_ARCHIVE_REP_ID=<parameter
value>&PURGE_AFTER_LOAD=<parameter value>
```

Move External Attachments Job

The Move External Attachments job moves external attachments along with their records during an archive or restore job. Run the Move External Attachments job only if all the following conditions apply:

- You enabled **Move Attachments in Synchronous Mode** on the **Create or Edit an Archive Source** page.
- The entities support external attachments.

The following table describes the parameters for the Move External Attachments job:

JSP Parameter Name	User Interface Parameter Name	Required
ADD_ON_URL	Add on URL	No
ARCHIVE_JOB_ID	Archive Job ID	Yes

Use the following syntax to formulate the URL that calls the `JobHandler.jsp`:

```
http://<host>:<port>/jsp/JobHandler.jsp?&action=<job
action>&ProcessName=MOVE_EXTERNAL_ATTACHMENTS&ADD_ON_URL=<parameter
value>&ARCHIVE_JOB_ID=<parameter value>
```

Purge Expired Records Job

The Purge Expired Records job purges records that are eligible for purge from the Data Vault. A record is eligible for purge when the retention policy associated with the record has expired and the record is not under legal hold. Create and assign retention policies before you run the Purge Expired Records job.

Run the Purge Expired Records job to perform one of the following tasks:

- Generate the Retention Expiration report. The report shows the number of records that are eligible for purge in each table. When you schedule the purge expired records job, you can configure the job to generate the retention expiration report, but not purge the expired records.
- Generate the Retention Expiration report and purge the expired records. When you schedule the purge expired records job, you can configure the job to pause after the report is generated. You can review the expiration report. Then, you can resume the job to purge the eligible records.

To determine the number of rows in each table that are eligible for purge, generate the detailed or summary version of the Retention Expiration report. To generate the report, specify a past or current date for Data Archive to base the report on. Data Archive generates a list of records that are eligible on that date. You can pause the job to review the report and then run the job again to purge the records.

You can choose to generate one of the following types of reports:

Retention Expiration Detail Report

Lists the tables in the archive folder or, if you specified an entity, the entity. Shows the total number of rows in each table, the number of records with an expired retention policy, the number of records on legal hold, and the name of the legal hold group. The report lists tables by retention policy.

Retention Expiration Summary Report

Lists the tables in the archive folder or, if you specified an entity, the entity. Shows the total number of rows in each table, the number of records with an expired retention policy, the number of records on legal hold, and the name of the legal hold group. The report does not categorize the list by retention policy.

To purge records, you must enable the purge step through the `SKIP_PURGE_STEP` parameter.

Note: Before you purge records, use the **Search Within an Entity in Data Vault** search option to review the records that the job will purge. Records that have an expired retention policy and are not on legal hold are eligible for purge.

When you run the Purge Expired Records job to purge records, Data Archive reorganizes the database partitions in the Data Vault, exports the data that it retains from each partition, and rebuilds the partitions. Based on the number of records, this process can increase the amount of time it takes for the Purge Expired Records job to run. After the Purge Expired Records job completes, you can no longer access or restore the records that the job purged.

Note: If you purge records, the Purge Expired Records job requires staging space in the Data Vault that is equal to the size of the largest table in the archive folder, or, if you specified an entity, the entity.

The following table describes the parameters for the Purge Expired Records job:

JSP Parameter Name	User Interface Parameter Name	Required or Optional
TARGET_REP_ID	Archive Store	Required. Enter the name of the folder in the Data Vault that contains the records that you want to purge.
REPORT_TYPE	Report Type	Required. Enter a value based on one of the following options: - To generate the Retention Expiration Detail report, enter: `DETAIL` - If you do not want a report with a list of expired records, enter: `NONE` - To generate the Retention Expiration Summary report, enter: `SUMMARY`
PAUSE_AFTER_REPORT	Pause After Report	Required. Enter a value based on one of the following options: - To pause the job after Data Archive generates the report, enter: `Y` or `Yes` - To continue the job after Data Archive generates the report, enter: `N` or `No`
SKIP_PURGE_STEP	Purge Deleted Records	Required. Enter a value based on one of the following options: - To logically delete records from the Data Vault, enter: `N` or `No` - To logically and physically delete records from the Data Vault, enter: `Y` or `Yes`
EAMIL_ID	Email ID for Report	Optional. Enter the email address you want Data Archive to send the report to.

Use the following syntax to formulate the URL that calls the `JobHandler.jsp`.

To generate the Retention Expiration report:

Specify the name of the Data Vault archive folder and the type of report. Specify that the job pauses after it generates the report. Specify that the job skips the purge step. Optionally, enter an email address for Data Archive to send the report to. Use the following syntax for the URL:

```
http://<host>:<port>/jsp/JobHandler.jsp?
&action=&ProcessName=PURGE_EXPIRED_RECORDS&TARGET_REP_ID= <name of the connection to
the Data Vault>&REPORT_TYPE=<NONE, DETAIL,
orSUMMARY>&PAUSE_AFTER_REPORT=Y&Purge_Expiry_Date=<DD-MMM-
YYYY>&SKIP_PURGE_STEP=Y&EAMIL_ID=<email address>
```

To logically and physically delete eligible records from the Data Vault:

Specify the name of the Data Vault archive folder and the type of report. Specify that the job does not pause after it generates the report. Specify that the job performs the purge step. Optionally enter an email address for Data Archive to send the report to. Use the following syntax for the URL:

```
http://<host>:<port>/jsp/JobHandler.jsp?
&action=&ProcessName=PURGE_EXPIRED_RECORDS&TARGET_REP_ID= <name of the connection to
the Data Vault>&REPORT_TYPE=<NONE, DETAIL,
orSUMMARY>&PAUSE_AFTER_REPORT=Y&Purge_Expiry_Date=<DD-MMM-
YYYY>&SKIP_PURGE_STEP=N&EAMIL_ID=<email address>
```

Restore External Attachments from Archive Folder

The Restore External Attachments from Archive Folder job restores external attachments from the Data Vault.

The following table describes the parameters for the Restore External Attachments from Archive Folder job:

JSP Parameter Name	User Interface Parameter Name	Required
ARCHIVE_JOB_ID	Job Id of Archive Digital Records	Yes

Use the following syntax to formulate the URL that calls the `JobHandler.jsp`:

```
http://<host>:<port>/jsp/JobHandler.jsp?&action=<job
action>&ProcessName=RESTORE_EXTERNAL_ATTACHMENTS&ARCHIVE_JOB_ID=<parameter value>
```

Sync with LDAP Server Job

The Sync with LDAP Server job synchronizes users between the LDAP directory service and Data Archive. Use the job to create users in Data Archive. Run the job when you initially set up LDAP authentication and after you create additional users in the LDAP directory service.

The following table describes the parameters for the Sync with LDAP Server job:

JSP Parameter Name	User Interface Parameter Name	Required
LDAP_HOST	Host of LDAP server	Yes
LDAP_PORT	Port of LDAP server	Yes
LDAP_USERNAME	User	Yes
LDAP_PASSWORD	Password	Yes
LDAP_SEARCH_BASE	Search Base	Yes
LDAP_USER_FILTER	User Filter	Yes
LDAP_GROUP_BASE	Group Base	No
LDAP_GROUP_FILTER	Group Filter	No
LDAP_SYSTEM	LDAP System	Yes

Use the following syntax to formulate the URL that calls the `JobHandler.jsp`:

```
http://<host>:<port>/jsp/JobHandler.jsp?&action=<job
action>&ProcessName=LDAP_SCRIPT&LDAP_HOST=<parameter value>&LDAP_PORT=<parameter
value>&LDAP_USERNAME=<parameter value>&LDAP_PASSWORD=<parameter
value>&LDAP_SEARCH_BASE=<parameter value>&LDAP_USER_FILTER=<parameter
value>&LDAP_GROUP_BASE=<parameter value>&LDAP_GROUP_FILTER=<parameter
value>&LDAP_SYSTEM=<parameter value>
```

Test Email Server Configuration Job

Run the Test Email Server Configuration job to verify that Data Archive can connect to your mail server. Run the job after you define the mail server properties in the system profile. After you verify the connection, users can configure email notifications when they schedule jobs or projects

The following table describes the parameters for the Test Email Server Configuration job:

JSP Parameter Name	User Interface Parameter Name	Required
EMAIL_RECIPIENT	Email Recipient	Yes

Use the following syntax to formulate the URL that calls the `JobHandler.jsp`:

```
http://<host>:<port>/jsp/JobHandler.jsp?&action=<job
action>&ProcessName=EMAIL_SERVER_TEST&EMAIL_RECIPIENT=<parameter value>
```

Test JDBC Connectivity Job

The Test JDBC Connectivity job connects to a repository and maintains the connection for a specified time period. If your connection to a repository intermittently drops, run the Test JDBC Connectivity job to troubleshoot the connection.

The following table describes the parameters for the Test JDBC Connectivity job:

JSP Parameter Name	User Interface Parameter Name	Required
REP_ID	Repository Name	Yes
RETENTION_INTERVAL	Connection Retention Time (in minutes)	Yes

Use the following syntax to formulate the URL that calls the `JobHandler.jsp`:

```
http://<host>:<port>/jsp/JobHandler.jsp?&action=<job
action>&ProcessName=JDBC_TEST&REP_ID=<parameter value>&RETENTION_INTERVAL=<parameter
value>
```

Run Update Retention JSP

You can update retention policies in Data Archive based on triggers or events from external applications. For example, you can update the retention policy on archived records when you process a termination for an employee in an external application.

The external application, such as a third-party scheduler, uses a specially formatted URL to call the `RunUpdateRetention.jsp`. When called, the JSP immediately starts the Update Retention Policy standalone job in Data Archive. The JSP uses the parameters that are specified in the URL to run the job.

Perform the following steps to update retention policies from external applications:

1. Configure security. Specify the machines that have access to call the JSP.
2. Form the URL that calls `RunUpdateRetention.jsp`. In the URL, specify the parameters for the job.
3. Add the URL to the external application code. When the external application calls the JSP, Data Archive uses the parameters in the URL to run the Update Retention Policy standalone job.
4. Form the URL that calls `GetJobStatus.jsp`.
5. Add the URL to the external application code. When the external application calls the JSP, Data Archive uses the parameters in the URL to provide the status of the Update Retention Policy standalone job. Optionally, use the job status to handle an event in the external application.

Step 1. Configure Security

You can run standalone jobs from an external application after you specify the IP address of the machine that hosts the external application in the `conf.properties` file.

In the `conf.properties` file, specify the machines that have access to call the JSP or API. By default, no machines have access.

1. Access `conf.properties` from the Web container folder of the Data Archive installation.
2. Configure the validHosts property.

   Enter the full IP address or the host name of the machines that can access the JSP or API from the external application. Use a comma to separate multiple values.

   The following text is an example of the property:

   ```
   validHosts=10.11.22.33, 192.168.20
   ```
3. Save the file.

Step 2. Form the URL that Calls RunUpdateRetention.jsp

`RunUpdateRetention.jsp` schedules the Update Retention Policy job and returns the reference ID.

To call the JSP, form a URL that includes the required parameters. The parameters determine how Data Archive runs the job. For example, on which Data Archive instance the jobs runs, which records you want to update, and the new retention policy.

The URL that calls `RunUpdateRetention.jsp` has the following syntax:

```
http://<Data Archive Host>:<Data Archive Port>/jsp/RunUpdateRetention.jsp?
DEST_REP_NAME=<Destination ILM Repository Name>&ENTITY_ID=<Entity
ID>&WHERE_CLAUSE=<WHERE Clause>&EXISTING_POLICY_NAME=<Existing
Policy>&NEW_POLICY_NAME=<New Policy>&COMMENTS=<Comments>&USER_NAME=<User
Name>&IS_INCLUDE_REFERENCE_TABLES=Yes
```

The parameters are included in brackets. When you formulate the URL, you insert values for all of the parameters.

The following table describes the parameters to formulate the URL:

Parameter	Description
Host	Name of the Data Archive host.
Port	Name of the Data Archive port.
DEST_REP_NAME or DEST_REP_ID	Destination ILM repository name or repository ID.
ENTITY_ID	Entity ID that you want to modify the retention policy for.
WHERE_CLAUSE	Clause that identifies the records in the entity that you want to change the retention policy for.
EXISTING_POLICY_NAME	Name of the policy that you want to update.
NEW_POLICY_NAME	Name of the new policy that you want to assign to the records.
COMMENTS	Optionally, enter comments that you want to associate with the updated records.
USER_NAME or USER_ID	Data Archive user name or user ID. **Note:** Use a user name or user ID that has privileges to update the retention policy. For example, the user should have access to the entity that you want to update.
RELATIVE_TO_DATE	Optionally, if the retention policy does not have column level retention, select a relative date when you assign a different retention policy to an entity. The expiration date equals the relative date plus the retention period. For example, you select an entity and assign a retention policy that has a retention period of five years. You enter January 1, 2011, as the relative date. The new expiration date is January 1, 2016.
IS_REPORT_REQUIRED	Generates a retention modification report if the value is set to Yes. Default is No.
IS_PAUSE_AFTER_REPORT	Pauses the job after generating the retention modification report if the value is set to Yes. Default is No.
IS_INCLUDE_REFERENCE_TABLES	Applies or updates the retention policy on the reference tables. Valid inputs are Yes and No. Default is No.

The WHERE_CLAUSE parameter uses special formatting for conditions and operators. The following table describes the formatting for conditions or operators:

Condition or Operator	WHERE_CLAUSE Format
>=	###GTEQ###
<=	###LTEQ###
=	###EQ###

Condition or Operator	WHERE_CLAUSE Format
<>	###NEQ###
>	###GT###
<	###LT###
In	###IN###
Not In	###NOT_IN###
Is Null	###IS_NULL###
Is Not Null	###NOT_NULL###
Like	###LIKE###
Like (Starts with)	###STARTS_WITH###
Like (Ends with)	###ENDS_WITH###
Like (Contains)	###CONTAINS###
Not Like (Does not start with)	###NOT_START_WITH###
Not Like (Does not end with)	###NOT_END_WITH###
Not Like (Does not contain)	###NOT_CONTAIN###

Forming the URL that Calls RunUpdateRetention.jsp

To call `RunUpdateRetention.jsp`, form a URL that includes the required parameters.

1. Use the following format to form the URL:

```
http://<Data Archive Host>:<Data Archive Port>/jsp/RunUpdateRetention.jsp?
DEST_REP_NAME=<Destination ILM Repository Name>&ENTITY_ID=<Entity
ID>&WHERE_CLAUSE=<WHERE Clause>&EXISTING_POLICY_NAME=<Existing
Policy>&NEW_POLICY_NAME=<New Policy>&COMMENTS=<Comments>&USER_NAME=<User
Name>&IS_INCLUDE_REFERENCE_TABLES=Yes
```

2. Enter the parameter values and formulate the WHERE_CLAUSE.
   Use the following values in the WHERE_CLAUSE:
   - `Brace` or `NOBrace` to indicate the presence or absence of the left brace.
   - Table name and column name separate by a period. For example, `CUSTOMERS.CUSTOMER_ID`.
   - Operator
   - Column value or, if the operator is `IS NULL` or `NOT NULL`, use `NOValue`.
   - `Brace` or `NOBrace` to indicate the presence or absence of the right brace.
   - `AND`, `OR`, or `NOCondition` for condition.

   Each criteria should be separated by four hash (#) symbols. Each value in a criteria should be separated by three hash (#) symbols.

For example, use the following format to form the WHERE_CLAUSE if you want to set column-level retention based on company ID 1 and customer name John Smith:

```
WHERE_CLAUSE=(##CUSTOMERS.CUSTOMERS_ID###EQ###1###NOBrace###false###OR####NOBrace###C
USTOMERS.CUST_FIRST_NAME###EQ###JOHN###)###false###AND####(###CUSTOMERS.CUSTOMERS_ID#
##EQ###2###NOBrace###false###OR####NOBrace###CUSTOMERS.CUST_LAST_NAME###EQ###SMITH###
)###false###NOCondition)
```

3. Form the URL.
   Replace the following in the WHERE_CLAUSE:
   - Replace all hash (#) symbols with %23. For example, change `###EQ###` to `%23%23%23EQ%23%23%23`.
   - Replace all left parentheses with %28.
   - Replace all left parentheses with %29.

   Example:

```
http://10.74.40.37:8080/jsp/RunUpdateRetention.jsp?DEST_REP_NAME=Copy+of
+DATES_23_SAND&ENTITY_ID=-2&WHERE_CLAUSE=%28+%28%23%23%23CUSTOMERS.CUSTOMER_ID
%23%23%23EQ%23%23%2381077%23%23%23NOBrace%23%23%23false%23%23%23AND
%23%23%23%23NOBrace%23%23%23CUSTOMERS.CUSTOMER_ID%23%23%23EQ
%23%23%2381081%23%23%23%29%23%23%23false%23%23%23OR%23%23%23%23NOBrace
%23%23%23CUSTOMERS.CUSTOMER_ID%23%23%23EQ%23%23%2381089%23%23%23%29%23%23%23false
%23%23%23NOCondition&EXISTING_POLICY_NAME=5+Years&NEW_POLICY_NAME=10+Years&COMMENTS=T
est&USER_NAME=amadmin&IS_REPORT_REQUIRED=Yes&IS_PAUSE_AFTER_REPORT=Yes
```

Example

The following table lists the parameter values you use to form the URL that calls `RunUpdateRetention.jsp`:

Parameter	Value
Host	10.74.40.31
Port	5330
DEST_REP_NAME or DEST_REP_ID	Denali
ENTITY_ID	-1
WHERE_CLAUSE	COMPANY_TAB.APPLIMATION_JOB_ID = 104 and COMPANY_TAB.COMPANY_ID = 10
EXISTING_POLICY_NAME	Termination Policy
NEW_POLICY_NAME	25YPolicy
COMMENTS	New Termination Policy Date
USER_NAME or USER_ID	Test User

Formulate the following URL to include the parameter values:

```
http://10.74.40.31:5330/jsp/RunUpdateRetention.jsp?
DEST_REP_NAME=Denali&ENTITY_ID=-1&WHERE_CLAUSE=COMPANY_TAB.APPLIMATION_JOB_ID
%23%23%23EQ%23%23%23 104%23%23%23%23false%23%23%23%23NOCondition
AND COMPANY_TAB.COMPANY_ID %23%23%23EQ%23%23%23
10%23%23%23%23false%23%23%23%23AND&COMMENTS=New Termination Policy Date&USER_ID=Test
User&EXISTING_POLICY_NAME=Termination Policy&NEW_POLICY_NAME=25YPolicy
```

Step 3. Add the URL to the External Application Code

Add the URL to the external application code to call `RunUpdateRetention.jsp`. When the external application calls the JSP, Data Archive uses the parameters in the URL to run the Update Retention Policy standalone job. Data Archive runs the job immediately.

The Run Update Retention JSP returns a scheduler ID. The scheduler ID is a required parameter for the URL that calls `GetJobStatus.jsp`. If there is an error, the JSP returns one of the following errors:

- -2. The remote IP address or host name is not valid. You may get this error if the IP address or host is not configured in the `conf.properties` file.
- -1. Another exception occurred. For example, the WHERE_CLAUSE or the policy is not valid, or the Data Vault Service is not available.

Step 4. Form the URL that Calls GetJobStatus.jsp

The `GetJobStatus.jsp` gets the status of the Update Retention Policy job. To call the JSP, form a URL that includes the required parameters. The URL uses the scheduler ID value returned by the `RunUpdateRetention.jsp`.

The following table describes the parameters to formulate the URL:

Parameter	Description
Host	Name of the Data Archive host.
Port	Name of the Data Archive port.
Scheduler ID	Scheduler ID that `RunUpdateRetention.jsp` returns.

Forming the URL that Calls GetJobStatus.jsp

To call `GetJobStatus.jsp`, form a URL that includes the required parameters.

1. Use the following format to form the URL:

```
http://<Data Archive Host>:<Data Archive Port>/jsp/GetJobStatus.jsp?
schedulerID=<Scheduler ID>
```

2. Enter the parameter values in the URL.

Example

The following table lists the parameter values that you use to form the URL that calls `GetJobStatus.jsp`:

Parameter	Value
Host	10.74.40.31
Port	5330
Scheduler ID	65

Formulate the following URL to include the parameter values:

```
http://10.74.40.31:5330/jsp/GetJobStatus.jsp?schedulerID=65
```

Step 5. Add the URL to the External Application Code

Add the URL to the external application code to call `GetJobStatus.jsp`. Optionally, use the job status to handle an event in the external application.

The output of `GetJobStatus.jsp` shows the current status of the Update Retention Policy job. It provides the same status that you see if you monitor the job status in Data Archive. The output shows one of the following values:

- C. Completed
- R. Running
- T. Terminated

The status is not refreshed automatically. Call `GetJobStatus.jsp` to get the updated status.

Run Definition JSP

You can use an external application, such as a third-party scheduler, to archive data.

To run an archive job from an external application, you first form a URL. When you submit the URL through the external application, the URL calls the Run Definition JSP that runs the archive job.

You can archive data to either the Data Vault or database. Before you can use an external application to archive data, you must first create an archive project in Data Archive where you specify details about the data you want to archive. When you create the archive project, Data Archive assigns a definition ID to the archive project. Use the definition ID in the URL to specify the data you want to archive.

To view the status of the archive job, form a URL to call the Get Job Status JSP. To get a more detailed job status, form a URL that calls the handler that returns the detailed job status.

To run the archive job from an external application, complete the following steps:

1. Configure security. Specify the machines that have access to call the JSP.
2. Form the URL that calls `RunDefinition.jsp`. Specify the archive definition ID in the URL.
3. Add the URL to the external application code. When the external application calls the JSP, Data Archive uses the parameters in the URL to run the archive job.
4. Form a URL that returns the job status or detailed job status.
5. Add the URL to the external application code. When the external application calls the JSP or handler, Data Archive uses the parameters in the URL to provide the status of the archive job. Optionally, use the job status to handle an event in the external application.

Step 1. Configure Security

You can run standalone jobs from an external application after you specify the IP address of the machine that hosts the external application in the `conf.properties` file.

In the `conf.properties` file, specify the machines that have access to call the JSP or API. By default, no machines have access.

1. Access `conf.properties` from the Web container folder of the Data Archive installation.
2. Configure the validHosts property.

Enter the full IP address or the host name of the machines that can access the JSP or API from the external application. Use a comma to separate multiple values.

The following text is an example of the property:

```
validHosts=10.11.22.33, 192.168.20
```

3. Save the file.

Step 2. Form the URL that Calls RunDefinition.jsp

To call `RunDefinition.jsp`, form a URL that includes the definition ID of the archive project.

The URL that calls `RunDefinition.jsp` has the following syntax:

```
http://<Data Archive Host>:<Data Archive Port>/jsp/RunDefinition.jsp?
DefinitionID=<Definition ID>
```

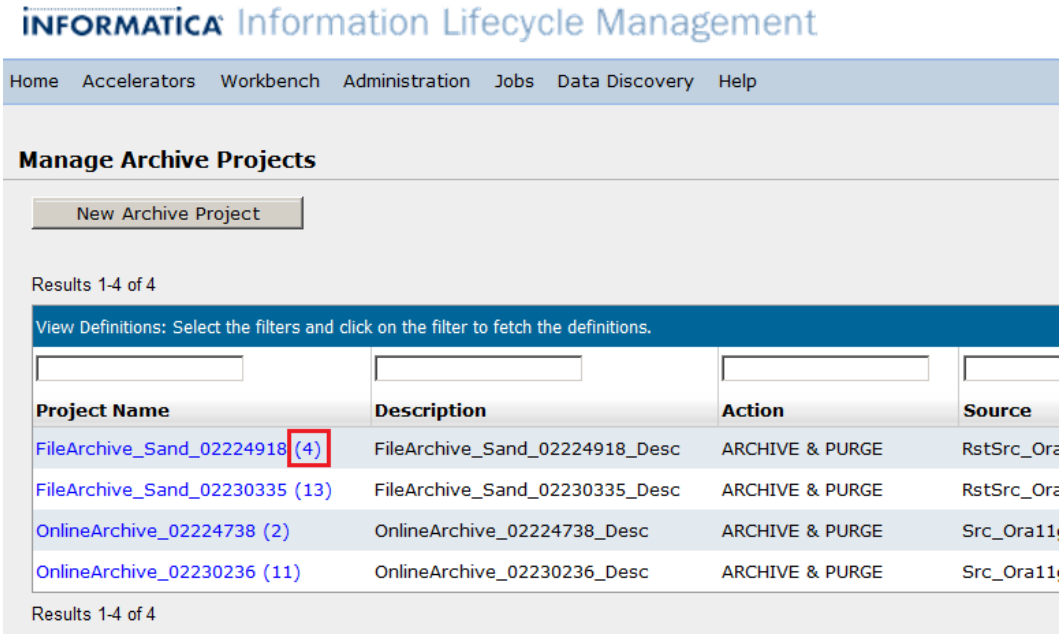
The parameters are included in brackets. When you formulate the URL, you insert values for all of the parameters.

The following table describes the parameters to formulate the URL:

Parameter	Description
Host	Name of the Data Archive host.
Port	Name of the Data Archive port.
Definition ID	The definition ID of the archive project.

Query the ILM repository to find the definition ID of an archive project. You can also find the definition ID of an archive project on the **Manage Archive Projects** page in Data Archive. To access the **Manage Archive Projects** page, click **Workbench** > **Manage Archive Projects**. The definition ID is the number inside the parenthesis located next to the project name.

The following image shows a list of archive projects and their definition IDs on the **Manage Archive Projects** page:



The following URL is an example of a URL that calls `RunDefinition.jsp`:

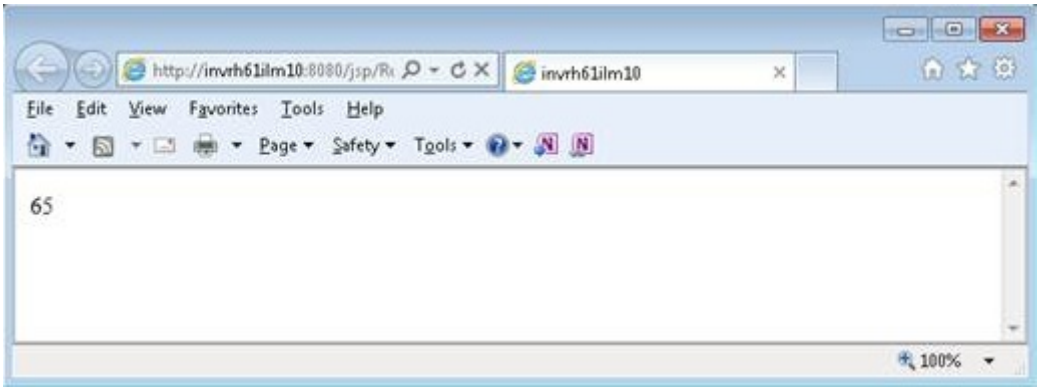
```
http://invr28ilm156:8080/jsp/RunDefinition.jsp?DefinitionID=4
```

Step 3. Add the URL to the External Application Code

Add the URL to the external application code to call `RunDefinition.jsp`. When the external application calls the JSP, Data Archive uses the definition ID in the URL to run the archive job immediately.

`RunDefinition.jsp` returns a scheduler ID. The scheduler ID is a required parameter for the URL that calls `GetJobStatus.jsp` and the URL that calls the handler that returns the detailed job status.

The following image shows the scheduler ID returned by `RunDefinition.jsp`:



If there is an error, `RunDefinition.jsp` returns one of the following errors:

- -2. The remote IP address or host name is not valid. You may get this error if the IP address or host is not configured in the `conf.properties` file.

- -1. Another exception occurred. For example, the definition ID is missing or incorrect.

Step 4. Form a URL that Returns the Job Status

To call `GetJobStatus.jsp` or the handler that returns the detailed job status, form a URL that includes the scheduler ID returned from `RunDefinition.jsp`.

Both `GetJobStatus.jsp` and the handler that returns the detailed job status retrieve the status of the archive job. `GetJobStatus.jsp` returns the status of the highest-priority job on the archive definition. If the archive project contains more than two entities, call the handler that returns the detailed job status to get the status of all jobs in the archive definition.

To call either job status, form a URL that includes the required parameters. The URLs use the scheduler ID returned from the `RunDefinition.jsp`.

The following table describes the parameters to formulate a URL:

Parameter	Description
Host	Name of the Data Archive host.
Port	Name of the Data Archive port.
Scheduler ID	Scheduler ID that `RunDefinition.jsp` returns.
Response Type	Provide one of the following output types: - XML - JSON If you provide any other value, the output type is linear. Applies to the detailed job status only.

Forming the URL that Returns the Job Status

To call `GetJobStatus.jsp` or the handler that returns the detailed job status, form a URL that includes the required parameters.

1. Use one of the following formats to form the URL:
   - To form the URL that calls `GetJobStatus.jsp`:

     ```
     http://<Data Archive Host>:<Data Archive Port>/jsp/GetJobStatus.jsp?
     schedulerID=<Scheduler ID>
     ```

   - To form the URL that calls the handler that returns the detailed job status:

     ```
     http://<Data Archive Host>:<Data Archive Port>/JobDetailStatus.htm?
     schedulerID=<Scheduler ID>&responseType=<Response Type>
     ```

2. Enter the parameter values in the URL.

Example

The following table lists the parameter values that you use to form the URLs that call `GetJobStatus.jsp` or the handler that returns the detailed job status:

Parameter	Value
Host	10.74.40.31
Port	5330
Scheduler ID	65
Response Type	xml

To call `GetJobStatus.jsp`, formulate the following URL to include the parameter values:

```
http://10.74.40.31:5330/jsp/GetJobStatus.jsp?schedulerID=65
```

To call the handler that returns the detailed job status, formulate the following URL to include the parameter values:

```
http://10.74.40.31:5330/jsp/GetJobStatus.jsp?schedulerID=65&responsetype=xml
```

Step 5. Add the URL to the External Application Code

Add a URL to the external application code to call `GetJobStatus.jsp` or the handler that returns the detailed job status. Optionally, use the job status to handle an event in the external application.

The output of `GetJobStatus.jsp` shows the current status of the archive job. The output of the handler that returns the detailed job status shows the status of all the archive jobs running. The outputs show one of the following values:

- C. Completed
- R. Running
- T. Terminated

The status is not refreshed automatically. You must call `GetJobStatus.jsp` or the handler again to get the updated status.

The following image displays C to indicate that the archive job is complete:

Legal Hold API

You can create legal hold groups, apply legal holds, or remove legal holds in Data Archive based on triggers or events from external applications.

The external application, such as a third-party scheduler, uses a specially formatted URL to call the legal hold API. When called, the API immediately starts the Apply or Remove Legal Hold job in Data Archive. The API uses the parameters that are specified in the URL to run the job.

Perform the following steps to create, apply, or remove legal holds from external applications:

1. Configure security. Specify the machines that have access to call the API.
2. Form the URL that calls the legal hold API. In the URL, specify the parameters for the job.
3. Add the URL to the external application code. When the external application calls the API, Data Archive uses the parameters in the URL to run the Apply Legal Hold or Remove Legal Hold job.
4. Form the URL that calls `GetJobStatus.jsp`.
5. Add the URL to the external application code. When the external application calls the API, Data Archive uses the parameters in the URL to provide the status of the legal hold job. Optionally, use the job status to handle an event in the external application.

To create, manage, or remove a legal hold, you must have one of the following system-defined roles:

- Administrator
- Legal Hold User

For more information on legal holds, see the *Informatica Data Archive User Guide*.

Step 1. Configure Security

You can run standalone jobs from an external application after you specify the IP address of the machine that hosts the external application in the `conf.properties` file.

In the `conf.properties` file, specify the machines that have access to call the JSP or API. By default, no machines have access.

1. Access `conf.properties` from the Web container folder of the Data Archive installation.
2. Configure the validHosts property.

   Enter the full IP address or the host name of the machines that can access the JSP or API from the external application. Use a comma to separate multiple values.

   The following text is an example of the property:

   ```
   validHosts=10.11.22.33, 192.168.20
   ```
3. Save the file.

Step 2. Form the Legal Hold URL

The legal hold API schedules the Add or Remove Legal Hold job and returns the reference ID.

To call the API, form a URL that includes the required parameters. The parameters determine how Data Archive runs the job. For example, the parameters determine the name of the legal hold group that you want to create, or the entity that you want to apply a legal hold to.

There are different URL formats for creating, applying, and removing legal holds.

Create Legal Hold Group URL Syntax

The URL that creates a legal hold group has the following syntax:

```
http://<Data Archive Host>:<Data Archive Port>/manageLegalHoldAPI.htm?
action=manageLegalHold&EVENT_TYPE=CREATE_LEGAL_HOLD&COMMENTS=<Comments>&LEGAL_HOLD_GROUP=
<Legal Hold Group>&USER_NAME=<User Name>
```

The parameters are included in brackets. When you formulate the URL, you insert values for all of the parameters and remove the brackets.

The following table describes the parameters to formulate the URL:

Parameter	Description
Host	Name of the Data Archive host.
Port	Name of the Data Archive port.
COMMENTS	Comments that you can use for audit purposes. For example, enter why you are creating the legal hold group. Optional.
LEGAL_HOLD_GROUP	Name for the legal hold group.
USER_NAME	Data Archive user name or user ID.

Forming the URL to Create a Legal Hold Group

To create a legal hold through the legal hold API, form a URL that includes the required parameters.

1. Use the following format to form the URL: `http://<Data Archive Host>:<Data Archive Port>/manageLegalHoldAPI.htm?action=manageLegalHold&EVENT_TYPE=CREATE_LEGAL_HOLD&COMMENTS=<Comments>&LEGAL_HOLD_GROUP=<Legal Hold Group>&USER_NAME=<User Name>`
2. Replace the text in brackets with the appropriate parameter values.

Example

The following URL creates a legal hold group named "LG1":

```
http://10.65.140.182:8013/manageLegalHoldAPI.htm?
action=manageLegalHold&EVENT_TYPE=CREATE_LEGAL_HOLD&COMMENTS=APITest&LEGAL_HOLD_GROUP=LG1
&USER_NAME=amadmin
```

Apply Legal Hold URL Syntax

The URL that applies a legal hold has the following syntax:

```
http://<Data Archive Host>:<Data Archive Port>/manageLegalHoldAPI.htm?
action=manageLegalHold&ENTITY_ID=<Entity
ID>&EVENT_TYPE=ADD_LEGAL_HOLD&COMMENTS=<Comments>&WHERE_CLAUSE=<WHERE
Clause>&TARGET_REP_ID=<Target Repository ID>&LEGAL_HOLD_GROUP=<Legal Hold
Group>&USER_NAME=<User Name>
```

The parameters are included in brackets. When you formulate the URL, you insert values for all of the parameters and remove the brackets.

The following table describes the parameters to formulate the URL:

Parameter	Description
Host	Name of the Data Archive host.
Port	Name of the Data Archive port.
ENTITY_ID	Entity that contains the records that you want to put on legal hold.
COMMENTS	Comments that you can use for audit purposes. For example, enter why you are creating the legal hold. Optional.
WHERE_CLAUSE	Clause that identifies the records that you want to put on legal hold.
TARGET_REP_ID	ID of the Data Vault target repository that contains the records that you want to put on legal hold.
LEGAL_HOLD_GROUP	Name of the legal hold group.
USER_NAME	Data Archive user name or user ID.

The WHERE_CLAUSE parameter uses special formatting for conditions and operators. The following table describes the formatting for conditions and operators:

Condition or Operator	WHERE_CLAUSE Format
>=	###GTEQ###
<=	###LTEQ###
=	###EQ###
<>	###NEQ###
>	###GT###
<	###LT###
In	###IN###
Not In	###NOT_IN###
Is Null	###IS_NULL###

Condition or Operator	WHERE_CLAUSE Format
Is Not Null	###NOT_NULL###
Like	###LIKE###
Like (Starts with)	###STARTS_WITH###
Like (Ends with)	###ENDS_WITH###
Like (Contains)	###CONTAINS###
Not Like (Does not start with)	###NOT_START_WITH###
Not Like (Does not end with)	###NOT_END_WITH###
Not Like (Does not contain)	###NOT_CONTAIN###

Forming the URL to Apply a Legal Hold

To apply a legal hold through the legal hold API, form a URL that includes the required parameters.

1. Use the following format to form the URL: `http://<Data Archive Host>:<Data Archive Port>/manageLegalHoldAPI.htm?action=manageLegalHold&ENTITY_ID=<Entity ID>&EVENT_TYPE=ADD_LEGAL_HOLD&COMMENTS=<Comments>&WHERE_CLAUSE=<WHERE Clause>&TARGET_REP_ID=<Target Repository ID>&LEGAL_HOLD_GROUP=<Legal Hold Group>&USER_NAME=<User Name>`
2. Replace the text in brackets with the appropriate parameter values.
3. Use the following values to create the WHERE clause that identifies the records that you want to place on legal hold:
   - `Brace` or `NOBrace` to indicate the presence or absence of the left brace.
     Use the `Brace` or `NOBrace` signifiers to indicate whether or not the search criteria contains a bracket.

     For example, the following type of search criteria does not require brackets: `First_Name EQ John and middle_name EQ rob OR last_name EQ smith`

     However, you can also create search criteria that contains brackets. For example: `(First_Name EQ John and middle_name EQ rob) OR last_name EQ smith`.

     Or, `First_Name EQ John and (middle_name EQ Rob OR last_name EQ smith)`

     The URL requires the use of `NOBrace` if the search criteria does not contain brackets.
   - Table name and column name separate by a period. For example, `CUSTOMERS.CUSTOMER_ID`.
   - Operator
   - Column value or, if the operator is `IS NULL` or `NOT NULL`, use `NOValue`.
   - `Brace` or `NOBrace` to indicate the presence or absence of the right brace.
   - `AND`, `OR`, or `NOCondition` for condition.

   Each criteria should be separated by four hash (#) symbols. Each value in a criteria should be separated by three hash (#) symbols.
4. Form the URL.

Replace the following in the WHERE clause:

- Replace all hash (#) symbols with %23. For example, change `###EQ###` to `%23%23%23EQ%23%23%23`.
- Replace all left parentheses with %28.
- Replace all left parentheses with %29.

Example

You want to apply a legal hold on a specific record in a table called LEVEL6. The table repository ID is "8" and the entity ID is "-11." The name of the legal hold group is "LG1."

For example:

COLUMN1	COLUMN2	COLUMN3
1	TEST1	VALUE1
2	TEST2	VALUE2
3	TEST3	VALUE3

The following URL places a legal hold on the second record in the table:

```
http://10.65.140.136:8016/manageLegalHoldAPI.htm?
action=manageLegalHold&ENTITY_ID=-11&EVENT_TYPE=ADD_LEGAL_HOLD&COMMENTS=APITest&WHERE_CLA
USE=NOBrace%23%23%23LEVEL6.COLUMN1%23%23%23IN%23%23%232%23%23%23NOBrace%23%23%23false
%23%23%23NOCondition&TARGET_REP_ID=8&LEGAL_HOLD_GROUP=LG1&USER_NAME=AMADMIN
```

In this example, the "false" parameter value is used to indicate that the search is not case sensitive. If you want the search to be case sensitive, use value "true."

Remove Legal Hold URL Syntax

The URL that removes a legal hold has the following syntax:

```
http://<Data Archive Host>:<Data Archive Port>/manageLegalHoldAPI.htm?
action=manageLegalHold&EVENT_TYPE=REMOVE_LEGAL_HOLD&COMMENTS=<Comments>&LEGAL_HOLD_GROUP=
<Legal Hold Group>&USER_NAME=<User Name>
```

The parameters are included in brackets. When you formulate the URL, you insert values for all of the parameters and remove the brackets.

The following table describes the parameters to formulate the URL:

Parameter	Description
Host	Name of the Data Archive host.
Port	Name of the Data Archive port.
COMMENTS	Comments that you can use for audit purposes. For example, enter why you are creating the legal hold. Optional.
LEGAL_HOLD_GROUP	Name of the legal hold group.
USER_NAME	Data Archive user name or user ID.

Forming the URL to Remove a Legal Hold

To remove a legal hold through the legal hold API, form a URL that includes the required parameters.

1. Use the following format to form the URL: `http://<Data Archive Host>:<Data Archive Port>/manageLegalHoldAPI.htm?action=manageLegalHold&EVENT_TYPE=REMOVE_LEGAL_HOLD&COMMENTS=<Comments>&LEGAL_HOLD_GROUP=<Legal Hold Group>&USER_NAME=<User Name>`
2. Replace the text in brackets with the appropriate parameter values.

Example

The following URL removes the legal hold group named "LG1":

```
http://10.65.140.182:8013/manageLegalHoldAPI.htm?
action=manageLegalHold&EVENT_TYPE=REMOVE_LEGAL_HOLD&COMMENTS=APITest&LEGAL_HOLD_GROUP=LG1
&USER_NAME=amadmin
```

Step 3. Add the URL to the External Application Code

Add the URL to the external application code to call the legal hold API. When the external application calls the API, Data Archive uses the parameters in the URL to run the Add or Remove Legal Hold job. Data Archive runs the job immediately.

The legal hold API returns a scheduler ID. The scheduler ID is a required parameter for the URL that calls `GetJobStatus.jsp`. If there is an error, the API returns one of the following errors:

- -2. The remote IP address or host name is not valid. You may get this error if the IP address or host is not configured in the `conf.properties` file.
- -1. Another exception occurred. For example, the WHERE clause is not valid, or the Data Vault Service is not available.

Step 4. Form the URL that Calls GetJobStatus.jsp

The `GetJobStatus.jsp` gets the status of the legal hold job. To call the JSP, form a URL that includes the required parameters. The URL uses the scheduler ID value returned by the legal hold API.

The following table describes the parameters to formulate the URL:

Parameter	Description
Host	Name of the Data Archive host.
Port	Name of the Data Archive port.
Scheduler ID	Scheduler ID that the legal hold API returns.

Forming the URL that Calls GetJobStatus.jsp

To call `GetJobStatus.jsp`, form a URL that includes the required parameters.

1. Use the following format to form the URL:

```
http://<Data Archive Host>:<Data Archive Port>/jsp/GetJobStatus.jsp?
schedulerID=<Scheduler ID>
```

2. Enter the parameter values in the URL.

Example

The following table lists the parameter values that you use to form the URL that calls `GetJobStatus.jsp`:

Parameter	Value
Host	10.74.40.31
Port	5330
Scheduler ID	65

Formulate the following URL to include the parameter values:

```
http://10.74.40.31:5330/jsp/GetJobStatus.jsp?schedulerID=65
```

Step 5. Add the URL to the External Application Code

Add the URL to the external application code to call `GetJobStatus.jsp`. Optionally, use the job status to handle an event in the external application.

The output of `GetJobStatus.jsp` shows the current status of the legal hold job. It provides the same status that you see if you monitor the job status in Data Archive. The output shows one of the following values:

- C. Completed
- E. Error
- P. Paused
- P. Pending
- R. Running
- T. Terminated
- W. Warning

The status is not refreshed automatically. Call `GetJobStatus.jsp` to get the updated status.

File Archive Transaction Restore API

You can run a transaction restore to restore data in the Data Vault back to the source database with the file archive transaction restore API.

An external application, such as a third-party scheduler, uses a specially formatted URL to call the file archive transaction restore API. When called, the API immediately runs the file archive transaction restore job in Data Archive. The API uses the parameters that are specified in the URL to run the job.

Perform the following steps to run the file archive transaction restore API:

1. Configure security. Specify the machines that have access to call the API and verify that the user has the required access role for the entity that you want to restore data from.
2. In the Data Archive user interface, create the file archive transaction restore definition.
3. Form the URL that calls the file archive transaction restore API. In the URL, specify the parameters for the job.
4. Add the URL to the external application code. When the external application calls the API, Data Archive uses the parameters in the URL to run the transaction restore job.

For more information on restoring data from the Data Vault, see the *Informatica Data Archive User Guide*.

Step 1. Configure Security

You can run a file archive transaction restore from an external application after you specify the IP address of the machine that hosts the external application in the `conf.properties` file. You must also verify that the user running the API has access to the entity that contains the transactions that you want to restore.

1. Access `conf.properties` from the Web container folder of the Data Archive installation.
2. Configure the validHosts property.

   Enter the full IP address or the host name of the machines that can access the API from the external application. Use a comma to separate multiple values.

   The following text is an example of the property:
   validHosts=10.11.22.33, 192.168.203.
3. Save the file.
4. To verify that the user running the API has access to the required entity, check the "Assigned Users" and "Associated Entities" reports from the **Manage Roles** page.
5. In the Data Archive user interface, click **Administration** > **Manage Roles**.
6. From the **Access Roles** tab, click the **View Users** link for an access role to verify the access role assigned to your login user.
7. Click the **View Entities** link for the user's access role and verify that the assigned role has been associated to the entity that contains the transaction that you want to restore.

Step 2. Create the File Archive Transaction Restore API Definition

Before you can run the API, you must create the definition for the file archive transaction restore in the Data Archive user interface.

1. Click **Workbench** > **Restore File Archive**.
2. Select the **Transaction Restore** tab.
3. Select the source Data Vault connection and the target database connection (the original source database) that you want to restore the transaction to.
4. Select the entity that contains the transaction or transactions that you want to restore.
5. Optionally, enter a maximum number of results records.
6. Enter the selection criteria to filter the desired transaction.
7. Click **View**.

   Data Archive searches for the transactions based on the selection criteria you entered. The results are displayed underneath the selection criteria.
8. Select the check box next to the transactions that you want to restore.
9. Click **Next**.

   The **Manage Restore Steps** page appears.
10. Select the actions you want the restore job to complete after each step in the restore process.

    To prevent the restore job from deleting the restored data in the database archive, select the **Skip** check box for the Delete from Archive step.
11. Click **Schedule**.

    This creates the required definition for the file archive transaction restore API. You can then begin to form the URL to call the API.

Step 3. Form the File Archive Transaction Restore URL

The file archive transaction restore API schedules the file archive transaction restore job and returns the schedule ID.

To call the API, form a URL that includes the required parameters. The parameters determine how Data Archive runs the job. For example, the parameters contain the WHERE clause that filters the records that you want to restore.

File Archive Transaction Restore URL Syntax

The URL that runs the file archive transaction restore has the following syntax:

http://<host>:<port>/API/fileArchiveTxnRestore.htm?definitionId=<definitionId>& whereClause=<whereClase>& username=<username>& password=<encryptedPassword>

The parameters are included in brackets. When you formulate the URL, you insert values for all of the parameters and remove the brackets.

The following table describes the parameters to form the URL:

Parameter	Description
definitionId	File archive transaction restore API definition ID. Required. After you create the definition in the user interface, you can obtain the definition ID by running the following command on the ILM repository database: `SELECT MAX(DEFN_ID) from AM_DEFN`
whereClause	Where clause to filter and restore the data. Required.
Username	Data Archive username. Required.
Password	Encrypted password returned by the encryptPassword.bat or encryptPassword.sh utility. Required when the property "informia.api.authentication" is enabled in the `conf.properties` file. Use this parameter only when "informia.api.authentication" is enabled in the `conf.properties` file. If "informia.api.authentication" is not enabled, you do not need to use this parameter in the URL. For more information, refer to the topic "API Authentication" in this chapter.

The whereClause parameter uses special formatting for conditions and operators. The following table describes the formatting for conditions and operators:

Condition or Operator	WHERE_CLAUSE Format
>=	###GTEQ###
<=	###LTEQ###
=	###EQ###
<>	###NEQ###
>	###GT###
<	###LT###
In	###IN###
Not In	###NOT_IN###
Is Null	###IS_NULL###
Is Not Null	###NOT_NULL###
Like	###LIKE###
Like (Starts with)	###STARTS_WITH###
Like (Ends with)	###ENDS_WITH###
Like (Contains)	###CONTAINS###
Not Like (Does not start with)	###NOT_START_WITH###

Condition or Operator	WHERE_CLAUSE Format
Not Like (Does not end with)	###NOT_END_WITH###
Not Like (Does not contain)	###NOT_CONTAIN###

Forming the URL for the File Archive Transaction Restore API

To restore a file archive transaction or transactions through the API, form a URL that contains the required parameters.

1. Use the following format to form the URL:

   ```
   http://<host>:<port>/API/fileArchiveTxnRestore.htm?definitionId=<definitionId>&
   whereClause=<whereClase>& username=<username>& password=<encryptedPassword>
   ```

2. Replace the text in brackets with the appropriate parameter values.
3. Use the following values to create the WHERE clause that identifies the records that you want to restore:
   - `Brace` or `NOBrace` to indicate the presence or absence of the left brace.
     Use the `Brace` or `NOBrace` signifiers to indicate whether or not the search criteria contains a bracket.

     For example, the following type of search criteria does not require brackets: `First_Name EQ John and middle_name EQ rob OR last_name EQ smith`

     However, you can also create search criteria that contains brackets. For example: `(First_Name EQ John and middle_name EQ rob) OR last_name EQ smith`.

     Or, `First_Name EQ John and (middle_name EQ Rob OR last_name EQ smith)`

     The URL requires the use of `NOBrace` if the search criteria does not contain brackets.
   - Table name and column name separate by a period. For example, `CUSTOMERS.CUSTOMER_ID`.
   - Operator
   - Column value or, if the operator is `IS NULL` or `NOT NULL`, use `NOValue`.
   - `Brace` or `NOBrace` to indicate the presence or absence of the right brace.
   - `AND`, `OR`, or `NOCondition` for condition.

   Each criteria should be separated by four hash (#) symbols. Each value in a criteria should be separated by three hash (#) symbols.
4. Form the URL.
   Replace the following in the WHERE clause:
   - Replace all hash (#) symbols with %23. For example, change `###EQ###` to `%23%23%23EQ%23%23%23`.
   - Replace all left parentheses with %28.
   - Replace all left parentheses with %29.

Example

You have created one file archive transaction restore definition on the entity "XYZ_001." The definition ID is 10. The entity driving table is "XYZ" and the table has 2 columns, "ID" (integer) and "Name"(varchar). The

WHERE clause is "ID<10." You want to use the API to create and schedule the job with the following job parameters:

Parameter	Value	Description
Host	10.74.0.233	Must be configured as a valid host in the `conf.properties` file.
Port	8080	Port number.
Definition ID	10	ID for the definition created in the Data Archive user interface. After you create the restore definition from the Data Archive user interface, run the following query to return the definition ID: SELECT MAX(DEFN_ID) from AM_DEFN
Username	User1	Data Archive username.
Password	adscsadnacasdad	Encrypted password returned by the encryptPassword.bat or encryptPassword.sh utilities.
WhereClause	ID >= 10	WHERE clause to filter records for the transaction restore.

The following URL runs the file archive transaction restore with the specified job parameters:

```
http://10.74.0.233:8080/API/fileArchiveTxnRestore.htm?
definitionId=10&username=User1&password=adscsadnacasdad&whereClause=%23%23%23XYZ.ID
%23%23%23GTEQ%23%23%2310%23%23%23NoBrace%23%23%23false%23%23%23NOCondition
```

Step 4. Add the URL to the External Application Code

Add the URL to the external application code to call the file archive transaction restore API. When the external application calls the API, Data Archive uses the parameters in the URL to run the restore job. Data Archive runs the job immediately.

The API returns either "-1" or a positive integer. "-1" indicates that an error has occurred and you can check the logs for more details. A positive integer is the scheduler ID.

You can use the scheduler ID to retrieve the job ID. Run the following command on the ILM repository to retrieve the job ID: SELECT JOB_ID FROM AM_JOBS where SCHEDULE_ID = <scheduler_ID>

On the **Monitor Jobs** page in Data Archive, file archive transaction restores run through the API are appended with "- API" in the job name.

API Authentication

You can enable API authentication to check if a password given in the URL is valid. To enable or disable API authentication, configure the "informia.api.authentication" parameter in the `conf.properties` file.

When you configure the parameter to enable API authentication, you must also provide an encrypted user password in the URL parameters. To generate the required encrypted password, use the encryptPassword.bat or encryptPassword.sh utility. The utilities are available in the Data Archive installation folder.

Password authentication is available only for API's built with the authentication mechanism. Refer to the individual API documentation in this chapter to determine if an API supports password authentication.

1. To enable or disable API authentication, open the `conf.properties` file from the Data Archive installation folder and configure the following parameter:

   `Informia.api.authentication`
   The default value of this parameter is N. Valid inputs are Y and N.
2. To generate the encrypted password required for the URL password parameter, run the encryptPassword.bat (on Microsoft Windows) or the encryptPassword.sh (on Linux and UNIX) utility from the command line. Input the password for the username running the API.

   The utility returns an encrypted password that you can copy and paste into the URL parameters.

CHAPTER 18

Salesforce Archiving Administrator Tasks

This chapter includes the following topics:

Salesforce Archiving Administrator Tasks Overview

Before you can archive data from Salesforce, you must configure permissions in Salesforce.

When you configure permissions in Salesforce, you configure them for certain profiles, users, and objects. For more information on how to configure permissions in Salesforce, see the Salesforce documentation.

For more information about the Salesforce archiving process, see the chapter "Salesforce Archiving" in the *Data Archive User Guide*.

Configure Salesforce Permissions

Before you can archive or purge data from Salesforce, you must configure permissions in Salesforce for certain profiles, users, and objects.

Configure the following permissions in Salesforce:

Object Type	Interim Schema	Level	Required Permissions/Configurations
Standard/Custom	PUBLIC/SFORCE	Profile	Enable "API Enabled"
		Profile	Enable "View Encrypted Data"
		User	Enable "Active"
		Object	Enable "Visible" field-level security for objects
		Object	Disable "Read-Only" field-level security for objects

Object Type	Interim Schema	Level	Required Permissions/Configurations
Standard	PUBLIC	Profile	Enable "Modify All Data"
		Object	Enable "Feed Tracking" for task and event object
		Object	Enable "Topics" for task and event
Custom	PUBLIC	Profile	Enable "View All" for custom object
		Profile	Enable "Modify All" for custom object
Standard/Custom	SFORCE	Profile	Enable "Customize Application"
		Profile	Enable "Manage Profiles and Permission Sets"
		Profile	Enable "Modify All Data"
		Profile	Enable "Create" on objects

CHAPTER 19

Upgrading Oracle History Data

This chapter includes the following topics:

Upgrading Oracle History Data Overview

You might need to upgrade Oracle ERP history, or archive, data when you upgrade an Oracle ERP application from version 11i to R12. Upgrade the history data so that you can view it in the upgraded version of the application.

For example, your organization archives data from Oracle E-Business Suite. The organization upgrades the version of Oracle E-Business Suite from version 11.5.10 to version R12. The data in the history database is compatible with version 11.5.10. You must be able to view the data in version R12.

Note: Contact Informatica Global Customer Service three months before you plan to upgrade to a newer version of Oracle E-Business Suite. Informatica Global Customer Service will send you the history upgrade scripts required for the version you are upgrading to.

Upgrading Oracle History Data Prerequisites

Before you upgrade Oracle ERP history data, verify the Data archive version, the accelerator version, the source connection, and the history database indexes. To increase performance of the database and the upgrade scripts, gather statistics for the history database.

Complete the following tasks:

1. Verify that you have Data Archive version 5.3 or later. You cannot use these instructions to upgrade the history data if you have an older version of Data Archive.
2. Verify that you have the most recent version of the accelerator. If you do not have the most recent version of the accelerator, contact Informatica Global Customer Support.
3. Verify the source connection. Do not change the source connection to reflect the updated version of the application. For example, if you upgrade from Oracle E-Business Suite 11.5.10 to R12, the application version for the source connection is Oracle Applications 11.5.10.
4. Verify that all indexes exist in the history database and that the indexes are valid. If indexes do not exist or are not valid, run the Create Indexes job.
5. Gather database statistics for the history database. Gathering database statistics optimizes the database and increases performance of the upgrade scripts.

Steps to Upgrade the History Data

You must perform multiple steps to upgrade the history data. When you complete the steps, you can view data that was archived through the original version of the application in the upgraded version of the application.

Perform each step in the order listed. Do not perform a step unless the previous step has completed successfully.

To upgrade the history data, perform the following steps:

1. Run the history upgrade scripts.
2. Run the Create History Table job with the source connection set to the original application version.
3. Update the application version for the source and target connections.
4. Run the Create History Table job with the source connection set to the upgraded application version.
5. Run seamless access with the source connection set to the upgraded application version.
6. Run the Create Indexes job.
7. Gather database statistics.

Step 1. Run the History Upgrade Scripts

Run the history upgrade scripts to upgrade key history tables with new columns and to populate data in new columns that cannot contain null values.

This step might create new tables and populate them with data. When you run the upgrade scripts, you can upgrade specific modules, or you can upgrade all supported modules in the history database.

History Upgrade Scripts Parameters

Each history upgrade script requires input parameters. Different scripts might require different parameters.

The following table describes the parameters that the history upgrade scripts use:

Parameter	Description
User name	Name of the history database user.
Password	Password for the history database user.
Database name	Name of the history database.
Tablespace name	Name of the tablespace.
Number of batches	Number of batches of rows that the script processes at the same time. For example, you specify two batches for a database with 1 million rows. The script processes two batches with 500,000 rows in each batch.
Degree of parallelism	Number of parallel processes that the SQL statements in the script can use
Modules to upgrade	Comma-separated list of codes that represent the Oracle modules to upgrade in the history database. For example, enter "GL" for General Ledger, "AP" for Accounts Payable, or "AR" for Accounts Receivable. For the list of supported modules, see the ILM Product Availability Matrix. When you enter the modules to upgrade, include all modules that were archived in the original application version.

Running the History Upgrade Scripts

To run the history upgrade scripts, copy the zip file to a directory on the history database server, unzip the file, create a link to the upgraded database, and run the scripts.

1. Copy AM_DATA_MIGRATION.zip to a directory on the history database server and unzip it.
2. Log in to the target history database as the Oracle owner.
3. Create a database link on the history database that connects to the upgraded version of Oracle E-Business Suite.

   Use the following syntax:

   ```
   create database link <link name> connect to <database> identified by <database password>
   using <E-Business TNS entry>;
   ```

4. Use SQL*Plus to connect to the history database as the administrator, and then run the following commands:

   ```
   Grant select on dba_extents to <history database schema name>;
   Grant select on dba_objects to <history database schema name>;
   Grant all on dbms_rowid to <history database schema name>;
   Exec dbms_stats.gather_table_stats('SYS','X$KTFBUE');
   Exit
   ```

5. Go to the directory where you extracted the zip file, and then run the following commands:

   ```
   chmod u+x install_common.sh
   chmod u+x alter.sh
   chmod u+x migrate.sh
   chmod u+x post_migration.sh
   ```

6. Run the install_common.sh script:

```
./install_common.sh <user name>/<password>@<database name>
```

For example:

```
./install_common.sh sarhist/amhistory@Ora11DB
```

7. Run the alter.sh script:

```
./alter.sh <user name>/<password>@<database name> <modules to upgrade>
```

For example:

```
./alter.sh sarhist/amhistory@Ora11DB GL,AP,AR,INV,WIP
```

8. Run the R12_table_indexs.sh script and include the tablespace name:

```
./R12_table_indexes.sh <user name>/<password>@<database name> <tablespace name> <modules to upgrade>
```

For example:

```
./R12_table_indexs.sh sarhist/amhistory@Ora11DB AM_HISTORY_INDEX GL,AP,AR,INV,WIP
```

9. Run the migrate.sh script:

```
./migrate.sh <user name>/<password>@<database name> <number of batches> <degree of parallelism> <modules to upgrade>
```

For example:

```
./migrate.sh sarhist/amhistory@Ora11DB 2 8 GL,AP,AR,INV,WIP
```

10. Run the post_migration.sh script:

```
./post_migration.sh <user name>/<password>@<database name> <modules to upgrade>
```

For example:

```
./post_migration.sh sarhist/amhistory@Ora11DB GL,AP,AR,INV,WIP
```

Step 2. Run the Create History Table Job with the Original Source

Run the Create History Table job with the source connection version set to the original application version.

The Create History Table job uses the metadata for the original application version. For example, if you upgrade Oracle E-Business Suite from version 11.5.10 to R12, the job uses the metadata for version 11.5.10.

The job performs the following actions:

- Adds new columns to existing tables in the history database.
- Processes tables that the upgrade scripts did not process.
- Adds columns that can contain null values to the history database tables.
- Verifies that tables not managed in the upgraded application version are updated properly.

This Create History Table job fails if it attempts to add a column that cannot hold null values to a table with data. If the job fails, run the history upgrade scripts for the module to which the table belongs.

Running the Create History Table Job

Run the Create History Table job through Data Archive.

1. Log in to Data Archive.
2. Select **Jobs** > **Schedule a Job**.
3. In the **Projects/Programs to Run** area select **Standalone Programs**, and click **Add Item**.

   The **Program** dialog box appears.
4. Select **Create Tables**, and click **Select**.
5. Select the source repository.
6. Select the target history database as the destination repository.
7. In the **Schedule** area, select the option to start the job immediately.
8. Click **Schedule**.

Step 3. Update the Application Version for Source and Target Connections

Update the source and target repository versions in the source and target connections to the upgraded application version.

Update the connections to ensure that Data Archive uses the correct accelerators when it archives and purges data, creates history tables and indexes, and creates seamless access.

Updating the Application Version for Source and Target Connections

Update the application version for source and target connections through Data Archive.

1. Log in to Data Archive.
2. Select **Administration** > **Manage Connections**.
3. Click **Edit** next to the source connection.
4. Select the upgraded application version from the **Application Version** list.
5. Click **Save**.
6. Repeat steps 3 through 5 for the target connection.

Step 4. Run the Create History Table Job with the Updated Source

Run the Create History Table job with the source connection set to the upgraded application version.

Running the Create History Table Job with the source connection set to the updated source creates new managed tables in the history database that the history upgrade scripts and the Create History Table job for the original application version did not create.

Running the Create History Table job with the updated source creates the following tables:

- Tables that are new to the application. For example, the AP_INVOICE_LINES_ALL table exists in Oracle E-Business Suite R12 but not in earlier versions of Oracle E-Business Suite.
- Tables that Data Archive supports in the upgraded application version but not the original application version. For example, the HXT_TIMECARDS_F table exists in Oracle E-Business Suite 11i and R12. Data Archive supports archiving and purging of this table for Oracle E-Business Suite R12.

Running the Create History Table Job

Run the Create History Table job through Data Archive.

1. Log in to Data Archive.
2. Select **Jobs** > **Schedule a Job**.
3. In the **Projects/Programs to Run** area select **Standalone Programs**, and click **Add Item**.

   The **Program** dialog box appears.
4. Select **Create Tables**, and click **Select**.
5. Select the source repository.
6. Select the target history database as the destination repository.
7. In the **Schedule** area, select the option to start the job immediately.
8. Click **Schedule**.

Step 5. Run the Seamless Data Access Job for the Updated Source

Run the Seamless Data Access job with the source connection version set to the upgraded application version.

Running the job with the updated source connection updates or recreates the seamless data access schemas for the updated application version.

Running the Seamless Data Access Job

Run the Seamless Data Access job through Data Archive.

1. Log in to Data Archive.
2. Select **Jobs** > **Schedule a Job**.
3. In the **Projects/Programs to Run** area select **Standalone Programs**, and click **Add Item**.

   The **Program** dialog box appears.
4. Select **Create Seamless Data Access**, and click **Select**.
5. Enter the job parameters.
6. In the **Schedule** area, select the option to start the job immediately.
7. Click **Schedule**.

Step 6. Run the Create Indexes Job

After you run the Seamless Data Access job, run the Create Indexes job.

Run the Create Indexes job to create indexes for new managed tables and to create indexes that did not exist in the original application version.

Running the Create Indexes Job

Run the Create Indexes job through Data Archive.

1. Log in to Data Archive.
2. Select **Jobs** > **Schedule a Job**.
3. In the **Projects/Programs to Run** area select **Standalone Programs**, and click **Add Item**.

   The **Program** dialog box appears.
4. Select **Create Indexes**, and click **Select**.
5. Select the source repository.
6. Select the target history database as the destination repository.
7. In the **Schedule** area, select the option to start the job immediately.
8. Click **Schedule**.

Step 7. Gather Database Statistics

Because the upgrade scripts make changes to the history database, gather database statistics for the upgraded history database. Gathering database statistics optimizes the database and increases performance of seamless access.

CHAPTER 20

Upgrading PeopleSoft History Data

This chapter includes the following topics:

Upgrading PeopleSoft History Data Overview

You might need to upgrade PeopleSoft history, or archive, data when you upgrade PeopleSoft from version 8.9 to 9.1. Upgrade history data so that you can view it in the upgraded version of the application.

For example, your organization archives data from PeopleSoft. The organization upgrades PeopleSoft from version 8.9 to version 9.1. The data in the history database is compatible with version 8.9. You must be able to view the data in version 9.1.

Note: Contact Informatica Global Customer Service three months before you plan to upgrade to a newer version of PeopleSoft. Informatica Global Customer Service will send you the history upgrade scripts required for the version you are upgrading to.

Upgrading PeopleSoft History Data Prerequisites

Before you upgrade PeopleSoft history data, verify the Data archive version, the accelerator version, the source connection, and the history database indexes. To increase performance of the database and the upgrade scripts, gather statistics for the history database.

Complete the following tasks:

1. Verify that you have Data Archive version 5.3 or later. You cannot use these instructions to upgrade the history data if you have an older version of Data Archive.
2. Verify that you have the most recent version of the accelerator. If you do not have the most recent version of the accelerator, contact Informatica Global Customer Support.
3. Verify that the PSAESTEPDEFN, PSAESTMTDEFN, and PS_UPG_DATACONV tables in the upgraded environment have records. You can use these tables for troubleshooting if you encounter problems during the PeopleSoft upgrade.
4. Verify the source connection. Do not change the source connection to reflect the updated version of the application. For example, if you upgrade from PeopleSoft 8.9 to 9.1, the application version for the source connection is PeopleSoft Applications 8.9.
5. Verify that all indexes exist in the history database and that the indexes are valid. If indexes do not exist or are not valid, run the Create Indexes job.
6. Gather database statistics for the history database. Gathering database statistics optimizes the database and increases performance of the upgrade scripts.

Steps to Upgrade the History Data

You must perform multiple steps to upgrade the history data. When you complete the steps, you can view data that was archived through the original version of the application in the upgraded version of the application.

Perform each step in the order listed. Do not perform a step unless the previous step has completed successfully.

To upgrade the history data, perform the following steps:

1. Run the history upgrade scripts.
2. Run the Create History Table job with the source connection set to the original application version.
3. Update the application version for the source and target connections.
4. Run the Create History Table job with the source connection set to the upgraded application version.
5. Run seamless access with the source connection set to the upgraded application version.
6. Run the Create Indexes job.
7. Gather database statistics.

Step 1. Run the History Upgrade Scripts

Run the history upgrade scripts to upgrade key history tables with new columns and to populate data in new columns that cannot contain null values. This step might create new tables and populate them with data.

The upgrade scripts are included in a zip file, for example, HRMS_HPY_UPGRADE_89To91.zip. If you customized any PeopleSoft module accelerator, you must update the "application changes" script to reflect the changes. For example, if you upgrade from PeopleSoft 8.9 to 9.1, the application changes script is HRMS_HPY_PS89_To_PS91_Upgrade_Appl_Changes.sql.

Running the History Upgrade Scripts

To run the history upgrade scripts, copy the zip file to a directory on the history database server, unzip the file, create a link to the upgraded database, and run the scripts.

1. Copy the upgrade scripts zip file to a directory on the history database server and unzip it.
2. If you customized any PeopleSoft module accelerator, update the application changes script to reflect the changes.

   Replace the interim table name in the script with the interim table name you use in the customized module, and update the script to reflect any change to the default columns or the business rule names.

   For example, you customized the PY - N.A. Payroll module. You changed the interim table name to "XA_1111_HPY1" and added column "ORG_NAME" with datatype "VARCHAR2(15 CHAR)." Replace all occurrences of "XA_10969_HPY1" with "XA_1111_HPY1." Add the ORG_NAME column to the CREATE TABLE statement.
3. Log in to the target history database as the owner.
4. Create a database link on the history database that connects to the upgraded version of the PeopleSoft application.

   Use the following syntax:

   ```
   create database link <link name> connect to <history database> identified by <database password> using <E-Business TNS entry>;
   ```
5. Go to the directory where you extracted the zip file, and then run the following command:

   ```
   chmod u+x migrate.sh
   ```
6. Run the migrate.sh script:

   ```
   ./migrate.sh <history database user name>/<database password>@<link name>
   ```

Original Application Changes Script Example

The application changes script creates the tables and application data changes in the history database schema to replicate the PeopleSoft upgrade process. To support restoring the old version of data with the new entities, this script also moves data from the old interim tables to the new interim tables.

If you upgrade from PeopleSoft 8.9 to 9.1, the application changes script is HRMS_HPY_PS89_To_PS91_Upgrade_Appl_Changes.sql.

This script contains the following commands:

```
if NOT migration_util.table_exists('XA_10969_HPY1) THEN

migration_util.execute_immediate('CREATE TABLE XA_10969_HPY1
(COMPANY VARCHAR2(3),
 PAYGROUP VARCHAR2(3),
 PAY_END_DT DATE,
 OFF_CYCLE VARCHAR2(1),
```

```
 PAGE_NUM NUMBER(4),
 LINE_NUM NUMBER(2),
 EMPLID VARCHAR2(11),
 PURGEABLE_FLAG CHAR(1),
 STATS_DATE DATE,
 ORG_ID VARCHAR2(5),
 ALL_PAGES_CONFIRMD NUMBER(31),
 ALL_EARNS_FINALIZD NUMBER(31),
 APPLIMATION_JOB_ID NUMBER(38) NOT NULL)');

end if;

if migration_util.table_exists('XA_10969_HPY1') THEN

migration_util.execute_immediate('INSERT INTO XA_10969_HPY1
(COMPANY, PAYGROUP, PAY_END_DT, OFF_CYCLE, PAGE_NUM, LINE_NUM, EMPLID, PURGEABLE_FLAG,
STATS_DATE, ORG_ID, ALL_PAGES_CONFIRMD, ALL_EARNS_FINALIZD, APPLIMATION_JOB_ID)
SELECT COMPANY, PAYGROUP, PAY_END_DT, OFF_CYCLE, PAGE_NUM, LINE_NUM, EMPLID,
PURGEABLE_FLAG, STATS_DATE, ORG_ID, ALL_PAGES_CONFIRMD, ALL_EARNS_FINALIZD,
APPLIMATION_JOB_ID FROM XA_797_HPY1');

end if;
```

Updated Application Changes Script Example

If you customized any PeopleSoft module accelerator, you must update the application changes script to reflect the changes.

For example, you upgraded from PeopleSoft 8.9 to 9.1. You customized the PY-N.A. Payroll module by changing the interim table name to "XA_1111_HPY1" and adding a column "ORG_NAME" with datatype "VARCHAR2(15 CHAR)."

Update the application changes script, HRMS_HPY_PS89_To_PS91_Upgrade_Appl_Changes.sql, as follows:

```
if NOT migration_util.table_exists('XA_1111_HPY1') THEN

migration_util.execute_immediate('CREATE TABLE XA_1111_HPY1
(COMPANY VARCHAR2(3),
 PAYGROUP VARCHAR2(3),
 PAY_END_DT DATE,
 OFF_CYCLE VARCHAR2(1),
 PAGE_NUM NUMBER(4),
 LINE_NUM NUMBER(2),
 EMPLID VARCHAR2(11),
 PURGEABLE_FLAG CHAR(1),
 STATS_DATE DATE,
 ORG_ID VARCHAR2(5),
 ALL_PAGES_CONFIRMD NUMBER(31),
 ALL_EARNS_FINALIZD NUMBER(31),
 ORG_NAME VARCHAR2(15),
 APPLIMATION_JOB_ID NUMBER(38) NOT NULL)');
end if;

if migration_util.table_exists('XA_1111_HPY1') THEN

migration_util.execute_immediate('INSERT INTO XA_1111_HPY1
(COMPANY, PAYGROUP, PAY_END_DT, OFF_CYCLE, PAGE_NUM, LINE_NUM, EMPLID, PURGEABLE_FLAG,
STATS_DATE, ORG_ID, ALL_PAGES_CONFIRMD, ALL_EARNS_FINALIZD, APPLIMATION_JOB_ID)
SELECT COMPANY, PAYGROUP, PAY_END_DT, OFF_CYCLE, PAGE_NUM, LINE_NUM, EMPLID,
PURGEABLE_FLAG, STATS_DATE, ORG_ID, ALL_PAGES_CONFIRMD, ALL_EARNS_FINALIZD,
APPLIMATION_JOB_ID FROM XA_797_HPY1');

end if;
```

Step 2. Run the Create History Table Job with the Original Source

Run the Create History Table job with the source connection version set to the original application version.

This step uses the metadata for the original application version. For example, if you upgrade PeopleSoft from version 8.9 to 9.1, this step uses the metadata for version 8.9.

This step performs the following actions:

- Adds new columns to existing tables in the history database.
- Processes tables that the upgrade scripts did not process.
- Adds columns that can contain null values to the history database tables.
- Verifies that tables not managed in the upgraded application version are updated properly.

This step fails if the Create History Table job attempts to add a column that cannot hold null values to a table with data. If this step fails, run the history upgrade scripts for the module to which the table belongs.

Running the Create History Table Job

Run the Create History Table job through Data Archive.

1. Log in to Data Archive.
2. Select **Jobs** > **Schedule a Job**.
3. In the **Projects/Programs to Run** area select **Standalone Programs**, and click **Add Item**.

   The **Program** dialog box appears.
4. Select **Create Tables**, and click **Select**.
5. Select the source repository.
6. Select the target history database as the destination repository.
7. In the **Schedule** area, select the option to start the job immediately.
8. Click **Schedule**.

Step 3. Update the Application Version for Source and Target Connections

Update the source and target repository versions in the source and target connections to the upgraded application version.

Update the connections to ensure that Data Archive uses the correct accelerators when it archives and purges data, creates history tables and indexes, and creates seamless access.

Updating the Application Version for Source and Target Connections

Update the application version for source and target connections through Data Archive.

1. Log in to Data Archive.

2. Select **Administration** > **Manage Connections**.
3. Click **Edit** next to the source connection.
4. Select the upgraded application version from the **Application Version** list.
5. Click **Save**.
6. Repeat steps 3 through 5 for the target connection.

Step 4. Run the Create History Table Job with the Updated Source

Run the Create History Table job with the source connection set to the upgraded application version.

Running the Create History Table Job with the source connection set to the updated source creates new managed tables in the history database that the history upgrade scripts and the Create History Table job for the original application version did not create.

Running the Create History Table job with the updated source creates the following tables:

- Tables that are new to the application. For example, the PS_IN_DEMAND table exists in PeopleSoft 9.1 but not in earlier versions of PeopleSoft.
- Tables that Data Archive supports in the upgraded application version but not the original application version. For example, the PS_MSR_HDR_INV table exists in PeopleSoft 8.9 and 9.1. Data Archive supports archiving and purging of this table for PeopleSoft 9.1.

Running the Create History Table Job

Run the Create History Table job through Data Archive.

1. Log in to Data Archive.
2. Select **Jobs** > **Schedule a Job**.
3. In the **Projects/Programs to Run** area select **Standalone Programs**, and click **Add Item**.

   The **Program** dialog box appears.
4. Select **Create Tables**, and click **Select**.
5. Select the source repository.
6. Select the target history database as the destination repository.
7. In the **Schedule** area, select the option to start the job immediately.
8. Click **Schedule**.

Step 5. Run the Seamless Data Access Job for the Updated Source

Run the Seamless Data Access job with the source connection version set to the upgraded application version.

Running the job with the updated source connection updates or recreates the seamless data access schemas for the updated application version.

Running the Seamless Data Access Job

Run the Seamless Data Access job through Data Archive.

1. Log in to Data Archive.
2. Select **Jobs** > **Schedule a Job**.
3. In the **Projects/Programs to Run** area select **Standalone Programs**, and click **Add Item**.

   The **Program** dialog box appears.
4. Select **Create Seamless Data Access**, and click **Select**.
5. Enter the job parameters.
6. In the **Schedule** area, select the option to start the job immediately.
7. Click **Schedule**.

Step 6. Run the Create Indexes Job

After you run the Seamless Data Access job, run the Create Indexes job.

Run the Create Indexes job to create indexes for new managed tables and to create indexes that did not exist in the original application version.

Running the Create Indexes Job

Run the Create Indexes job through Data Archive.

1. Log in to Data Archive.
2. Select **Jobs** > **Schedule a Job**.
3. In the **Projects/Programs to Run** area select **Standalone Programs**, and click **Add Item**.

   The **Program** dialog box appears.
4. Select **Create Indexes**, and click **Select**.
5. Select the source repository.
6. Select the target history database as the destination repository.
7. In the **Schedule** area, select the option to start the job immediately.
8. Click **Schedule**.

Step 7. Gather Database Statistics

Because the upgrade scripts make changes to the history database, gather database statistics for the upgraded history database. Gathering database statistics optimizes the database and increases performance of seamless access.

CHAPTER 21

Data Archive Maintenance

This chapter includes the following topics:

Data Archive Maintenance Overview

To ensure a smooth recovery of Data Archive components after an unexpected server shut down, periodically backup and maintain Data Archive components.

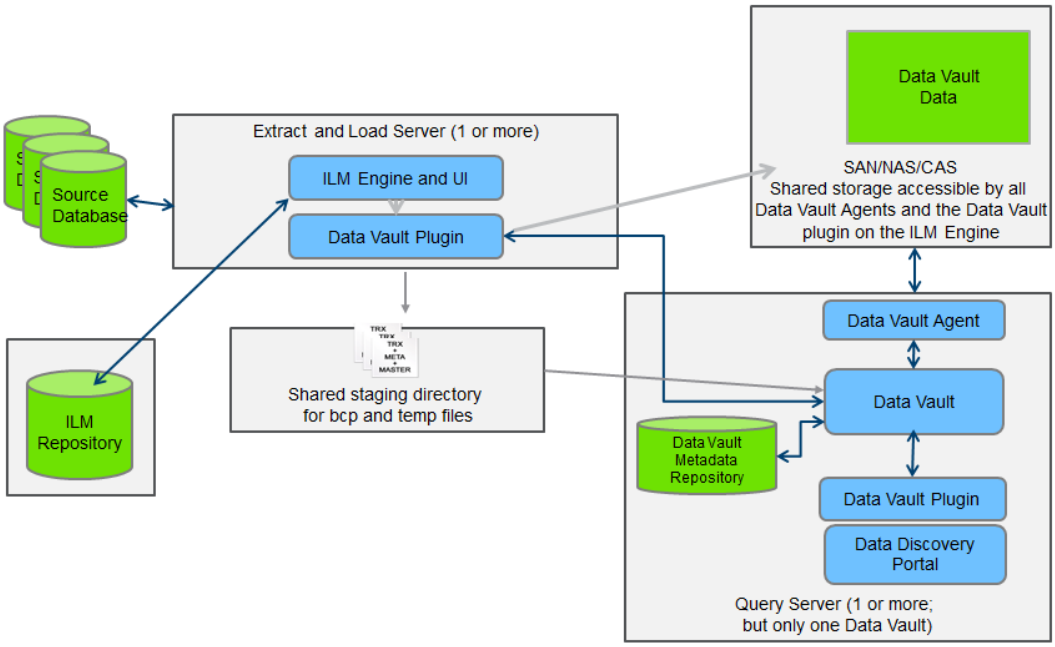
Back up the following Data Archive components:

- ILM repository
- Data archived in Data Vault
- Directory containing the search indexes used by Search Data Vault

Maintain the Data Vault repository once a month.

Data Archive Component Integration

The following diagram shows how Data Archive components integrate with each other in a typical data center:



Backing up Data Archive Components

Back up the ILM repository, data in the Data Vault, and the directories that contain the search indexes. Back up the Data Archive components at the same time. Follow the same backup schedule as your production environment.

ILM Repository

Back up the ILM repository just as you would a standard database. Incrementally back up the ILM repository schema.

Data in the Data Vault

Use your standard file system backup process to back up the data in the Data Vault. Incrementally back up the data files and metadata repository database.

Directory Containing the Search Indexes

Back up the directory containing the search indexes used by Search Data Vault. The location of the directory containing the search indexes is configured in the `conf.properties` file.

1. Open the `conf.properties` file.

   The `conf.properties` file is located in the root Data Archive installation directory.
2. Locate the informia.keywordSearchIndexDir property.

   The property specifies the file path to the search-indexes directory.

Maintaining the Data Vault Repository

Complete maintenance tasks on the Data Vault repository on a monthly basis. Schedule three to four hours to complete the maintenance tasks.

1. Stop the Data Vault components.
2. Create a copy of the current metadata database file.

   Example: `$SSA_INI_DIR/meta.n00 $SSA_INI_DIR/meta.n00.current_backup`
3. Manually start the Data Vault repository server.
   a. Enter the following command: `ssaeng meta meta &`

      The following message appears: Accepting connection
   b. Enter the following command: `ssasql meta meta DBA/dba`

      The command prompt appears.
   c. Enter the following commands:

      ```
      alter session set dismountgc=4;
      shutdown;
      .exit
      ```

      The garbage collection routine starts.
4. Optionally, to compact the repository, enter the following command: `ssadbchk -p meta`

   Performance remains the same whether you compact the database or not. Based on the internal database structure, the size of the database file size might decrease.

   The following success message appears:

   ```
   Warnings: 0 errors: 0
   ```

   If you do not see the success message, restore the metadata database file that you created. Start Data Vault and send the error details to Informatica Global Customer Support.
5. Start the Data Vault components.
6. Verify that the Data Archive components start up.

APPENDIX A

Datetime and Numeric Formatting

This appendix includes the following topics:

Datetime and Numeric Formatting Overview

You can specify the default search and display formats for datetime values in data discovery searches and results.

You can also specify the data format for numeric and datetime columns when you define search options for an entity in Data Discovery. The data format determines how Data Archive displays the data in Data Vault search and browse data search results.

Enter datetime formats through a format string that contains one or more alphabetic characters. Data Archive interprets alphabetic characters in a format string as the components of a date or time value unless you enclose them within single quotation marks. Data Archive interprets characters that appear within single quotation marks as literal strings. It also interprets nonalphabetic characters as literal strings. For example, Data Archive interprets the letter `E` in a format string as the day of the week, but it interprets `'E'` as the letter "E."

Enter numeric formats through a format string that contains one or more reserved characters. Data Archive interprets the reserved characters as the components of a numeric value unless you enclose them within single quotation marks. Data Archive interprets characters that appear within single quotation marks and characters that are not reserved characters as literal strings. For example, Data Archive interprets `#` as one digit, but it interprets `'#'` as the pound sign character. If you do not enter a numeric format string, Data Archive displays the number as it appears in the database.

Datetime Format Strings

Enter datetime formats through a format string that contains one or more alphabetic characters. The number of characters in the string determines how Data Archive interprets the datetime value.

The following table describes the datetime format characters that Data Archive uses:

Character	Description
a	Time period for a 12-hour clock as text. For example, "PM."
d	Day in month as a number. Range is 1 through 31.
D	Day in year as a number. Range is 1 through 366.
E	Day of the week as text. For example, "Tuesday" or "Tue."
f	Fractional seconds as a number. Display fractional seconds as they are stored in the database.
F	Day of the week in the month as a number. Range is 1 through 5. For example, "3" for the third Tuesday of the month.
G	Era as text. For example, "AD" or "BC."
h	Hour in day for a 12-hour clock as a number. Range is 1 through 12.
H	Hour in day for a 24-hour clock as a number. Range is 0 through 23.
k	Hour in day for a 24-hour clock as a number. Range is 1 through 24.
K	Hour in day for a 12-hour clock as a number. Range is 0 through 11.
m	Minute in hour as a number. Range is 0 through 59.
M	Month in year as a number or text. If you enter one or two characters, Data Archive interprets the month as a number. Range is 1 through 12. For example, "7" or "07." If you enter three or more characters, Data Archive interprets the month as text. Range is January through December. For example, "Jul" or "July."
s	Second in minute as a number. Range is 0 through 59.
S	Fractional seconds as a number. Display trailing zeros. Range is 0 through 999,999. For example, "SSS" for milliseconds.
w	Week in year as a number. Range is 1 through 52.
W	Week in month as a number. Range is 1 through 5.
y	Year. For example, "1995" or "95."

Character	Description
z	General time zone. For example, "Pacific Standard Time," "PST," or "GMT-08:00." For time zones represented by a GMT offset value, the range for hours is 0 - 23, and the range for minutes is 00 through 59.
Z	RFC 822 time zone. For example, "-0800." The range for hours is 00 - 23, and the range for minutes is 00 through 59.

The following table describes how Data Archive interprets the number of characters in display formats and search formats for text, numeric, year, and time zone values:

Datetime Value Type	Display Formats	Search Formats
Text values	To display the abbreviated form of a datetime value, enter three or fewer characters. To display the full form of a datetime value, enter four or more characters.	Data Archive accepts both the abbreviated form and the full form of datetime values, no matter how many characters that you enter.
Numeric values	The number of characters that you enter determines the number of digits that Data Archive displays. If you enter extra digits, Data Archive pads the result with leading zeros. For example, if the display format is `DDD` and the day of the year is January 1, Data Archive displays the day as 001.	Data Archive ignores the number of characters.
Year	To display a two-digit year, enter two characters. To display a four-digit year, enter four characters.	Data Archive interprets two characters as a two-digit year. To determine the century, Data Archive adjusts the date to be within 80 years before and 20 years after the current year. For example, it is 2011, and the search format is `mm/dd/yy`. If you enter 10/31/65, Data Archive interprets the year as 1965. If you enter 10/31/05, Data Archive interprets the year as 2005. Data Archive interprets one character and three or more characters literally. For example, the search format is `mm/dd/yyyy`. If you enter 12/25/11, Data Archive interprets the date as December 25, 11 AD. If you enter 12/25/-3, Data Archive interprets the date as December 25, 4 BC.
General time zone	Data Archive displays the time zone name. If the time zone does not have a name, Data Archive uses a GMT offset value with the following syntax: `GMT +/- HH:mm`	Data Archive accepts the general time zone or the RFC 822 time zone.
RFC 822 time zone	Data Archive uses the following syntax: `+/- HH:mm`	Data Archive accepts the general time zone or the RFC 822 time zone.

Datetime Format String Examples

Data Archive interprets datetime values differently based on the datetime format string.

The following table shows how Data Archive interprets the datetime value "July 4, 2001 12:08:56 Pacific Daylight Time" when you enter different datetime format strings:

Datetime Format String	Result
`yy.MM.dd G 'at' HH:mm:ss z`	2001.07.04 AD at 12:08:56 PDT
`EEE, MMM d, ''yy`	Wed, Jul 4, '01
`h:mm a`	12:08 PM
`hh 'o''clock' a, zzzz`	12 o'clock PM, Pacific Daylight Time
`K:mm a, z`	0:08 PM, PDT
`yyyyy.MMMMM.dd GGG hh:mm aaa`	02011.July.04 AD 12:08 PM
`EEE, d MMM yyyy HH:mm:ss Z`	Wed, 4 Jul 2001 12:08:56 -0700
`yyMMddHHmmssZ`	010704120856-0700

Numeric Format Strings

Enter numeric formats through a format string that contains one or more characters. The number of characters in the string determines how Data Archive displays the numeric value.

When the number of digits in the Data Vault target or target database exceeds the number of digits in the format string, Data Archive uses half even rounding. For example, if the value in the target database is 12.345, and the format string is `###.00`, Data Archive displays 12.35.

The following table describes the numeric format characters that Data Archive uses:

Character	Description
0	Digit. Display leading and trailing zeros. For example, if the value in the target database is 1.2, and the format string is `000.00`, Data Archive displays 001.20.
#	Digit. Do not display leading or trailing zeros. For example, if the value in the target database is 1.2, and the format string is `###.##`, Data Archive displays 1.2.
.	Decimal separator or monetary decimal separator.
-	Negative sign.
,	Grouping separator. In most countries, the grouping separator separates thousands, but in some countries, it separates ten thousands. If you specify a format string with multiple grouping separators, Data Archive uses the interval between the last grouping separator and the end of the integer. For example, you enter `#,####,###`. Data Archive displays the number 12345678 as 12,345,678.

Character	Description
E	Exponent character. Separates the mantissa from the exponent in scientific notation.

;	Separator for positive and negative patterns in a format string. To enter a numeric format string that provides formatting for positive and negative values, enter the positive format string, the semicolon character (;), and the negative format string.
	For example:
	`#,##0.00;(#,##0.00)`
	Or:
	`0.00;-0.00`
	If you do not provide a negative format string, Data Archive uses the positive format string preceded with the negative sign.

%	Percent sign. Data Archive multiplies the numeric value by 100 and displays it as a percent value.
\u2030	Per mille sign. Data Archive multiplies the numeric value by 1000 and displays it as a per mille value.
¤ (\u00A4)	Currency symbol. Data Archive interprets ¤¤ as the international currency symbol.

Numeric Format String Examples

Data Archive displays numeric values differently based on the numeric format string.

The following table shows how Data Archive interprets different numeric format strings:

Numeric Format String	Numeric Input Value	Result
`###,###.##`	1234356.789	123,456.79
`###.##`	1234356.789	123456.79
`00000.000`	123.45	00123.450
`$###,###.##`	12345.67	$12,345.67
`\u00A5###,###.##`	12345.67	¥ $12,345.67
`0.###E0`	1234	1.234E3
`##0.#####E0`	123456	123.456E3
`00.###E0`	0.00123	12.3E-4
`##0.##E0`	12345	12.3E3

APPENDIX B

Data Archive Connectivity

This appendix includes the following topics:

Data Archive Connectivity Overview

Data Archive uses different types of connectivity to communicate with source and target databases.

Native connectivity

The ILM Engine connects to the data source using a vendor-provided JDBC driver or other vendor proprietary connectivity.

DataDirect Connect JDBC connectivity

The ILM Engine connects to the data source using DataDirect JDBC drivers provided by Informatica. DataDirect JDBC drivers are packaged with Data Archive and are the preferred method of connecting to data sources for which Data Archive does not provide native connectivity.

PowerExchange connectivity

The ILM Engine connects to a mainframe data source through the PowerExchange ODBC drivers that are included with PowerExchange.

Third-party JDBC connectivity

Data Archive supports connectivity to any JDBC data source. If Informatica does not provide an adapter for your desired data source, you may be able to get a third-party JDBC adapter.

Native Connectivity

You can configure Data Archive to use database native utilities for data movement.

Native utilities are the preferred method of connectivity because the ILM Engine can leverage database-specific SQL or vendor-specific operations to improve the performance of retrieving and inserting data.

Native connectivity also enables seamless access from the original application interface.

You can use native connectivity with the following types of databases:

- Oracle
- Microsoft SQL Server
- IBM DB2
- Teradata
- Netezza

You can also use the Teradata Parallel Transporter to export data from Teradata sources on Linux to the Data Vault.

You can configure Netezza databases to use the nzLoad utility for bulk data movement.

DataDirect Connect JDBC Connectivity

You can configure Data Archive to use DataDirect Connect JDBC drivers to connect to data sources.

Data Archive uses JDBC drivers to connect to source, target, and lookup databases. DataDirect Connect JDBC drivers are packaged with Data Archive and are the preferred method of connectivity if native connectivity is not supported.

You can use DataDirect Connect JDBC connectivity with the following types of databases:

- IBM DB2 LUW
- IBM DB2 for i, i5/OS or AS/400
- IBM DB2 for z or z/OS
- Informix
- Sybase ASE

PowerExchange Connectivity

Data Archive can use PowerExchange ODBC drivers to access nonrelational data on z/OS.

Data Archive uses a JDBC-ODBC bridge to connect to the z/OS source data through the ODBC drivers that are included with PowerExchange.

You must install the UNIX ODBC package if your PowerExchange adapter is on one of the following operating systems:

- UX
- Linux

For information on installing the UNIX ODBC package, go to: http://www.unixodbc.org/

After you install the UNIX ODBC package, set the environmental variable `LD_PRELOAD` with the file path of the `libodbc.so` file. For example:

```
export LD_PRELOAD=/usr/lib64/libodbc.so
```

You can retire the following types of nonrelational z/OS data:

- Adabas
- C-ISAM
- Datacom
- IDMS
- IMS
- VSAM
- Sequential data sets

Third-Party JDBC Connectivity

You can configure Data Archive to use third-party JDBC adapters to connect to data sources.

Data Archive supports connectivity to any JDBC data source through third-party adapters.

INDEX

G

H

I

J

L

M

O

P

Q

R

S

T

U

Z