![Informatica logo]

Informatica® Data Archive
6.4.3

User Guide

Informatica Data Archive User Guide
6.4.3
December 2016

© Copyright Informatica LLC 2003, 2018

This product includes software licensed under the terms at http://www.tcl.tk/software/tcltk/license.html, http://www.bosrup.com/web/overlib/?License, http://www.stlport.org/doc/ license.html, http://asm.ow2.org/license.html, http://www.cryptix.org/LICENSE.TXT, http://hsqldb.org/web/hsqlLicense.html, http://httpunit.sourceforge.net/doc/ license.html, http://jung.sourceforge.net/license.txt , http://www.gzip.org/zlib/zlib_license.html, http://www.openldap.org/software/release/license.html, http://www.libssh2.org, http://slf4j.org/license.html, http://www.sente.ch/software/OpenSourceLicense.html, http://fusesource.com/downloads/license-agreements/fuse-message-broker-v-5-3- license-agreement; http://antlr.org/license.html; http://aopalliance.sourceforge.net/; http://www.bouncycastle.org/licence.html; http://www.jgraph.com/jgraphdownload.html; http://www.jcraft.com/jsch/LICENSE.txt; http://jotm.objectweb.org/bsd_license.html; . http://www.w3.org/Consortium/Legal/2002/copyright-software-20021231; http://www.slf4j.org/license.html; http://nanoxml.sourceforge.net/orig/copyright.html; http://www.json.org/license.html; http://forge.ow2.org/projects/javaservice/, http://www.postgresql.org/about/licence.html, http://www.sqlite.org/copyright.html, http://www.tcl.tk/software/tcltk/license.html, http://www.jaxen.org/faq.html, http://www.jdom.org/docs/faq.html, http://www.slf4j.org/license.html; http://www.iodbc.org/dataspace/iodbc/wiki/iODBC/License; http://www.keplerproject.org/md5/license.html; http://www.toedter.com/en/jcalendar/license.html; http://www.edankert.com/bounce/index.html; http://www.net-snmp.org/about/license.html; http://www.openmdx.org/#FAQ; http://www.php.net/license/3_01.txt; http://srp.stanford.edu/license.txt; http://www.schneier.com/blowfish.html; http://www.jmock.org/license.html; http://benalman.com/about/license/; https://github.com/CreateJS/EaselJS/blob/master/src/easeljs/display/Bitmap.js; http://www.h2database.com/html/license.html#summary; http://jsoncpp.sourceforge.net/LICENSE; http://jdbc.postgresql.org/license.html; http://protobuf.googlecode.com/svn/trunk/src/google/protobuf/descriptor.proto; https://github.com/rantav/hector/blob/master/LICENSE; http://web.mit.edu/Kerberos/krb5-current/doc/mitK5license.html; http://jibx.sourceforge.net/jibx-license.html; https://github.com/lyokato/libgeohash/blob/master/LICENSE; https://github.com/hjiang/jsonxx/blob/master/LICENSE; https://code.google.com/p/lz4/; https://github.com/jedisct1/libsodium/blob/master/LICENSE; http://one-jar.sourceforge.net/index.php?page=documents&file=license; https://github.com/EsotericSoftware/kryo/blob/master/license.txt; http://www.scala-lang.org/license.html; https://github.com/tinkerpop/blueprints/blob/master/LICENSE.txt; http://gee.cs.oswego.edu/dl/classes/EDU/oswego/cs/dl/util/concurrent/intro.html; https://aws.amazon.com/asl/; https://github.com/twbs/bootstrap/blob/master/LICENSE; https://sourceforge.net/p/xmlunit/code/HEAD/tree/trunk/LICENSE.txt; https://github.com/documentcloud/underscore-contrib/blob/master/LICENSE, and https://github.com/apache/hbase/blob/master/LICENSE.txt.

This product includes software licensed under the Academic Free License (http://www.opensource.org/licenses/afl-3.0.php), the Common Development and Distribution License (http://www.opensource.org/licenses/cddl1.php) the Common Public License (http://www.opensource.org/licenses/cpl1.0.php), the Sun Binary Code License Agreement Supplemental License Terms, the BSD License (http:// www.opensource.org/licenses/bsd-license.php), the new BSD License (http://opensource.org/licenses/BSD-3-Clause), the MIT License (http://www.opensource.org/licenses/mit-license.php), the Artistic License (http://www.opensource.org/licenses/artistic-license-1.0) and the Initial Developer's Public License Version 1.0 (http://www.firebirdsql.org/en/initial-developer-s-public-license-version-1-0/).

This product includes software copyright © 2003-2006 Joe Walnes, 2006-2007 XStream Committers. All rights reserved. Permissions and limitations regarding this software are subject to terms available at http://xstream.codehaus.org/license.html. This product includes software developed by the Indiana University Extreme! Lab. For further information please visit http://www.extreme.indiana.edu/.

This product includes software Copyright (c) 2013 Frank Balluffi and Markus Moeller. All rights reserved. Permissions and limitations regarding this software are subject to terms of the MIT license.

See patents at https://www.informatica.com/legal/patents.html.

DISCLAIMER: Informatica LLC provides this documentation "as is" without warranty of any kind, either express or implied, including, but not limited to, the implied warranties of noninfringement, merchantability, or use for a particular purpose. Informatica LLC does not warrant that this software or documentation is error free. The information provided in this software or documentation may include technical inaccuracies or typographical errors. The information in this software and documentation is subject to change at any time without notice.

NOTICES

This Informatica product (the "Software") includes certain drivers (the "DataDirect Drivers") from DataDirect Technologies, an operating company of Progress Software Corporation ("DataDirect") which are subject to the following terms and conditions:

1. THE DATADIRECT DRIVERS ARE PROVIDED "AS IS" WITHOUT WARRANTY OF ANY KIND, EITHER EXPRESSED OR IMPLIED, INCLUDING BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE AND NON-INFRINGEMENT.
2. IN NO EVENT WILL DATADIRECT OR ITS THIRD PARTY SUPPLIERS BE LIABLE TO THE END-USER CUSTOMER FOR ANY DIRECT, INDIRECT, INCIDENTAL, SPECIAL, CONSEQUENTIAL OR OTHER DAMAGES ARISING OUT OF THE USE OF THE ODBC DRIVERS, WHETHER OR NOT INFORMED OF THE POSSIBILITIES OF DAMAGES IN ADVANCE. THESE LIMITATIONS APPLY TO ALL CAUSES OF ACTION, INCLUDING, WITHOUT LIMITATION, BREACH OF CONTRACT, BREACH OF WARRANTY, NEGLIGENCE, STRICT LIABILITY, MISREPRESENTATION AND OTHER TORTS.

Publication Date: 2018-06-25

Table of Contents

Preface

The *Informatica Data Archive User Guide* is written for the Database Administrator (DBA) that performs key tasks involving data backup, restore, and retrieval using the Informatica Data Archive user interface. This guide assumes you have knowledge of your operating systems, relational database concepts, and the database engines, flat files, or mainframe systems in your environment.

Informatica Resources

Informatica Network

Informatica Network hosts Informatica Global Customer Support, the Informatica Knowledge Base, and other product resources. To access Informatica Network, visit https://network.informatica.com.

As a member, you can:

- Access all of your Informatica resources in one place.

- Search the Knowledge Base for product resources, including documentation, FAQs, and best practices.

- View product availability information.

- 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 search Informatica Network for product resources such as documentation, how-to articles, best practices, and PAMs.

To access the Knowledge Base, visit https://kb.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

To get the latest documentation for your product, browse the Informatica Knowledge Base at https://kb.informatica.com/_layouts/ProductDocumentation/Page/ProductDocumentSearch.aspx.

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

Informatica Product Availability Matrixes

Product Availability Matrixes (PAMs) indicate the versions of operating systems, databases, and other types of data sources and targets that a product release supports. If you are an Informatica Network member, you can access 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. Developed from the real-world experience of hundreds of data management projects, Informatica Velocity represents the collective knowledge of our consultants who have worked with organizations from around the world to plan, develop, deploy, and maintain successful data management solutions.

If you are an Informatica Network member, you can access 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 augment, extend, or enhance your Informatica implementations. By leveraging any of the hundreds of solutions from Informatica developers and partners, you can improve your productivity and speed up time to implementation on your projects. You can access Informatica Marketplace at https://marketplace.informatica.com.

Informatica Global Customer Support

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

To find your local Informatica Global Customer Support telephone number, visit the Informatica website at the following link:
http://www.informatica.com/us/services-and-training/support-services/global-support-centers.

If you are an Informatica Network member, you can use Online Support at http://network.informatica.com.

CHAPTER 1

Introduction

This chapter includes the following topics:

Data Archive Overview

Data Archive is a comprehensive solution for relocating application data to meet a range of business requirements.

Features

- Enterprise archive engine for use across multiple databases and applications
- Streamlined flow for application retirement
- Analytics of data distribution patterns and trends
- Multiple archive formats to meet different business requirements
- Seamless application access to archived data
- Data Discovery search of offline archived or retired data
- Accessibility or access control of archived or retired data
- Retention policies for archived or retired data
- Purging of expired data based on retention policies and expiration date
- Catalog for the retired applications to see retired application details
- Complete flexibility to accommodate application customizations and extensions

Benefits

- Improve production database performance
- Maintain complete application integrity
- Comply with data retention regulations
- Retire legacy applications
- High degreee of compression reduces application storage footprint
- Enable accessibility to archived data

- Immutability ensures data authenticity for compliance audits
- Reduce legal risk for data retention compliance

Data Archive Use Cases

Improving application performance, complying with data retention regulations, and retiring legacy applications are all use cases for Data Archive.

Data Archive for Performance

Customers need to continuously improve the operational efficiency of their production environment. The Data Archive for Performance solution provides flexible archive and purge functionality to quickly reduce the overall size of production databases. Proactive archiving enables you to keep pace with the increasing rate of your business activity. Data Archive reduces transaction volume for intensively used modules, thereby ensuring that inactive data does not interfere with the day to day use of active data.

Data Archive for Compliance

Data Archive for Compliance retires data with encapsulated BCP files to meet long term data retention, legal, and regulatory compliance requirements. Each object contains the entire configuration, operational, and transactional data needed to identify any business transaction.

With Enterprise Data Manager, accelerators for Oracle E-Business Suite, PeopleSoft, and Siebel applications are available to "jump start" your compliance implementation. In addition, you can use Enterprise Data Manager for additional metadata customization.

Data Archive for Retirement

Organizations upgrading or implementing newer versions of enterprise applications often face the challenge of existing legacy applications that have outlived their usefulness. To eliminate infrastructure and licensing costs, as well as functional redundancy, organizations must shut down these applications. However, the question of application data remains. Many organizations want to ensure accessibility to this older data without either converting it to the format of a new application or maintaining the technology necessary to run the legacy application. Data Archive provides all functionality necessary to retire legacy applications.

CHAPTER 2

Accessing Data Archive

This chapter includes the following topics:

Data Archive URL

Access Data Archive with the following URL format:

```
http://<host name>:<port number>/<web application name>
```

Or, if SSL is enabled:

```
https://<host name>:<port number>/<web application name>
```

The URL includes the following variables:

Host Name

Name of the machine where Data Archive is installed.

Port Number

Port number of the machine where Data Archive is installed.

Web Application Name

Name of the Data Archive web application, typically `Informatica`. Required if you do not use embedded Tomcat. Omit this portion of the URL if you use embedded Tomcat.

When you launch the URL, the Data Archive login page appears. Enter your user name and password to log in.

Log In and Out of Data Archive

Besides initial login, Data Archive provides its user the capability to log in again. Typically, this functionality is meant for switching between users. However, it might also be applicable when a user's access level is modified. In such a circumstance, a re-login is necessary.

A user can log in again from Home > Login Again and log out from Home > Logout.

Update User Profile

View and edit your user profile from **Home** > **User Profile**. You can update your name, email and password. You can also view the system-defined roles assigned to you.

CHAPTER 3

Working with Data Archive

This chapter includes the following topics:

- Data Archive Editions, 21
- Organization of Menus, 21
- Predefined List of Values (LOVs), 22
- Online Help, 22

Data Archive Editions

Data Archive is available in Standard and Enterprise editions. This manual looks at the application from the Enterprise edition perspective and explains functionality differences within editions, wherever applicable.

Note: The Standard edition of Enterprise Data Management Suite allows the user to select only a single ERP, whereas the Enterprise edition encompasses all supported platforms.

Organization of Menus

The functionality of Data Archive is accessible through a menu system consisting of seven top-level menus and respective submenus as shown in the following figure:
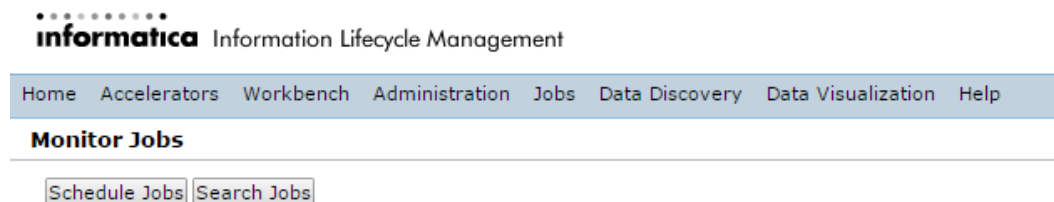
Note: Some of these menu options might not be available in the Standard edition of Data Archive. System-defined roles restrict the menu paths that are available to users. For example, only users with the discovery user role have access to the Data Discovery menu.

Predefined List of Values (LOVs)

Most tasks performed from the Data Archive user interface involve selecting a value from a list. Such lists are abbreviated as LOVs and their presence is marked by the LOV button, usually followed by an asterisk sign (*) to indicate that it is mandatory to specify a value for the preceding text box.

The list is internally generated by a query constructed from Constraints specified on a number of tables through the Enterprise Data Manager.

Mandatory Fields

Throughout the application, red asterisks * indicate mandatory information that is essential for subsequent processes.

Edit and Delete Options

Most sections of the application provide edit and delete options for elements like Policies, Users, Security Groups, Repositories, and Jobs, in the form of an Edit button and a Delete button respectively.

Clicking Edit usually leads you to the Create/Edit page for the corresponding element in the section, while clicking Delete triggers a confirmation to delete the element permanently.

Navigation Options

The breadcrumb link in each page denotes the current location in the application.

When the number of elements in a section, represented as rows (for example, Roles i.e. "elements" in the Role Workbench i.e. "section") exceed ten in number, navigation options are introduced, to navigate to the first, previous, next, and last page respectively.

The user can change the number of rows to be displayed per page.

The number of elements (represented as rows) in a section are specified at the end of a page.

Expand and Collapse Options

In certain sections, elements (i.e. rows) contain detailed information and are expandable. For such sections, Expand All and Collapse All options are available on the top right corner of the section.

Additionally, for each row, a Collapse All button to the left suggests that further detail can be extracted by clicking on it.

Online Help

The Data Archive user interface is enabled with a context sensitive online help system which can be accessed by clicking on the Help button on the top right corner of the screen.

CHAPTER 4

Scheduling Jobs

This chapter includes the following topics:

Scheduling Jobs Overview

Data Archive allows a user to run one-time jobs or schedule jobs to run periodically. Users can also search, view, monitor, and manage jobs.

Standalone Jobs

A standalone job is a job that is not linked to a Data Archive project. Standalone jobs are run to accomplish a specific task.

You can run standalone jobs from Data Archive. Or, you can use a JSP-API to run standalone jobs from external applications.

Allocate Datafiles to Mount Points Job

The Allocate Datafiles to Mount Points standalone job determines how many datafiles to allocate to a particular tablespace, based on the maximum datafile size you enter. Before allocation, the job estimates the size of tables in the segmentation group based on row count.

Before you run this job, run the Get Table Row Count Per Segment job to estimate the size of the segments that the smart partitioning process will create. Then run the Allocate Datafiles to Mount Points job before you create a segmentation policy.

Provide the following information to run this job:

SourceRep

Source connection for the segmentation group. Choose from available source connections.

SegmentationGroupName

Name of the segmentation group that contains the datafiles you want to allocate. Choose from available segmentation groups.

File Size(MB)

Size of the datafiles you want to allocate. Default is 20000 MB.

Auto Extend Increment(MB)

Increment by which the datafiles will automatically extend. Default is 1000 MB.

Max Size(MB)

Maximum size of the datafiles you want to allocate. Default is 30000 MB.

IndexCreationOption

Type of index that will be created for the segment tablespaces. Select global, local, or global for unique indexes.

Archive Structured Digital Records

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.

In the job step log on the **Monitor Jobs** page, you can view a row count report for the records that were successfully loaded to Data Vault.

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`.

After you run the job, assign a Data Vault access role to the entity. Then, assign users to the Data Vault access role. Only users that have the same role assignment as the entity can use Browse Data to view the structured digital records in Data Discovery.

Metadata File

The metadata file is an XML file that contains the details of the entity that the job creates. The file also contains the structure of the digital records, such as the column and row separators, and the table columns and datatypes.

The following example shows a sample template file for the structured metadata file:

```xml
<?xml version="1.0" encoding="UTF-8"?>

<ATTACHMENT_DESCRIPTOR>
    <ENTITY>
            <APP_NAME>CDR</APP_NAME>
            <APP_VER>1.10</APP_VER>
            <MODULE>CDR</MODULE>
            <ENTITY_NAME>CDR</ENTITY_NAME>
            <ENTITY_DESC>CDR Records</ENTITY_DESC>
```

```
        </ENTITY>
        <FILE_DESC>
                <COLUMN_SEPARATOR>,</COLUMN_SEPARATOR>
                <ROW_SEPARATOR>%#%#</ROW_SEPARATOR>
                <DATE_FORMAT>yyyy-MM-dd-HH.mm.ss</DATE_FORMAT>
                <TABLE NAME="CDR">
                        <COLUMN PRIMARY_KEY="Y" ORDER="1" DATATYPE="NUMERIC"
PRECISION="10">CALL_FROM</COLUMN>
                        <COLUMN ORDER="2" DATATYPE="NUMERIC" PRECISION="10">CALL_TO</COLUMN>
                        <COLUMN ORDER="3" DATATYPE="DATETIME">CALL_DATE</COLUMN>
                        <COLUMN ORDER="4" DATATYPE="VARCHAR" LENGTH="100">LOCATION</COLUMN>
                        <COLUMN ORDER="5" DATATYPE="NUMERIC" PRECISION="10"
SCALE="2">DURATION</COLUMN>
                </TABLE>
        </FILE_DESC>
</ATTACHMENT_DESCRIPTOR>
```

You can also indicate whether or not a column is nullable. Use the following syntax: `<COLUMN NULLABLE="N">
</COLUMN>` or `<COLUMN NULLABLE="Y"></ COLUMN>`.

For example:

```
<TABLE NAME="SAMPLE">

<COLUMN PRIMARY_KEY="Y" ORDER="9" DATATYPE="DATETIME" LENGTH="1000"
NULLABLE="N">callDate_callTime</COLUMN>

</TABLE>
```

Archive Structured Digital Records Job Parameters

Each standalone job requires program-specific parameters as input.

The following table describes the job parameters for the Archive Structured Digital Records job:

Parameter	Description
Metadata File	XML file that contains the details of the entity that the job creates and the structure of the digital records, such as the column and row separators, and the table columns and datatypes.
Target Archive Store	Destination where you want to load the structured digital records. Choose the corresponding Data Vault target connection. The list of values is filtered to only show Data Vault target connections.
Purge After Load	Determines whether the job deletes the attachments from the staging directory that is configured in the Data Vault target connection. - Yes. Deletes attachments from the directory. - No. Keeps the attachments in the directory. You may want to select no if you plan to manually delete the attachments after you run the standalone job.

Attach Segments to Database Job

The attach segments to database job reattaches a segment set that you previously detached from a database with the detach segments from database job. You might use the attach segments to database job if you want to move segments that you previously detached and archived back to into a production system.

The attach segments job does not move the .dmp and database files for the transportable table space. When you run the job to reattach the segments, you must manually manage the required .dmp and database files for the table space.

You cannot detach or reattach the default segment.

Provide the following information to run this job:

SourceRep

Source connection for the segmentation group. Choose from available source connections.

SegmentationGroupName

Name of the segmentation group that contains the segments you want to attach. Choose from available segmentation groups.

SegmentSetName

Name of the segment set that you want to attach to the database.

DmpFileFolder

Directory where the detached segment files exist.

DmpFileName

Name of the file you want to attach to the database.

Check Indexes for Segmentation Job

The Check Indexes for Segmentation job determines if smart partitioning will use existing table indexes for segmentation.

Table indexes can help optimize smart partitioning performance and significantly reduce smart partitioning run time. Typically smart partitioning uses the index predefined by the application. If multiple potential indexes exist in the application schema for the selected table, the Check Indexes for Segmentation job determines which index to use. You must register the index after you add the table to the segmentation group.

Before you run the Check Indexes for Segmentation job, you must configure a segmentation policy and generate metadata for the segmentation group. Save the segmentation policy as a draft and run the job before you run the segmentation policy.

Provide the following information to run this job:

SourceRep

Source connection for the segmentation group. Choose from available source connections.

SegmentationGroupName

Name of the segmentation group that you want to check indexes for. Choose from available segmentation groups.

Related Topics:

- "Smart Partitioning Segmentation Policies Overview" on page 225
- "Segmentation Policy Process" on page 226
- "Creating a Segmentation Policy" on page 230

Clean Up After Merge Partitions

After you run the merge partitions into single partition job, run the clean up after merge partition job. The clean up after merge partitions job drops the empty partitions and tablespaces, and merges the partition metadata.

Only run the clean up job after you have merged all of the segments that you want to merge for a data classification. After you run the clean up after merge partitions job, you cannot replace the merged partitions with the original partitions.

Provide the following information to run this job:

SourceRep

Source connection for the segmentation group. Choose from available source connections.

DataClassificationName

Name of the data classification that contains the segments that you merged. Choose from available data classifications.

Clean Up After Segmentation Job

The clean up after segmentation job cleans up temporary tables after the smart partitioning process is complete. The job also gives you the option to drop the original, unpartitioned tables.

SourceRep

Source connection for the segmentation group. Choose from available source connections.

SegmentationGroupName

Name of the segmentation group that you ran a segmentation policy on. Choose from available segmentation groups.

DropManagedRenamedTables

Drops the original unpartitioned table. Choose Yes or No.

Default is No.

Collect Data Growth Stats

The Collect Data Growth Stats job allows you to view database statistics for the selected data source.

Provide the following information to run this job:

- Source repository. Database with production data.

Compress Segments Job

The compress segments job uses native database functionality to compress the tablespaces and local indexes for a selected segment set.

SourceRep

Source connection for the segmentation group. Choose from available source connections.

SegmentationGroupName

Name of the segmentation group containing the set of segments you want to compress. Choose from available segmentation groups.

SegmentCompressOption

> Option to compress the segments. Choose Compress or No Compress.

RemoveOldTablespaces

> Option to remove the old tablespaces after compression.

> Default is yes.

SegmentSetName

> Name of the segment set you want to compress. Choose from available segments sets.

LoggingOption

> The option to log your changes. If you select logging, you can revert your changes.

> Default is NOLOGGING.

ParallelDegree

> The degree of parallelism to be used to compress the segments.

> Default is 4.

Copy Data Classification Job

The copy data classification job copies a data classification from one source connection to another. Use the copy data classification job when you want to use the same data classification, including the dimensions and dimension slices, on a different source connection.

This job uses the following parameters:

SourceRep

> Source connection where the data classification you want to copy exists. Choose from available source connections.

Data Classification

> The data classification that you want to copy to a different source connection. Choose from available data classifications.

To SourceRep

> The source connection that you want to copy the data classification to. Choose from available source connections.

Create Archive 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.

Provide the following information to run this job:

- Target repository. Database with archived data.

- Entity name. Name of the entity for which database indexes need to be created.

Create Optimization Indexes Job

The create optimization indexes job creates a temporary optimization index for a segmentation group when the segmentation group does not have a pre-defined index that is sufficient for smart partitioning. Optimization indexes increase smart partitioning performance.

Provide the following information to run this job:
SourceRep

> Source connection for the segmentation group. Choose from available source connections.

SegmentationGroupName

> Name of the segmentation group that you want to create an optimization index for. Choose from available segmentation groups.

Create Audit Snapshot Job

The create audit snapshot job creates a pre-segmentation audit of the tables that you want to run a segmentation policy on. The audit records table row counts as well as objects like triggers and indexes.

Provide the following information to run this job:
SourceRep

> Source connection for the segmentation group. Choose from available source connections.

SegmentationGroupName

> Name of the segmentation group that you want to audit. Choose from available segmentation groups.

Description

> Enter a description to be used in the compare audit snapshots jobs.

CountHistory

> Option to collect the history table row counts.

> Default is no.

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.

Provide the following information to run this job:

- Destination Repository. Archive folder you want to load the archive data to. Choose the archive folder from the list of values. The target connection name appears before the archive folder name.

Create History Index

The Create History Index job creates database indexes for all tables involved during seamless data access. Run this job after running the Create Seamless Data Access job.

Provide the following information to run this job:

- Output location of generated script file. The output location of the generated script file.

- Source repository. Database with production data.

- Target repository. Database with archived data.

Create History Table

The Create History Table job creates table structures for all tables that will be involved in seamless data access for a data source and data target. Run this job prior to running the Create Seamless Data Access job.

Note: If additional tables are moved in a second archive job execution from the same data source, seamless data access will not be possible for the archived data, unless the Create History Tables job is completed.

Provide the following information to run this job:

- Output location of generated script file. The output location of the generated script file.
- Source repository. Database with production data.
- Target repository. Database with archived data.

Create Indexes on Data Vault

Before you can search for records across applications in Data Vault, you must first create search indexes. Run the Create Indexes on Data Vault job to create a search index for each table that you want to include in the search.

If you add or remove columns from the search index, delete the search indexes first. Then run the Create Indexes on Data Vault to create the new indexes.

Provide the following information to run this job:

Destination Repository

Required. Database with archived data.

Entity

Optional. The name of the entity for the table with the columns you want to add or remove from the search index.

Table

Optional. The name of the table with the columns you want to add or remove from the search index.

Create Materialized Views Job

You can define materialized views in the Enterprise Data Manager. After you define the materialized views, run the create materialized views job to create the views in the Data Vault.

Provide the following information to run the job:

Destination Repository

The archive folder in the Data Vault where you retired the application data. Click the list of values button and select from the available folders.

Schema Name

The schema in which you want to create the materialized views.

Entity

The entity for the view. This field is optional.

View

The name of the view. If you selected a value for the entity parameter, the entity value overrides the view parameter. This field is optional.

Create Seamless Data Access Job

The Data Archive Seamless Data Access job creates the two schemas and populates them with the appropriate synonyms and views.

Provide the following information to run this job:

- 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.

Create Seamless Data Access Script Job

To configure seamless access, you must create views and synonyms in the database. If your policies do not allow external applications to create database objects, use the Create Seamless Data Access Script standalone job to create a script that you 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

Provide the following information to run the Seamless Data Access Script job:

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.

Create Seamless Data Access Physical and Logical Script

You can create seamless access between an IBM DB2 for AS/400 source connection with physical and logical files and a IBM DB2 for AS/400 target connection with archived data. Use the Create Seamless Data Access Physical and Logical Script standalone job to create a script that you can review and then run on the

target database. The script includes statements to create the required database objects for seamless access between the source and target connections.

Provide the following information to run this job:

SourceRep

The IBM DB2 for AS/400 source connection that stores the source or production data.

DestRep

The IBM DB2 for AS/400 target connection that stores the archived data.

Script Location

Location where you want to save the scripts that the job creates. This location must be accessible to the ILM application server.

After the job completes successfully, download the script files from the **Monitor Jobs** page. To download the script files, expand the Job ID and click the **Download Physical and Logical Script** link. Then extract the downloaded file, which is named `ScriptGeneration_xx.zip`.

Within the extracted folder, the `bin` folder contains a connections properties file. Update the parameters to be specific to your environment. The `lib_name` and `file_name` parameters should be eight characters or less.

After you edit the connections file, run the `GenerateScriptForDb2AS400.bat` file on a Microsoft Windows operating system or the `GenerateScriptForDb2AS400.sh` file on a UNIX/Linux operating system. When the shell script or batch file runs, it creates the seamless access scripts in the location that you specified in the job parameters.

For more information, see the *Informatica Data Archive Administrator Guide*.

Copy Application Version for Retirement

Use the Copy Application Version for Retirement job to copy metadata from a pre-packaged application version to a customer-defined application version. Run the job to modify the metadata in a customer-defined application version. The metadata in pre-packaged application versions is read-only.

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.

Provide the following information to run this job:

- Product Family Version. The pre-packaged application version that you want to copy.
- New Product Family Version Name. A name for the application version that the job creates.

Copy Source Metadata Job

The copy source metadata job copies the metadata from one ILM repository to another ILM repository on a cloned source connection.

You might need to clone a production database and still want to manage the database with smart partitioning. Or, you might want to clone a production database to subset the database. Run the copy source metadata job if you want to use the same ILM repository information on the cloned database as the original source connection.

Before you run the job, clone the production database and create a new source connection in the Data Archive user interface. Then, run the define staging schema standalone job or select the new staging schema from the **Manage Segmentation** page.

Provide the following information to run this job:

From Source Repository

>   The staging schema of the source repository that you want to copy metadata from.

To Source Repository

>   The staging schema of the source repository that you want to copy metadata to.

Define Staging Schema Job

The define staging schema job defines or updates the staging schema for a source connection. Smart partitioning creates database objects such as packages, tables, views, and procedures in the staging schema during initial configuration. You might need to update the staging schema to update a view definition, or if you upgrade the Data Archive software to a more recent version.

Provide the following information to run this job:

SourceRep

>   Source connection for the segmentation group. Choose from available source connections.

DefinitionOption

>   The choice to define or update the staging schema. If you choose define staging schema, the ILM Engine performs a check to determine if the staging schema has been configured previously. If you choose update staging schema, the job updates the staging schema regardless of whether it has been previously configured.

Delete Indexes on Data Vault

The Delete Indexes on Data Vault job deletes all search indexes used by Search Data Vault. Run this job every time you add or remove columns from search indexes. You must also run this job after you purge an application that was included in the search indexes. After this job completes, run the Create Indexes on Data Vault job to create the new search indexes.

To run the Delete Indexes on Data Vault job, a user must have one of the following system-defined roles:

- Administrator
- Retention Administrator
- Scheduler

Detach Segments from Database Job

The detach segments from database job detaches a segment set from a database with transportable table space functionality. You might use the detach segments job if you want to archive a set of segments to preserve space in your production system. When you detach the segments, you can archive or move them and reattach them later with the attach segments to database job, if you choose.

If you run the detach segments from database job, you can no longer run the replace segmented tables with original tables job. You cannot detach or reattach the default segment.

Provide the following information to run this job:

SourceRep

>   Source connection for the segmentation group. Choose from available source connections.

SegmentationGroupName

Name of the segmentation group that contains the segments you want to detach. Choose from available segmentation groups.

SegmentSetName

Name of the segment set that you want to detach from the database.

DmpFileFolder

Directory where you want to save the detached segment data files.

DmpFileName

File name of the detached segments data files. Enter a name.

Disable Access Policy Job

The disable access policy job disables the access policy in use for a specific segmentation group.

Provide the following information to run this job:

SourceRep

Source connection for the segmentation group. Choose from available source connections.

SegmentationGroupName

Name of the segmentation group that you want to disable the access policy on. Choose from available segmentation groups.

Drop History Segments Job

The drop history segments job drops the history segments that the smart partitioning process creates when you run a segmentation policy. If you want to create a subset of history segments for testing or performance purposes, you must drop the original history segments.

Before you create a subset of the history segments, you must clone both the application database and the ILM repository. You must also create a copy of the ILM engine and edit the `conf.properties` file to point to the new ILM repository. Then you can create a segment set that contains the segments you want to drop and run the drop history segments job.
Provide the following information to run this job:

SourceRep

Source connection for the segmentation group. Choose from available source connections.

SegmentationGroupName

Name of the segmentation group that contains the history segments you want to drop. Choose from available segmentation groups.

SegmentSetName

Name of the segment set that contains the history segments you want to drop.

GlobalIndexTablespaceName

Name of the global index tablespace.

EstimatePercentage

The sampling percentage. Valid range is 0-100.

Default is 10.

ParallelDegree

> The degree of parallelism the ILM Engine uses to drop history segments.
>
> Default is 4.

Drop Interim Tables Job

The drop interim tables job drops the interim tables that the smart partitioning process creates when you configure interim tables to pre-process business rules.

Provide the following information to run this job:

SourceRep

> Source connection for the segmentation group. Choose from available source connections.

SegmentationGroupName

> Name of the segmentation group that contains the interim tables you want to drop. Choose from available segmentation groups.

Enable Access Policy Job

The enable access policy job enables the access policy you have configured for a segmentation group.

Provide the following information to run this job:

SourceRep

> Source connection for the segmentation group. Choose from available source connections.

SegmentationGroupName

> Name of the segmentation group with the access policy you want to enable. Choose from available segmentation groups.

DbPolicyType

> Type of the row-level security policy on the application database. Choose from context sensitive, dynamic, or shared context sensitive.

Export Data Classification Job

The export data classification standalone job exports the details of a data classification in XML format, including all dimensions and dimension slices associated with the data classification. After you export a data classification, you can import it to a different ILM repository with the import data classification job. The export and import data classification jobs allow you to share data classifications between Data Archive installations.

This job uses the following parameters:

SourceRep

> Source connection where the data classification exists. Choose from available source connections.

Data Classification

> Name of the data classification you want to export. Choose from available data classifications.

Exported file location (on server)

> The directory location where you want to export the data classification XML file to.

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.

Provide the following information to run this job:

- Archive Job ID. Job ID generated after you publish an archive project. Locate the archive job ID from the **Monitor Job** page.

Note: Data Vault supports a maximum of 4093 columns in a table. If the Data Vault Loader tries to load a table with more than 4093 columns, the job fails.

Generate Explain Plan Job

The Generate Explain Plan job creates a plan for how the smart partitioning process will run the selection statements that create segments. This plan can help determine how efficiently the statements will run.

You must create a segmentation group in the Enterprise Data Manager and configure a segmentation policy in the Data Archive interface before you run the Generate Explain Plan job. Save the segmentation policy as a draft and run the job before you run the segmentation policy.

If you enabled interim table processing to pre-process business rules for the segmentation group, you must run the Pre-process Business Rules job before you run the Generate Explain Plan job.

Provide the following information to run this job:

SourceRep

Source connection for the segmentation group. Choose from available source connections.

SegmentationGroupName

Name of the segmentation group you want to generate the explain plan for. Choose from available segmentation groups.

SelectionStatementType

Type of selection statements. Select ongoing, base, or history.

OutFileFullName

File name of the explain plan.

ScriptOnly

Default is no.

Get Table Row Count Per Segment Job

The Get Table Row Count Per Segment job estimates the size of the segments that will be created when you run a segmentation policy.

If you enabled interim table processing to pre-process business rules for the segmentation group, you must run the Pre-process Business Rules job before you run the Get Table Row Count Per Segment job. Provide the following information to run this job:

SourceRep

Source connection for the segmentation group. Choose from available source connections.

SegmentationGroupName

Name of the segmentation group that you want to estimate segment row counts for. Choose from available segmentation groups.

SelectionStatementType

Type of selection statement. Select ongoing, base, or history.

IBM DB2 Bind Package

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.

Provide the following information to run this job:

- 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.

Import Data Classification

The import data classification job imports a data classification to an ILM repository. You must export the data classification with the export data classification job before you run the import data classification job. The export and import data classification jobs allow you to share data classifications between Data Archive installations.

To successfully import a data classification, you must create identical dimensions in the Enterprise Data Manager for both ILM repositories. The dimensions must have the same name, datatype, and dimension type.

This job uses the following parameters:

SourceRep

Source connection that you want to import the data classification to. Choose from available source connections.

Exported file path

The complete file path, including the XML name, of the data classification XML file that you want to import.

Load External Attachments

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 job creates an application module in EDM, called EXTERNAL_ATTACHMENTS. Under the application module, the job creates an entity and includes the AM_ATTACHMENTS table. You specify the entity in the job parameters. The entity allows you to use Browse Data to view the external attachments in Data Discovery without a stylesheet.

After you run the job, assign a Data Vault access role to the entity. Then, assign users to the Data Vault access role. Only users that have the same role assignment as the entity can use Browse Data to view the attachments in Data Discovery. A stylesheet is still required to view attachments in the App View.

You can run the job for attachments that you did not archive yet or for attachments that you archived in a previous release. You may want to run the job for attachments that you archived in a previous release to use Browse Data to view the attachments in Data Discovery.

If the external attachments you want to load have special characters in the file name, you must configure the LANG and LC_ALL Java environment variables on the operating system where the ILM application server runs. Set the LANG and LC_ALL variables to the locale appropriate to the special characters. For example, if the file names contain United States English characters, set the variables to `en_US.utf-8`. Then restart the ILM application server and resubmit the standalone job.

Load External Attachments Job Parameters

Each standalone job requires program-specific parameters as input.

The following table describes the job parameters for the Load External Attachments job:

Parameter	Description
Directory	Directory that contains the attachments you want to archive. For upgrades, choose a directory that does not contain any files.
Attachment Entity Name	Entity name that the job creates. You can change the default to any name. If the entity already exists, the job does not create another entity. Default is `AM_ATTACHMENT_ENTITY`.
Target Archive Store	Data Vault destination where you want to load the external attachments. Choose the corresponding archive target connection. For upgrades, choose the target destination where you archived the attachments to.
Purge After Load	Determines whether the job deletes the attachments from the directory that you provide as the source directory. - Yes. Deletes attachments from the directory. - No. Keeps the attachments in the directory. You may want to select no if you plan to manually delete the attachments after you run the standalone job.

Merge Archived History Data While Segmenting Production Job

The merge archived history data while segmenting production job merges previously archived data with a configured segmentation group before it creates segments.

For example, a previously archived table may contain accounts receivable transactions from 2008 and 2009. You want to merge this data with a segmentation group you have created that contains accounts receivable transactions from 2010, 2011, and 2012, so that you can create segments for each of the five years. Run this job to combine the archived data with the accounts receivable segmentation group and create segments for each year of transactions.

The merge archived history data job applies the business rules that you configured to the production data. Run smart partitioning on the history data separately before you run the job.

If you choose not to merge all of the archived history data into its corresponding segments in the production database, the Audit Compare Snapshots job returns a discrepancy that reflects the missing rows in the row count report.

Provide the following information to run this job:

SourceRep

Source connection for the segmentation group. Choose from available source connections.

SegmentationGroupName

Name of the segmentation group that you want to merge with archived data. Choose from available segmentation groups.

LoggingOption

Option to log your changes. If you select logging, you can revert your changes.

Default is no.

AddPartColumnToNonUniqueIndex

Option to add the partition column to a non-unique index.

Default is no.

IndexCreationOption

Type of index the smart partitioning process creates for the merged and segmented tables.

Default is global.

BitmapIndexOption

Type of bitmap index the smart partitioning process creates for the merged and segmented tables.

Default is local bitmap.

GlobalIndexTablespaceName

Name of the tablespace you want to the global index to exist in. This field is optional.

CompileInvalidObjects

Option to compile invalid objects after the job recreates indexes.

Default is no.

ReSegmentFlag

Default is no.

CheckAccessPolicyFlag

Checks for an access policy on the segmentation group.

Default is yes.

ParallelDegree

Degree of parallelism used when the smart partitioning process creates segments.

Default is 4.

DropTempTablesBeforeIndexRebuild

Drops the temporary tables before rebuilding the index after segmentation.

Default is no.

DropTempSelTables

Drops the temporary selection tables after segmentation.

Default is yes.

ArchiveRunIdView

Database view that maps the partition ID to the archive ID.

ArchiveRunIdColumn

Name of the archive column.

Merge Partitions Into Single Partition Job

Use the merge partitions into single partition job to merge a segment set containing multiple segments into one segment. For instance, the job can merge four segments of history data from four different quarters into one segment that contains data for the entire year.

Provide the following information to run this job:

SourceRep

Source connection for the segmentation group. Choose from available source connections.

SegmentationGroupName

Name of the segmentation group that contains the segments that you want to merge. Choose from available segmentation groups.

SegmentSetName

Name of the segment set that contains the segments to merge.

LoggingOption

The option to log your changes. If you select logging, you can revert your changes.

Default is NOLOGGING.

ParalellDegree

The degree of parallelism to be used to merge the segments.

FinalMergeJob

Parameter to designate whether or not the job is the final merge job that you want to run for a segmentation group. If you have multiple segment sets that you want to merge in a segmentation group, select "No" unless the job is the final segment set that you want to merge in that group. Select "Yes" if the job is the last merge job that you want to run for the segmentation group.

MergedPartitionTablespaceName

Leave empty to use the existing tablespace. You can enter a new tablespace for the merged segment, but you must create the tablespace before running the job.

Move External Attachments

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.

Provide the following information to run this job:

- Add on URL. URL to the Data Vault Service for external attachments.

- Archive Job ID. Job ID generated after you publish an archive project. Locate the archive job ID from the **Monitor Job** page.

Move From Default Segment to History Segments Job

The Move from Default Segment to History Segments job moves closed transactions from the default segment to the existing history segment where the transactions belong, according to the business rules applied.

If you enabled interim table processing to pre-process business rules for the segmentation group, you must run the Pre-process Business Rules job before you run the Move from Default Segment to History Segments job.

Provide the following information to run this job:

SourceRep

Source connection for the segmentation group. Choose from available source connections.

SegmentationGroupName

Name of the segmentation group that contains the default and history segments that you want to move transactions between. Choose from available segmentation groups.

SegmentSetName

Name of the segment set that you want to move from default to history.

This field is optional.

BatchSize

Size of the selection, in table rows, that is moved at once.

Default is 100000.

RefreshSelectionStatements

Refreshes the metadata constructed as base queries for the partitioning process. If you have updated table relationships for the segmentation group in the Enterprise Data Manager, select Yes to refresh the base queries.

Default is no.

CheckAccessPolicyFlag

Checks for an access policy on the segmentation group.

Default is yes.

AnalyzeTables

Calls the Oracle API to collect database statistics at the end of the job run.

Default is no.

DropTempSelTables

> Drops the temporary selection tables after the transactions are moved.
>
> Select yes.

Move From History Segment to Default Segments

The Move from History Segment to Default Segments job moves transactions from an existing history segment back to the default segment for the segmentation group.

If you enabled interim table processing to pre-process business rules for the segmentation group, you must run the Pre-process Business Rules job before you run the Move from History Segment to Default Segments job.

Provide the following information to run this job:

SourceRep

> Source connection for the segmentation group. Choose from available source connections.

SegmentationGroupName

> Name of the segmentation group that contains the history segments you want to move data from. Choose from available segmentation groups.

SegmentSetName

> Name of the segment set that you want to move from history to default.
>
> This field is optional.

BatchSize

> Size of the selection, in table rows, that is moved at once.
>
> Default is 100000.

RefreshSelectionStatements

> Refreshes the metadata constructed as base queries for the partitioning process. If you have updated table relationships for the segmentation group in the Enterprise Data Manager, select Yes to refresh the base queries.
>
> Default is no.

CheckAccessPolicyFlag

> Checks for an access policy on the segmentation group.
>
> Default is yes.

AnalyzeTables

> Calls the Oracle API to collect database statistics at the end of the job run.
>
> Default is no.

DropTempSelTables

> Drops the temporary selection tables after the transactions are moved.
>
> Default is yes.

Move Segment to New Storage Class

The move segment to new storage class job moves the segments in a segmentation group from one storage classification to another.

When you move a segment to a new storage classification, some data files might remain in the original storage classification tablespace. You must remove these data files manually.

If the segment and new storage classification are located on an Oracle 12c database, the job moves the segment without taking the tablespace offline.

Provide the following information to run this job:

SourceRep

Source connection for the segmentation group. Choose from available source connections.

TablespaceName

Name of the tablespace that contains the new storage classification. Choose from available tablespaces.

StorageClassName

Name of the storage classification that you want to move the segments to. Choose from available storage classifications.

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.

When you run the Purge Expired Records job, by default, the job searches all entities in the Data Vault archive folder for records that are eligible for purge. To narrow the scope of the search, you can specify a single entity. The Purge Expired Records job searches for records in the specified entity alone, which potentially decreases the amount of time in the search stage of the purge process.

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, select a date for Data Archive to base the report on. If you select a past date or the current date, Data Archive generates a list of tables with the number of records that are eligible for purge on that date. You can pause the job to review the report and then schedule the job to purge the records. If you resume the job and continue to the delete step, the job deletes all expired records up to the purge expiration date that you used to generate the report. If you provide a future date, Data Archive generates a list of tables with the number of records that will be eligible for purge by the future date. The job stops after it generates the report.

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.

The reports are created with the Arial Unicode MS font. To generate the reports, you must have the font file ARIALUNI.TTF saved in the `<Data Archive installation>\webapp\WEB-INF\classes` directory.

To purge records, you must enable the purge step through the **Purge Deleted Records** parameter. You must also provide the name of the person who authorized the purge.

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.

Purge Expired Records Job Parameters

The following table describes the job parameters for the Purge Expired Records job:

Parameter	Description	Required or Optional
Archive Store	The name of the archive folder in the Data Vault that contains the records that you want to delete. Select a folder from the list.	Required.
Purge Expiry Date	The date that Data Archive uses to generate a list of records that are or will be eligible for delete. Select a past, the current, or a future date. If you select a past date or the current date, Data Archive generates a report with a list of all records eligible for delete on the selected date. You can pause the job to review the report and then schedule the job to purge the records. If you resume the job and continue to the delete step, the job deletes all expired records up to the selected date. If you select a future date, Data Archive generates a report with a list of records that will be eligible for delete by the selected date. However, Data Archive does not give you the option to delete the records.	Required.
Report Type	The type of report to generate when the job starts. Select one of the following options: - **Detail**. Generates the Retention Expiration Detail report. - **None**. Does not generate a report. - **Summary**. Generates the Retention Expiration Summary report.	Required.
Pause After Report	Determines whether the job pauses after Data Archive creates the report. If you pause the job, you must resume it to delete the eligible records. Select **Yes** or **No** from the list.	Required.

Parameter	Description	Required or Optional
Entity	The name of the entity with related or unrelated tables in the Data Vault archive folder. To narrow the scope of the search, select a single entity.	Optional.
Purge Deleted Records	Determines whether to delete the eligible records from the Data Vault. Select **Yes** or **No** from the list.	Required.
Email ID for Report	Email address to which to send the report.	Optional.
Purge Approved By	The name of the person who authorized the purge.	Required if you select a past date or the current date for **Purge Expiry Date**.

Recreate Indexes Job

The recreate indexes job re-creates the indexes for a segmentation group after you run a segmentation policy to create segments. You can use the recreate indexes job to change the type of index from global to local, or local to global.

Provide the following information to run this job:

SourceRep

Source connection for the segmentation group. Choose from available source connections.

SegmentationGroupName

Name of the segmentation group that you want to re-create indexes for. Choose from available segmentation groups.

AddPartColumnToNonUniqueIndex

Option to add the partition column to a non-unique index.

Default is no.

AnalyzeTables

Collects table statistics.

Default is no.

IndexCreationOption

Type of index the smart partitioning process creates.

Default is global.

GlobalIndexTablespaceName

Name of the global index tablespace. Optional.

CompileInvalidObjects

Compiles invalid objects after the job re-creates indexes.

Default is no.

ParallelDegree

The degree of parallelism to be used for re-creating indexes.

Default is 4.

RecreateAlreadySegmentedIndexes

Option to re-create indexes that were previously segmented.

If you upgrade to Oracle E-Business Suite Release 12 from a previous version, some of the segmented local indexes might become unsegmented. When you run the recreate indexes standalone job, the job re-creates and segments the indexes that became unsegmented during the upgrade. You have the option of choosing whether or not you want to re-create the indexes that did not become unsegmented during the upgrade.

To re-create the indexes that are still segmented when you run the recreate indexes job, select **Yes** from the list of values for the parameter.

If you do not want to re-create the already segmented indexes, select **No** from the list of values for the parameter.

Default is yes.

Do not set this parameter to yes if you set the IndexCreationOption parameter to global.

Refresh Materialized Views Job

The refresh materialized views job refreshes selected materialized views that exist in the Data Vault. To run the refresh materialized views job, select the job from the standalone programs list. Then define the job parameters and schedule the job.

Provide the following information to run the job:

Destination Repository

Archive folder in the Data Vault where you retired the application data. Click the list of values button and select from the available folders.

Schema

Name of the retired schema.

Entity

Name of the entity that contains the view or views you want to refresh.This field is optional.

Refresh All Views

Determines whether to refresh all of the materialized views. If you do not know the entity or name of the view you need to refresh, click the list of values button and select **Yes**. If you select **No**, you must provide the entity or the name of the view you want to refresh.

View

Name of view you want to refresh. Click the list of values button and select the view. This field is optional.

Refresh Selection Statements Job

The refresh selection statements job updates the selection statements for a segmentation group, so that you do not have to regenerate the segmentation group before you run a segmentation policy.

Use the refresh selection statements job if you modify a segmentation group parameter after the group has been created. For example, you may want to add a new index or query hint to a segmentation group and validate the selection statement before you run the segmentation policy.

SourceRep

Source connection for the segmentation group. Choose from available source connections.

SegmentationGroupName

Name of the segmentation group that you want to update selection statements for. Choose from available segmentation groups.

SelectionStatementType

Type of selection statement used in the segmentation group. Choose from base, ongoing, or history.

- Base. Choose the base type if you plan to use the split segment method of periodic segment creation for this segmentation group.

- Ongoing. Choose the ongoing type if you plan to use the new default or new non-default method of periodic segment creation for this segmentation group.

- History. Choose the history type if you plan to merge the segmentation group data with history data.

Replace Merged Partitions with Original Partitions Job

Use the replace merged partitions with original partitions job if you need to reverse the merge partitions into single partition standalone job.

You cannot replace the merged partitions with the original partitions if you have already run the clean up after merge partitions job for a segment set.

Provide the following information to run the job:

SourceRep

Source connection for the segmentation group. Choose from available source connections.

SegmentationGroupName

Name of the segmentation group that contains the merged segments that you want to replace with the originals. Choose from available segmentation groups.

SegmentSetName

Name of the segment set that contains the merged segments that you want to replace.

DropBackupTables

Drops the backup tables created during the merge process.

Default is No.

Replace Segmented Tables with Original Tables Job

The replace segmented tables with original tables job mends the tables that you have run a segmentation policy on to return the tables to their original state. Use this job if you want to undo the smart partitioning process in a test environment.

If you run the replace segmented tables with original tables job and then run a segmentation policy on the tables a second time, the segmentation process will fail if any of the table spaces were marked as read-only in the first segmentation cycle. Ensure that the table spaces are in read/write mode before you run the segmentation policy again.

Provide the following information to run this job:

SourceRep

Source connection for the segmentation group. Choose from available source connections.

SegmentationGroupName

Name of the segmentation group that contains the segmented tables you want to replace. Choose from available segmentation groups.

AddPartColumnToNonUniqueIndex

Option to add the partition column to a non-unique index.

Default is no.

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`.

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.

Test JDBC Connectivity

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.

Provide the following information to run this job:

- Connection Retention Time. The length of time the Test JDBC Connectivity job maintains a connection to the repository.

- Repository Name. Source or target database.

Test Scheduler

The Test Scheduler job checks whether the inbuilt scheduler is up and running. You can run this job at anytime.

Turn Segments Into Read-only Mode

The turn segments into read-only job turns a set of segments into read-only mode to prevent their modification.

Provide the following information to run this job:

SourceRep

Source connection for the segmentation group. Choose from available source connections.

SegmentationGroupName

Name of the segmentation group that contains the history segments you want to drop. Choose from available segmentation groups.

SourceSegmentSetName

Name of the segment set you want to turn read-only.

Unpartition Tables Job

The unpartition tables job replaces all of the segmented tables in a segmentation group with tables identical to the original, unsegmented tables. If you have run a segmentation policy and dropped the original tables, you can run the unpartition tables job to revert the segmentation group to its original state. After you run the job, the segmentation group returns to a status of generated.

Provide the following information to run this job:

SourceRep

Source connection for the segmentation group. Choose from available source connections.

SegmentationGroupName

Name of the segmentation group that contains the tables you want to unpartition. Choose from available segmentation groups.

DataTablespaceName

Name of the tablespace where you want the job to create the unpartitioned tables. Enter the name of an existing tablespace.

IndexTablespaceName

Name of the tablespace where you want the job to create the unpartitioned index. Enter the name of an existing tablespace.

LoggingOption

The option to log your changes. If you select logging, you can revert your changes.

Default is NOLOGGING.

Scheduling Jobs

To schedule a standalone job, select the type of job you want to run and enter the relevant parameters.

1. Click **Jobs** > **Schedule a Job**.

2. Choose **Projects** or **Standalone Programs** and click **Add Item**.

    The **Select Definitions** window appears with a list of all available jobs.

3. Optionally, to filter the list of jobs, select **Program Type** and enter the name of a program in the adjacent text field. For example, type Data Vault in the text field to limit the list to Data Vault jobs.

4. Select the job you want to run. Repeat this step to schedule more than one job.

5. Enter the relevant job parameters.

6. Schedule the job to run immediately or on a certain day and time. Choose a recurrence option if you want the job to repeat in the future.

7. Enter an email address to receive notification when the job completes, terminates, or returns an error.

8. Click **Schedule**.

Job Logs

Data Archive creates job logs when you run a job. The log summarizes the actions the job performs and includes any warning or error messages.

You access the job log when you monitor the job status. For archive and retirement jobs, the jobs include a log for each step of the archive cycle. Expand the step to view the log. When you expand the step, you can view a summary of the job log for that step. The summary includes the last few lines of the job log. You can open the full log in a separate window or download a PDF of the log.

Delete From Source Step Log

For archive jobs, the log for the delete from source step includes information on the tables the job deletes records from and how many rows the job deletes from the tables.

For Oracle sources, the job log includes a real-time status of record deletion. You can use the real-time status to monitor the job progress. For example, to determine the amount of time that the job needs to complete the record deletion.

The real-time status is available when the following conditions are true:

- The source is an Oracle database.
- The entity tables include primary key constraints.
- The source connection uses staging and uses the Oracle row ID for deletion.
- The source connection does not use Oracle parallel DML for deletion.

The log includes the number of parallel threads that the job creates and the total number of rows that the job needs to delete for each table. The log includes a real-time status of record deletion for each thread. The job writes a message to the log after each time a thread commits rows to the database. The log displays the thread number that issued the commit and the number of rows the thread deleted in the current commit. The log also displays the total number of rows the thread deleted in all commits and the number of rows that the thread must delete in total. The log shows when each thread starts, pauses, resumes, and completes.

Monitoring Jobs

Current jobs can be viewed from **Jobs** > **Monitor Jobs** and future jobs can be viewed from **Jobs** > **Manage Jobs**.

The **Monitor Jobs** page lists jobs with their ID, status, user, items to execute, start date and time, and completion date and time.

A job can have one of the following values for status:

- Completed
- Error
- Not Executed
- Not Started
- Paused
- Pending

- Running

- Terminated

- Terminating

- Validated

- Warning

Jobs can be displayed in detail through expand and collapse options.

Click the job name, wherever an arrow button appears to the left of it, to display logging information of a particular job. The last few lines from the logging information are displayed inline. To change the number of lines for the display, select a value from the combo box, for "view lines." Click the View Full Screen to display the entire log in a new window. View Log, Summary Report or Simulation Report displays relevant reports as PDF files.

Options are also available to resume and terminate a job. To terminate a job, click **Terminate Job** and to resume a job, click **Resume Job**. Jobs can be resumed if interrupted either by the user or due to external reasons like power failure. When a job corresponding to a Data Archive project errors out, you should terminate that Job before scheduling the same project.

To refresh the status of a job, select **Auto Refresh Job Status**.

Pausing Jobs

You can pause archive jobs during specific steps of the archive job. For example, you can pause the job at the delete from source step for Oracle sources. To pause the job from a specific step, open the status for the step. When you resume the job, the job starts at the point where it paused.

For SAP retirement projects, Data Archive extracts data from special SAP tables as packages. When you resume the job, the job resumes the data extraction process, starting from the point after the last package was successfully extracted.

Delete From Source Step

You can pause and resume an archive job when the job deletes records from Oracle sources.

You can pause and resume an archive job during the delete from source step when the following conditions are true:

- The source is an Oracle database.

- The entity tables include primary key constraints.

- The source connection uses staging and uses the Oracle row ID for deletion.

- The source connection does not use Oracle parallel DML for deletion.

You may want to pause a job during the delete from source step if you delete from a production system and you only have a limited amount of time that you can access the system. If the job does not complete within the access time, you can pause the job and resume it the next time you have access.

When the job pauses depends on the delete commit interval and the degree of parallelism configuration for the archive source. The pause occurs at the number of threads multiplied by the commit interval. When you pause the job, each thread completes the number of rows in the current commit and then pauses. The job pauses after all threads complete the current number of rows in the commit interval.

For example, if you configure the source to use a delete commit interval of 30,000 and two delete parallel threads, then the job pauses after the job deletes 60,000 rows. The job writes a message to the log that the threads are paused.

You can pause the job when you monitor the job status. When you expand the delete from source step, you can click on a link to pause the job. You can resume the job from the step level or at the job level. When you resume the job, the job continues at the point from which you paused the job. The log indicates when the threads pause and resume.

Searching Jobs

To search jobs, you can conduct a quick search or an advanced search facility.

Quick Search

The generated results in quick search are filtered for the currently logged-in user.

The following table describes the quick search parameters:

Search Option	Description
My Previous Error / Paused Job	Last Job that raised an Error or was Paused.
All My Yesterday's Jobs	All Jobs that were scheduled for the previous day.
All my Error Jobs	All Jobs that were terminated due to an Error.
All my Paused Jobs	All Jobs that were Paused by the currently logged-in user.
Search For	A drop down list is available here, to search Jobs by their Job ID, Job Name, or Project name.

Advanced Search

Advanced Search is available for the following parameters, each of which is accompanied by a drop down list containing "Equals" and "Does not Equal", to include or exclude values specified for each parameter.

The following table describes the advanced search parameters:

Search Option	Description
Submitted By	Include or exclude a User who submitted a particular job.
Project/Program	When you search for a project, you can also specify a stage in the data archive job execution, for example, Generate Candidates.
Start Date	Include or exclude a Start Date for the Jobs.
Completion Date	Include or exclude a Completion Date for the Jobs.
Status	Include or exclude a particular Status for Jobs.

CHAPTER 5

Viewing the Dashboard

This chapter includes the following topics:
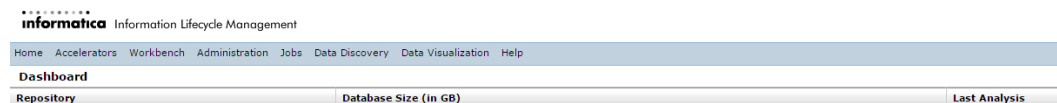
Viewing the Dashboard Overview

The dashboard is a point of reference for repositories configured on Data Archive. This can be accessed from **Administration** > **Dashboard**.

All repositories are displayed in a list, with information such as the database size (in GB), and the date it was last analyzed.

After the DGA_DATA_COLLECTION job is complete, the dashboard reflects an entry for the respective database as shown in the following figure.

Click the **Repository** link to view the graphical representation of Modules, Tablespaces, and Modules Trend.

Note: To view information about a repository in the dashboard, you must run the DGA_DATA_COLLECTION standalone program from the **Schedule Job** page (against that repository).

Scheduling a Job for DGA_DATA_COLLECTION

To run DGA_DATA_COLLECTION:

1. Click **Jobs** > **Schedule a Job**.

   The following page appears on the screen:

2. In the Projects/Programs to Execute area, select **Standalone Programs**.

3. Click **Add Items**.

   The following pop-up window appears on the screen:
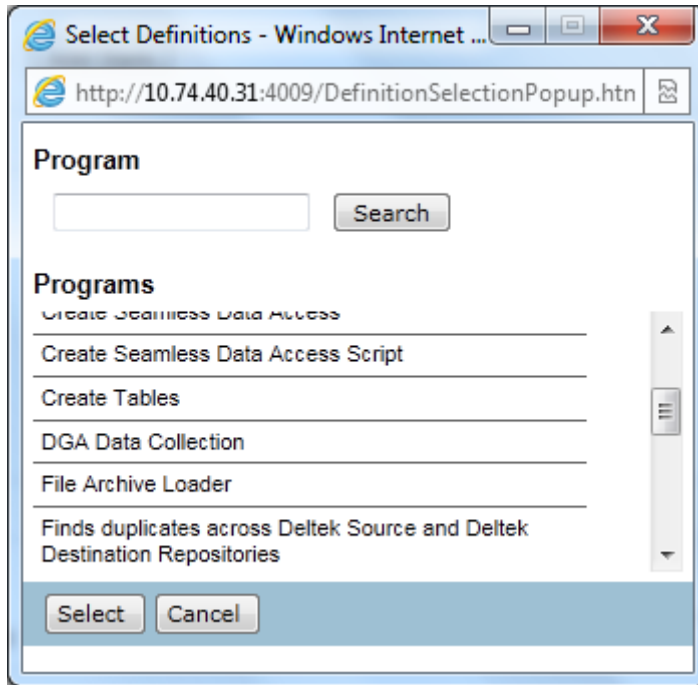
4. Select DGA_DATA_COLLECTION, and click Select.

   The DGA_DATA_COLLECTION section appears on the page as shown in the following figure:
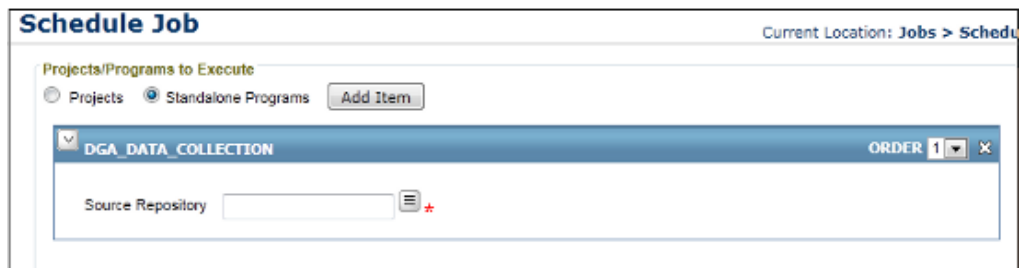
5. Select the Source Repository by clicking the LOV button to select from the LOVs.

6. Specify schedule time as `Immediate` from the Schedule section.

7. Setup notification information such as Email ID for messages on completion of certain events (Completed, Terminated, and Error).

8. Click **Schedule** to run the DGA_DATA_COLLECTON job.

CHAPTER 6

Creating Data Archive Projects

This chapter includes the following topics:

Setting Up Data Archive Projects

Transaction data can be archived either to a database archive or the Data Vault. Data Archive relies on metadata to define Data Archive transactions.

An archive project contains the following information:

- Source and target repository for the data. If you archive to a database, the source database and the target database must be the same type.
- The project action, which specifies whether to retain the data in the source database after archiving.

   The following table describes the project actions:

Action	Description
Archive Only	Data is extracted from the source repository and loaded into the target repository.
Purge Only	Data is extracted from the source repository and removed from the source repository.
Archive and Purge	Data is extracted from the source repository, loaded into the target repository, and data is removed from the source repository.

- Specification of entities from the source data and imposition of relevant retention policies defined earlier.
- Configuration of different phases in the archival process and reporting information.

To view archive projects, select **Workbench** > **Manage Archive Projects**. To create an archive project, click **New Archive Project** in the **Manage Archive Projects** page.

The following page appears:
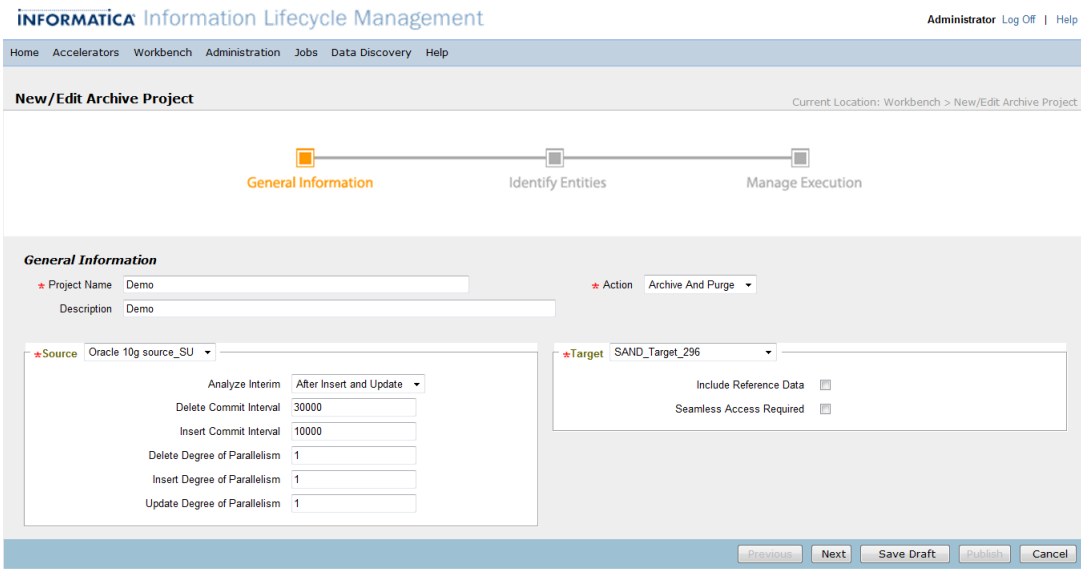
Specifying Project Basics

As a first step to archiving, the information to be specified is described in the following tables.

General Information

Field	Description
Project Name	A name for a Data Archive project is mandatory.
Description	A description is desirable.
Action	A mandatory value for Data Archive Action can be Archive Only, Purge Only, and Archive and Purge.

Source

Field	Description
Connection name	Source connection that you want to archive from. This field is mandatory.
Analyze Interim	Oracle sources only. Determines when the interim table is analyzed for table structure and data insertion.
	- After Insert. Analyzer runs after the interim table is populated.
	- After Insert and Update. Analyzer runs after the interim table is populated and updated.
	- After Update. Analyzer runs after the interim table is updated.
	- None. Analyzer does not run at all.
	Default is After Insert and Update. To optimize performance for best results, use the default value.

Field	Description
Delete Commit Interval	Number of rows per thread that the archive job deletes before the job issues a commit to the database.
	Default is 30,000. Performance is typically optimized within the range of 10,000 to 30,000. Performance may be impacted if you change the default value. The value you configure depends on several variables such as the number of rows to delete and the table sizes.
	Use the following rules and guidelines when you configure the commit interval:
	- If you configure an Oracle source connection to use Oracle parallel DML for deletion, the archive job creates a staging table that contains the row IDs for each commit interval. The archive job truncates the table and reinserts rows for each commit interval. Populating the staging table multiple times can decrease performance. To increase performance, configure the commit interval to either zero or a very high amount, such as 1,000,000, to generate a single commit.
	- If you pause the archive job when the job deletes from Oracle sources, the commit interval affects the amount of time that the job takes to pause. The job pauses after all threads commit the active transaction.
	- As the commit interval decreases, the performance may decrease. The amount of messages that the job writes to the log increases. The amount of time that the archive job takes to pause decreases.
	- As the commit interval increases, performance may increase. The amount of messages that the job writes to the log decreases. The amount of time that the archive job takes to pause increases.
Insert Commit Interval	Number of database insert transactions after which a commit point should be created. Commit interval is applicable when the archive job uses JDBC for data movement.
	If your source data contains the Decfloat data type, set the value to 1.
Delete Degree of Parallelism	Number of concurrent threads that are available for parallel deletion. The number of rows to delete for each table is equally distributed across the threads.
	Default is 1.
Insert Degree of Parallelism	The number of concurrent threads for parallel insertions.
	When the Oracle Degree of Parallelism for Insert is set to more than '1', the java parallel insert of tables is not used.
Update Degree of Parallelism	The number of concurrent threads for parallel updates.
Reference Data Store Connection	Required when you use the history database as the source and you only want to purge from the history database.
	Connection that contains the original source reference data for the records that you archived to the history database. The source includes reference data that may not exist in the history database. The archive job may need to access the reference data when the job purges from the history database.

Target

When you create an archive project, you choose the target connection and configure how the job archives data to the target. You can archive to the Data Vault or to another database.

Archive to the Data Vault to reduce storage requirements and to access the archived data from reporting tools, such as the Data Discovery portal. Archive to another database to provide seamless data access to the archived data from the original source application.

For archive jobs that include `archive only` and `archive and purge` cycles, the system validates that the connections point to different databases. If you use the same database for the source and target connections, the source data may be duplicated or deleted.

The following table describes properties for Data Vault targets:

Property	Description
Connection Name	List of target connections. The target connection includes details on how the archive job connects to the target database and how the archive job processes data for the target. Choose the Data Vault target that you want to archive data to. This field is mandatory.
Include Reference Data	Select this option if you want the archive project to include transactional and reference data from ERP tables.
Reference Data Store Connection	Required when you use a history database as the source. Connection that contains the original source reference data for the records that you archived to the history database. The source includes reference data that may not exist in the history database. The archive job may need to access the reference data when the job archives or purges from the history database.
Enable 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 create an encryption key. Data Archive stores the encrypted key in the ILM repository as a hidden table attribute. When you run the archive definition, the key is passed to Data Vault as a job parameter and is not stored in Data Vault or any log file. The encrypted key is unique to the archive definition and is generated only once for a definition. If you run an archive definition more than once, the same encryption key is used. If you have enabled data encryption for both the selected target connection and the archive definition, Data Archive uses the archive definition encryption details. If you do not configure data encryption at the definition level, then the job uses the details provided at the target connection level.
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.

The following table describes properties for database targets:

Property	Description
Connection name	List of target connections. The target connection includes details on how the archive job connects to the target database and how the archive job processes data for the target. Choose the database target that you want to archive data to.
Seamless Access Required	Select this option if you want to archive to a database repository at the same time as you archive to the Data Vault server. Allows seamless access to both production and archived data.

Property	Description
Seamless Access Data Repository	Repository that stores the views that are required for enabling seamless data access.
Commit Interval (For Insert)	Number of database insert transactions after which the archive job creates a commit point.
Parallelism Degree (Delete)	Number of concurrent threads for parallel deletions.
Parallelism Degree (Insert)	Number of concurrent threads for parallel insertions.

The available fields depend on the type of archive cycle you select. After you configure the target, you can continue to configure the project details or save the project.

Click **Next** to add entities to the project.

Click **Save Draft** to save a draft of the project. You can edit the project and complete the rest of the configuration at a later time.

Specifying Entities, Retention Policies, Roles, and Report Types

The Entities section allows you to specify entities and related information such as retention policies and roles when the target database is the Data Vault.

Note: If your source database is Oracle, you can archive tables that contain up to 998 columns.

Click **Add Entity**. A popup window displays available entities in the source database for selection.

On selecting the desired entity and clicking Select, a new entity is added to the Data Archive definition as shown in the following figure, based on the Data Archive action selected.

Parameter	Description
Candidate Generation Report Type	Determines the report that the project generates during the candidate generation step. - Summary: The report lists counts of purgeable and non-purgeable records for each entity, and counts of non-purgeable records by exception type. A record marked for archiving is considered as purgeable. An exception occurs when purging a record causes a business rule validation to fail. - Detail: In addition to the summary report content, if records belong to related entities, this report contains information about each exception. - Summary Exclude Non Exceptions: Similar to detail, the report contains detailed information where an exception did not occur. - None: No report is generated.
Enabled	Determines whether to exclude the entity from archiving. Disable to exclude the entity from archiving. Default is enabled.
Policy	When the target is the Data Vault, displays a list of retention policies that you can select for the entity. You can select any retention policy that has general retention. You can select a retention policy with column level retention if the table and column exist within the entity. The retention policy list appends the table and column name to the names of retention policies with column level associations. If you do not select a retention policy, the records in the entity do not expire.

Parameter	Description
Role	Data Vault access role for the archive project if the target is the Data Vault. The access role determines who can view data from Data Discovery searches.

If you assign an access role to an archive or retirement project, then access is restricted by the access role that is assigned to the project. Data Discovery does not enforce access from the access roles that are assigned to the entity within the project. Only users that have the Data Discovery role and the access role that is assigned to the project can access the archive data.

If you do not assign an access role to an archive or retirement project, then access is restricted by the access role that is assigned to the entity in the project. Only users that have the Data Discovery role and the access role that is assigned to the entity can access the archive data. |
| Operators and Values | Depending on the list of values that were specified for columns in tables belonging to the entity, a list of operators is displayed, whose values must be defined. |

Configuring Data Archive Run

The last step in the Data Archive Project is to configure a Data Archive run. The following page appears on your screen:

Internally, during Data Archive job execution, data is not directly copied from data source to data target. Instead, a staging schema and interim (temporary) tables are used to ensure that archivable data and associated table structures are sufficiently validated during Archive and Purge.

The following steps are involved in the process:

- Generate Candidates. Generates interim tables based on entities and constraints specified in the previous step.
- Build Staging. Builds the table structure in the staging schema for archivable data.
- Copy to Staging. Copies the archivable rows from source to the staging tablespace.

- Validate Destination (when project type demands data archive). Validates table structure and data in the target in order to generate DML for modifying table structure and/or adding rows of data.
- Copy to Destination (when project type demands data archive). Copies data to the target. If your target is the Data Vault, this step moves data to the staging directory. To move the data from the staging directory to the Data Vault, run the Data Vault Loader job after you publish the archive project.
- Delete from Source (when project type demands data purge). Deletes data from source.
- Purge Staging. Deletes interim tables from the staging schema.

Gaps in Execution of Steps

After each stage in the archival process, Data Archive provides the user with an ability to pause. Relevant check boxes (for that step) can be selected under the Pause After column.

Generating Row Count Reports

To generate ROWCOUNT reports for certain Steps in the Data Archive job execution, select relevant check boxes under the column Row Count Report.

This report gives information on affected ROWS for all selected stages in the process. ROWCOUNT(s) will be logged for the Steps: Generate Candidates and Copy to Staging for this Data Archive job execution.

Note: Whenever a Row Count Report is requested for the Generate Candidates stage in Data Archive job execution, a Simulation Report will also be generated, with information for table Size, Average Row Size and Disk Savings for that particular Data Archive job execution.

Running Scripts

You can also specify JavaScripts or Procedures to run Before or After a particular stage under the Run Before and Run After columns.

Notification Emails

After selecting a check box under the Notify column, a notification email with relevant status information is sent to the user when the Data Archive process is aborted due to an Error or Termination event.

Order of Execution for Step 5 and 6

Towards the bottom of the page there is a combo box to specify the order of execution for Steps 5 and 6.

One of the following options can be selected:

- Run Step 5 first
- Run Step 6 first
- Run Steps 5 and 6 simultaneously

Note: While adding a Data Source, if the Use Staging option is disabled, then the stages: "Build Staging" and "Copy to Staging" will not be included in the Data Archive job execution.

The Project is saved and the Schedule Job page is displayed, when Publish & Schedule is clicked.

On the other hand, the Project gets listed in the Manage Data Archive Projects page when either Publish or Save Draft is clicked. From a user's perspective, the former indicates that the Project is ready for scheduling, the latter means modifications are still required.

Viewing and Editing Data Archive Projects

On publishing a Data Archive Project, a message indicating successful completion of the same is displayed at the beginning of the Manage Data Archive Projects page. A list of projects created so far follow.

Edit and Delete options are available.

Running the Data Vault Loader Job

If you selected the Data Vault as your target connection in the archive project, you must run the Data Vault Loader job to complete the archive process.

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.

Note: Data Vault supports a maximum of 4093 columns in a table. If the Data Vault Loader tries to load a table with more than 4093 columns, the job fails.

Run the Data Vault Loader Job

Run the Data Vault Loader Job to complete the process of archiving data to the Data Vault.

You will need the archive project ID to run the Data Vault Loader job. To get the archive project ID, click **Workbench** > **Manage Archive Projects**. The archive project ID appears in parenthesis next to the archive project name.

1. Click **Jobs** > **Schedule a Job**.
2. Choose **Standalone Programs** and click **Add Item**.

   The **Select Definitions** window appears with a list of all available jobs.
3. Select **Data Vault Loader**.
4. Click **Select**.

   The Data Vault Loader job parameter appears on the **Schedule a Job** page.
5. Enter the Archive Job ID.

You can locate the archive job ID from the **Monitor Job** page.

6. Click **Schedule**.

Troubleshooting Data Vault Loader

Problems can occur at each step of the Data Vault Loader job. In addition, the Data Vault Loader job may have performance problems. If you encounter a problem, follow the troubleshooting tips in this section.

Collect Row Count Step

I cannot see the row count report.

Verify that you enabled the row count report for the Copy to Destination step in the archive project.

Create Tables Step

The Data Vault loader job fails at the create tables step.

Verify that the tables in each schema have different names. If two tables in different schemas have the same name, the job will fail.

Performance

For retirement projects, the Data Vault Loader process runs for days.

When you build retirement entities, do not use one entity for all of the source tables. If an entity includes a large number of tables, the Data Vault Loader may have memory issues. Create one entity for every 200-300 tables in the source application.

You can use a wizard to automatically generate entities. Or, you can manually create entities if you want the entities to have special groupings, such as by function or size.

When you create entities with smaller amounts of tables, you receive the following benefits:

- The Data Vault Loader job performance increases.
- Reduces staging space requirement as the retirement project uses less staging space.
- You can resolve code page configuration issues earlier.
- You can start to validate data earlier with the Data Validation Option.
- Overall, provides better utilization of resources.

Troubleshooting Data Archive Projects

Extraction fails with various Oracle errors.

The archive job fails due to known limitations of the Oracle JDBC drivers. Or, the archive job succeeds, but produces unexpected results. The following scenarios are example scenarios that occur due to Oracle JDBC driver limitations:

- You receive a `java.sql.SQLException` protocol violation exception.

- You receive an exception that OALL8 is in an inconsistent state.

- When you run an archive cycle for an entity with a Run Procedure step, an exception occurs due to an incomplete import. The job did not import the CLOB column, even though the job log shows that the column was imported successfully. The `applimation.log` includes an exception that an SQL statement is empty or null.

- The connection to an Oracle database fails.

- The Copy to Destination step shows an exception that the year is out of range.

- The select and insert job fails with an ORA-600 exception.

- The residual tables job fails with an ORA-00942 exception that a table or view does not exist.

To resolve the errors, use a different Oracle JDBC driver version. You can find additional Oracle JDBC drivers in the following directory:

`<Data Archive installation directory>/optional`

For more information, see Knowledge Base article 109857.

Extraction fails with out of memory or heap errors.

You may receive an out of memory error if several tables include LOB datatypes.

To resolve the errors, perform the following tasks:

- Lower the JDBC fetch size in the source connection properties. Use a range between 100-500.

- Reduce the number of Java threads in the `conf.properties` file for the informia.maxActiveAMThreads property. A general guideline is to use 2-3 threads per core.

Archive job runs for days.

To improve the archive job performance, perform the following tasks:

- Check for timestamp updates in the `host:port/jsp/tqm.jsp` or the BCP directory.

- Check for updates in table AM_JOB_STEP_THREADS in the ILM repository (AMHOME).

- For UNIX operating systems, run the following UNIX commands to see input, output, and CPU activity:

  - `sar -u 1 100`

  - `iostat-m 1`

  - `top`

  - `ps -fu 'whoami' -H`

- Use multiple web tiers to distribute the extraction.
  To use multiple web tiers, perform the following steps:

  1. Create a copy of the web tier.

  2. Change the port in the `conf.properties` file.

  3. Start the web tier.

When I use multiple web tiers, I receive an error that the archive job did not find any archive folders.

Run the Data Vault Loader and Update Retention jobs on the same web tier where you ran the Create Archive Folder job.

The archive job fails because the process did not complete within the maximum amount of time.

You may receive an error that the process did not complete if the archive job uses IBM DB2 Connect utilities for data movement. When the archive job uses IBM DB2 Connect utilities to load data, the job

waits for a response from the utility for each table that the utility processes. The archive job fails if the job does not receive a response within 3 hours.

Use the IBM DB2 trace files to locate the error.

The archive job fails at the Copy to Staging step with the following error: Abnormal end unit of work condition occurred.

The archive job fails if the source database is IBM DB2 and the data contains the Decfloat data type.

To resolve this issue, perform the following tasks:

1. Go to **Workbench** > **Manage Archive Projects**.
2. Click the edit icon for the archive project.
3. Set the **Insert Commit Interval** value to 1.
4. Complete the remaining steps in the archive project to archive the data.

The archive job fails at the Build Staging or Validate Destination step.

If the source database is Oracle and a table contains more than 998 columns, the archive job fails.

This issue occurs because of an Oracle database limitation on column number. Oracle supports a maximum number of 1000 columns in a table. The archive job adds two extra columns during execution. If the table contains more than 998 columns on the source database, the archive job fails because it cannot add any more columns.

CHAPTER 7

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 Process

Before you retire an SAP application, you must verify the installation prerequisites and complete the setup activities.

To retire an SAP application, perform the following high-level steps:

1. Install and apply the SAP transports. For more information, see the "SAP Application Chapter" of the *Informatica Data Archive Administrator Guide*.
2. Install and configure the SAP Java Connector. For more information, see the "SAP Application Chapter" of the *Informatica Data Archive Administrator Guide*.

3. Copy the `sapjco3.jar` file to the `webapp\web-inf\lib` directory in Data Archive.

4. If the SAP application is installed on a Microsoft SQL Server database, enable the following property in the `conf.properties` file by setting the value to "Y": `informia.sqlServerVarBinaryAsVarchar=Y`

5. Restart Data Archive.

6. Create a customer-defined application version under the SAP application in the Enterprise Data Manager.

7. Import metadata from the SAP application to the customer-defined application version.

8. Optionally, run the SAP smart retirement standalone job. For more information, see the topic "SAP Smart Retirement" in this chapter.

9. Create retirement entities.

   Use the generate retirement entity wizard in the Enterprise Data Manager to automatically generate the retirement entities. When you create the entities, you can specify a prefix or suffix that will help you identify the entities as retirement entities. If you ran the SAP smart retirement job, select "Group by SAP Module" when you configure the retirement entities.

10. Create a source connection.

    For the connection type, choose the database in which the SAP application is installed. Configure the SAP application login properties. You must correctly enter the SAP host name so that the SAP datatypes are correctly mapped to Data Vault datatypes during the retirement process.

11. Create a target connection to the Data Vault.

12. Run the Create Archive Folder job to create archive folders in the Data Vault.

13. If the SAP application is installed on a Microsoft SQL Server database, enable the following property in the `conf.properties` file by setting the value to "Y": `informia.sqlServerVarBinaryAsVarchar=Y`

14. Create and run a retirement project.

    When you create the retirement project, add the retirement entities and the pre-packaged SAP entities. Add the attachment link entities as the last entities to process. Optionally, add the Load External Attachments job to the project if you want to move attachments to the Data Vault.

15. Create constraints for the imported source metadata in the ILM repository.

    Constraints are required for Data Discovery portal searches. You can create constraints manually or use one of the table relationship discovery tools to help identify relationships in the source database.

    The SAP application retirement accelerator includes constraints for some tables in the SAP ERP system. You may need to define additional constraints for tables that are not included in the application accelerator.

16. Copy the entities from the pre-packaged SAP application version to the customer-defined application version.

17. Create entities for Data Discovery portal searches.

    Use the multiple entity creation wizard to automatically create entities based on the defined constraints. For business objects that have attachments, add the attachment tables to the corresponding entities.

    The SAP application retirement accelerator includes entities for some business objects in the SAP ERP system. You may need to create entities for business objects that are not included in the application accelerator.

18. Create stylesheets for Data Discovery portal searches.

    Stylesheets are required to view attachments and to view the XML transformation of SAP encoded data.

19. To access the SAP Archives, verify that your login user is assigned the SAP portal user role.

If you installed the SAP retirement accelerator, you can use pre-packaged data visualization reports to access retired data in the Data Vault. You can access the reports through the SAP Archives under the Data Visualization menu.

20. Optionally, validate the retired data.

    Use the Data Validation Option to validate the data that was archived to the Data Vault. Because SAP applications contain thousands of tables, you may want to validate a subset of the archived data. For example, validate the transactional tables that were most used in the source application.

After you retire the application, use the Data Discovery portal, SAP Archives, or third-party query tools to access the retired data.

SAP Smart Retirement

Before you create retirement entities, you can run the SAP smart retirement job. The job identifies the type of table and whether or not it contains data. This information can help you select which tables you want to retire.

When you run the SAP smart retirement standalone job, you provide the name of a source connection from the list of available connections. Based on the source connection properties, the job identifies the SAP application version as it appears in the Enterprise Data Manager. The job then updates the ILM repository with metadata from the SAP source connection. This metadata includes the type of table and whether or not the table contains data.

Note: The SAP smart retirement job updates the metadata for a single source connection imported in the Enterprise Data Manager. Do not export or import the pre-packaged accelerators entities from one environment to another, for example from a development to a production environment.

If you run the SAP smart retirement job and then generate the retirement entities in the Enterprise Data Manager, the resulting entities are grouped together by a common naming convention. The naming convention identifies the type of tables and whether or not the tables contains data.

When you configure the retirement entities, you can specify a prefix or suffix that will be appended to the entities when the Enterprise Data Manager creates them. The prefix or suffix can help you to identify the entity as a retirement entity. After you generate the retirement entities, the entities appear in the Enterprise Data Manager appended with the specified prefix or suffix. For example, the entity "V2_1CDOC_194" is a change documents entity that has the prefix "V2," which you specified when you created the entities.

The Enterprise Data Manager also groups the entities by the SAP module. The entity name reflects the module name as an abbreviation. For example, the entity "SD_001" is in the Sales and Distribution module, while the entity "FI_04" is in the Finance module. When you configure the entities, you specify the maximum amount of tables in the entity. The Enterprise Data Manager creates the required number of entities and names them according to the module. For example, if you have 1,000 Sales and Distribution tables and you specify a maximum of 200 tables per entity, the Enterprise Data Manager creates five different SD entities.

For tables that do not belong to a standard SAP module, the Enterprise Data Manager uses the following naming conventions when it creates the entities:

Naming Convention	Description
1RFC	RFC tables
1NODT	Tables with no data
1CDOC	Change document tables
1IDOC	IDOC tables
1OFFA	Office and attachment tables
1LOG	Log tables
1WF	Workflow tables
YYNM	Tables with no module assigned in the SAP system

You can use this information to decide whether or not you want to retire the entity when you create the retirement project. If you want to reduce the number of tables in the retirement project, you can exclude certain types of tables, such as workflow tables, or tables that do not contain any data, as designated by the "1NODT" in the naming convention.

The Enterprise Data Manager groups special tables in their own entities. Special tables are identified by the suffix "_SAP."

Tables with archived data are identified by the suffix "_SAP_ADK."

The following image shows an example of the entities created by the SAP smart retirement process in the Enterprise Data Manager. These entities were created with the prefix "UF2":
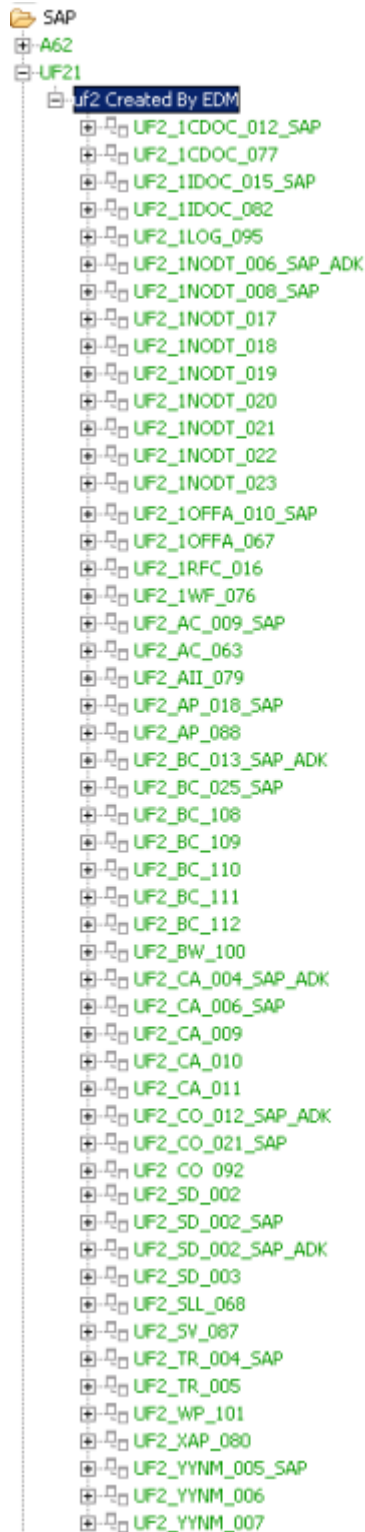
Running the SAP Smart Retirement Standalone Job

Before you generate retirement entities, run the SAP smart retirement standalone job.

1. Click **Jobs** > **Schedule a Job**.

   The **Schedule Job** page appears.

2. Select **Standalone Programs** and then **Add Item**.

   The **Select Definitions** window appears.

3. From the list of programs, select SAP Smart Retirement.

4. Select the source connection for the SAP application version that you want to retire.

5. Click **Submit**.

6. Schedule the job to run.

   You can monitor the job progress from the Monitor Jobs page. When the job completes, you can generate the retirement entities for the SAP application version in the Enterprise Data Manager. For more information about generating retirement entities, see the *Informatica Data Archive Enterprise Data Manager Guide*.

Retirement Job

The retirement job uses a combination of methods to access data from SAP applications. The job accesses data directly from the SAP database for normal physical tables. The job uses the SAP Java Connector to log in to the SAP application to access data in transparent HR and STXL tables, ADK files, and attachments.

When you run a retirement job, the job uses the source connection configured in the retirement project to connect to the source database and to log in to the SAP system. The job extracts data for the following objects:

Physical Tables

Physical tables store data in a readable format from the database layer. The retirement job extracts data directly from the database and creates BCP files. The job does not log in to the SAP application.

Transparent HR and STXL tables

Transparent HR PCL1-PCL5 and STXL tables store data that is only readable in the SAP application layer. The retirement job logs in to the SAP system. The job cannot read the data from the database because the data is encoded in an SAP format.

The job uses the source connection properties to connect to and log in to the SAP application. The job uses an ABAP import command to read the SAP encoded data. The job transforms the SAP encoded data into XML format and creates BCP files. The XML transformation is required to read the data after the data is retired. The XML transforms occurs during the data extraction.

ADK Files

ADK files store archived SAP data. The data is only readable in the SAP application layer. The retirement job uses the source connection properties to connect to and log in to the SAP application. The job calls a function module from the Archive Development Kit. The function module reads and extracts the archived ADK data and moves the data to the BCP staging area as compressed BCP files.

Attachments

Attachments store data that is only readable in the SAP application layer. The retirement job uses the SAP Java Connector to log in to the SAP application and to call a function module. The function module reads, decodes, and extracts the attachments that are stored in the database. The job moves the attachments to the attachment staging area that you specify in the source connection. Depending on your configuration, the job may also move attachments that are stored in an external file system.

When the retirement job creates BCP files, the job compresses and moves the files to a BCP staging area. For physical tables, the job stores the BCP files in the staging directory that is configured in the Data Vault target connection. For transparent HR and STXL tables, ADK files, and attachments, the job uses an SAP function module to generate the BCP files. The SAP function module stores the BCP files in the staging directory that is configured in the source connection.

The job uses the BCP file separator properties in the `conf.properties` file to determine the row and column separators. The job uses the Data Vault target connection to determine the maximum amount of rows that the job inserts into the BCP files. After the job creates the BCP files, the Data Vault Loader moves the files from the BCP staging area to the Data Vault.

Attachment Retirement

You can archive CRM, ERP, SCM, SRM, and GOS (Generic Object Services) attachments. The attachments can exist in the SAP database and in an external file system. Data Archive retires any objects that are attached to the business object, such as notes, URLs and files.

By default, the retirement job archives all attachments that are stored in the SAP database and in an external file system or storage system. If attachments are in a storage system, then archive link must be available. The job creates a folder in the staging location for each business object ID that includes attachments. The folder stores all attachments for the business object ID. You specify the staging location in the source connection. The job uses the following naming convention to create the folders:

```
<SAP client number>_<business object ID>
```

The job downloads all attachments for the business objects to the corresponding business object ID folder. The job downloads the attachments in the original file format and uses a gzip utility to compress the attachments. The attachments are stored in .gz format.

For attachments that are stored in the SAP database, the retirement job uses an ABAP import command to decode the attachments before the job downloads the attachments. The SAP database stores attachments in an encoded SAP format that are only readable in the SAP application layer. The job uses the SAP application layer to decode the attachments so you can read the attachments after you retire the SAP application.

For attachments that are stored in a file system, the job downloads the attachments from the external file system or storage to the staging location. Optionally, you can configure the attachment entity to keep attachments in the original location.

After the job creates the folder for the business object ID, the job appends the business object number as a prefix to the attachment name. The prefix ensures that attachment has a unique identifier and can be associated with the correct business object.

The job uses the following naming convention for the downloaded attachment names:

```
<business object number>_<attachment name>.<file type>.gz
```

After you run the retirement job, you choose the final storage destination for the attachments. You can keep the attachments in the staging file system, move the attachments to another file system, or move the

attachments to the Data Vault. There is no further compression of the attachments regardless of the final storage destination.

Attachment Storage

Attachment storage options depend on the location of the attachments in the source application. You can keep the attachments in the original location or move the attachments to a different location. It is mandatory to move attachments that are stored in the SAP database to a different storage destination. It is optional to move attachments that are stored in an external file system.

You can archive attachments to any of the following destinations:

Original File System

You can keep attachments in the original storage for attachments that are stored in an external file system. By default, the retirement job moves all attachments to the file system that is configured in the source connection.

To keep the attachments in the original file system, configure the run procedure parameter for the attachment entity.

Staging File System

By default, the retirement job moves attachments that are stored in the SAP database and in an external file system to a staging file system. The staging file system is configured in the source connection.

After the retirement job moves attachments to the staging file system, you can keep the attachments in this location or you can move the attachments to another file system, such as Enterprise Vault, or move the attachments to the Data Vault. Store attachments in a file system to optimize attachment performance when users view attachments.

Data Vault

After the retirement job archives attachments to the staging file system, you can move the attachments to the Data Vault.

Note: The retirement job compresses the attachments. The Data Vault Service does not add significant compression. In addition, users may experience slow performance when viewing attachments. To optimize attachment performance, store attachments in a file system instead.

To move attachments to the Data Vault, run the Load External Attachments job after the retirement job completes. The Load External Attachments job moves the attachments from the staging file system to the Data Vault.

Attachment Viewing

After you archive the attachments, you can view the attachments in the Data Discovery portal or with a third-party reporting tool. The ILM repository link table, ZINFA_ATTCH_LINK, includes information on the downloaded attachments. The table includes information to link business objects to the corresponding attachments. The link table is included in the pre-packaged SAP entities for business objects that have attachments.

To view attachments in the Data Discovery portal, you must configure entities for discovery. For any business object that may have attachments, you must add the attachment tables to the corresponding entity. If you store the attachments in a file system, add the ZINFA_ATTCH_LINK table to the entities. If you store the attachments in the Data Vault, add the ZINFA_ATTCH_LINK and AM_ATTACHMENTS tables to the entities.

To view attachments from the Data Discovery portal, you must create a stylesheet. In the stylesheet, build a hyperlink to the attachment location. Use the ZINFA_ATTCH_LINK table to build the path and file name. When users perform searches, business object attachments appear as hyperlinks. When users click a hyperlink, the

default browser opens the attachment. The attachments are compressed in gzip format. The client from which the user runs the Data Discovery portal searches must have a program that can open gzip files, such as 7zip.

Data Visualization Reports

The SAP application accelerator includes data visualization report templates. The reports include information that corresponds to key information in SAP ERP transactions. You can run the reports to access retired data in the Data Vault.

If you use the **Reports and Dashboards** window to access the reports, you must first copy the report templates to the folder that corresponds to the archive folder in the Data Vault target connection. The archive folder is equivalent to a Data Vault database. When you copy the reports, the system updates the report schemas based on the target connection configuration. Then you can run the reports.

If you use the SAP Archives to access the reports, all of the reports installed by the accelerator are copied to the corresponding SAP archive folder the first time that you launch the SAP Archives. If you plan to access the reports through the SAP Archives, you do not need to manually copy the reports to the archive folder, though you have the option to do so.

When you install the SAP application retirement accelerator, the installer publishes the reports to the data visualization server with the following properties:

Property	Value
Folder name	SAP
Catalog name	SAP_Retirement
Catalog Path	<Data Archive install directory>/webapp/visualization/reporthome/<system generated folders>/ SAP_Retirement.cat

Related Topics:

- "Copying Reports" on page 162

SAP Archives

If you have installed the SAP retirement accelerator, you can access the data visualization reports through the SAP Archives.

The SAP Archives contain all of the reports installed by the SAP application retirement accelerator. After you install the accelerator and launch the SAP Archives for the first time, Data Archive copies the report templates from the default folder to the archive folder where the application is retired.

If you have retired the SAP application to one archive folder, you can view all of the available reports when you launch the SAP Archives. If you have retired the SAP tables to more than one archive folder, you are given the option to select an archive folder when you launch the SAP Archives.

To access the SAP Archives, you must have the SAP portal user role. The SAP portal user role only grants access to the portal. To run, copy, or delete reports, you must have the appropriate report permissions. For more information about assigning system-defined roles, see the "Security" chapter in the *Informatica Data Archive Administrator Guide*. For more information about Data Visualization report permissions, see the "Data Visualization" chapter in this guide.

Running Reports in the SAP Archives

To run reports, launch the SAP Archives from the Data Visualization menu.

1. Click **Data Visualization** > **SAP Archives**.
2. If required, select the archive folder where the tables that you want to run a report on are retired.

   If you retired the SAP application to only one archive folder, you are not required to select a folder.

   The list of available reports appears.
3. To run a report, click the name of the report.
4. Enter the search criteria.
5. Click **Submit**.

Troubleshooting SAP Application Retirement

SAP application retirement messages contain message IDs that can help in troubleshooting or reporting issues that occur during the retirement process. This section explains message IDs and how to troubleshoot and resolve issues.

Message Format

Messages that are related to SAP application retirement convey information about a task that completed successfully, an error that occurred, a warning, or status updates.

Each message starts with a message ID. The message ID is a unique mix of numbers and letters that identify the type of message, type of component, and message number.

The following image shows a sample message ID:
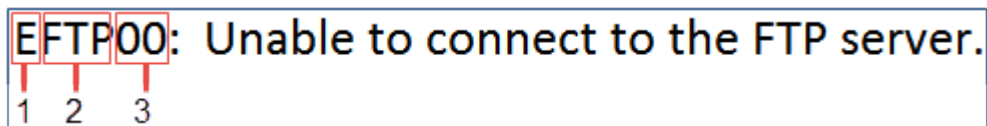
1. Message type
2. Component code
3. Message number

The message ID has the following parts:

Message type

    The first character in the message ID represents the type of message.

    The message types are represented by the following letters:

- E - Error
- I - Information
- S - Success
- W - Warning

Component code

The second, third, and fourth characters in the message ID represent the component code of the component that triggered the message.

The components are represented by the following codes:

- ADK - refers to the archive session.
- ATT - refers to attachments.
- FTP - refers to the FTP process.
- HR - refers to HR tables and attachments.
- JOB - refers to the Data Archive job.
- SPL - refers to special tables.
- STG - refers to the Staging folder.
- TXT - refers to data in text files.

Message number

The last two numbers in the message ID represent the message number.

Frequently Asked Questions

The following list explains how to troubleshoot and resolve issues that might occur during the SAP application retirement process.

When I run the retirement job, the job immediately completes with an error.

The job log includes the following message:

```
java.lang.NoClassDefFoundError: com/sap/mw/jco/IMetaData
```

The error occurs because the `sapjaco.jar` file is missing in the ILM application server. Verify that the SAP Java Connector installation is correct.

Perform the following steps to verify the SAP Java Connector installation:

1. From the command line, navigate to the following directory:

   ```
   <Data Archive installation>/webapp/WEB-INF/lib
   ```
2. Enter the following command:

   ```
   java -jar sapjco.jar
   ```

   If the installation is correct, you see the path to the SAP Java Connector library file, such as `C:\SAPJCO\sapjcorfc.dll`.

   If the installation is not correct, you see a Library Not Found message in the path for to the SAP Java Connector library file.

If the installation is not correct, perform the following steps to resolve the error:

1. Verify that you installed the SAP Java Connector.
2. Copy the `sapjaco.jar` file from the root directory of the SAP Java Connector installation and paste to the following directory:

   ```
   <Data Archive installation>/webapp/WEB-INF/lib
   ```
3. Restart the ILM application server.

When I run the retirement job, I receive an error that the maximum value was exceeded for no results from the SAP system.

For every package of rows that the job processes, the job waits for a response from the SAP system. The job uses the response time configured in the `conf.properties` file and the package size specified in the job parameters. If the SAP system does not respond within the configured maximum response time, the job fails.

The SAP system may not respond for one or more of the following reasons:

- The maximum amount of SAP memory was exceeded. The SAP system terminated the process. The SAP system may run out of memory because the package size that is configured in the job is too high.
- The SAP system takes a long time to read the data from the database because of the configured package size.
- The network latency is high.

To resolve the error, in the SAP system, use transaction `SM50` to verify if the process is still running. If the process is still running, then restart the job in Data Archive.

If the process is not running, perform one or more of the following tasks:

- Use SAP transactions `SM37`, `ST22`, and `SM21` to analyze the logs.
- Decrease the package size in the job parameters.
- Increase the informia.MaxNoResultsIteration parameter in the `conf.properties` file.

The job log indicates that the Data Vault Loader received an exception due to an early end of buffer and the load was aborted.

The error may occur because the SAP application is hosted on Windows.

Perform the following steps to resolve the error:

1. Configure the informia.isSAPOnWindows property to `Y` in the `conf.properties` file.
2. Reboot the ILM application server.
3. Create a new retirement project.

CHAPTER 8

Creating Retirement Archive Projects

This chapter includes the following topics:

Creating Retirement Archive Projects Overview

A retirement archive project contains the following information:

- Identifying the application to retire.
- Defining the source and target for the retirement data and imposition of relevant policies defined earlier.
- Optionally, enabling integrated validation before you run the project. Integrated validation ensures the integrity of the data that you retire.
- Configuration of different phases in the retirement process, and reporting information.

Existing retirement archive projects can be viewed from Workbench > Manage Retirement Projects. A new project can be created by clicking the New Application button from the Manage Retirement Projects page. The process of creating a retirement archive project is described in detail in the following section.

Identifying Retirement Application

As a first step to application retirement, the information to be specified is described in the following table:

Application Detail	Description
Application Name	A name for an application is mandatory.
Application Vendor Name	Vendor name of the selected application.
Application Owner	Application owner of the legacy application selected.
Vendor Contact	Contact information of the vendor.
Retirement Date	A mandatory field to enter proposed retirement date.
Vendor Phone	Vendor contact numbers.
Primary Contact	Vendor's primary contact information.
Vendor Contract Number	The signed contract number of the vendor.
Comments	A section where additional information related to this particular retirement can be added.
The Add More Attributes button	A facility to add any more attributes that you may want to include to the application details which would serve in the process of application retirement.
Use Integrated Validation	Enables integrated validation for the project.
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 create an encryption key. Data Archive stores the encrypted key in the ILM repository as a hidden table attribute in case the job fails and must be resumed. When you run the retirement job, the key is passed to Data Vault as a job parameter and is not stored in Data Vault or any log file. If the retirement job is successful, the key is deleted from the ILM repository. The encrypted key is unique to the retirement definition and is generated only once for a definition.
	If you have enabled data encryption for both the selected target connection and the retirement definition, Data Archive uses the retirement definition encryption details. If you do not configure data encryption at the definition level, then the job uses the details provided at the target connection level.
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.

Clicking the Next button advances to the section on defining source and target repositories. The Save Draft button saves the project for future modifications and scheduling.

Defining a Source Connection

Create a source connection for the source database that you want to retire. The source database is the database that stores the transactional data that you want to archive. For example, the source can be an ERP application such as PeopleSoft. The source database is usually the production database or application database.

Define a connection to the application that you want to retire. Select the application version. Then, create or choose a source connection.

Important:

- If the source connection is a Microsoft SQL Server, clear the **Compile ILM Functions** check box on the **Create or Edit an Archive Source** page before you start a retirement job. This ensures that the staging database user has read-only access and will not be able to modify the source application.

- Clear the **Use Staging** check box on the **Create or Edit an Archive Source** page before you start a retirement job.

Defining a Target Repository

Define a target connection, which is the location where you want to archive data to. For a retirement project, the target is an optimized file archive, also called the Data Vault.

Select the application version. Then, create or choose a target connection.

Adding Entities

Add the entities from the application that you want to retire.

Defining Retention and Access

This section allows you to define the retention policies and associated access policies for an application.

The following table describes the fields in the Retention group:

Field	Description
Policy	Retention policy that determines the retention period for records in the entity. You can create a retention policy or select a policy. You can select any retention policy that has general retention. If you do not create or select a retention policy, the records in the entity do not expire.
Name	Name of the retention policy. Required if you create a retention policy.
Description	Optional description of the retention policy.
Retention Period	The retention period in years or months. Data Archive uses the retention period to determine the expiration date for records in the entity. If you select or create a policy with general retention, the expiration date equals the retirement archive job date plus the retention period.

The following table describes the fields in the Access Roles group:

Field	Description
Access Roles	Data Vault access role for the retirement archive project. You can create an access role or select a role. The access role determines who can view data from Data Discovery searches.
Role Name	Specific responsibility of a user who initiates a retention policy.
Description	Any intuitive description on the access roles.
Valid From	Date when a role is assigned to a user.
Valid Until	Optional date when a role assigned to a user expires.

When you click the **Publish and Schedule** button, the project is saved and the **Schedule Job** page appears.

When you click the **Publish** button or the **Save Draft** button, Data Archive lists the project in the **Manage Retirement Projects** page. Publishing the project indicates that the project is ready for scheduling. Saving a draft means that modifications are still required.

Clicking the **Next** button opens the **Review and Approve** page. The **Save Draft** button saves the project for future modifications and scheduling.

Review and Approve

The Review and Approve page in the Data Archive for Application Retirement lets the user not only completely evaluate the configured source and target repositories, but also facilitates viewing and approving the selected application's retrial.

The following table lists possible available fields:

Field	Possible Values
Application Name	Name of the legacy application which is being retired.
Retired Date	The Date the application retirement is scheduled.
Retired Object Name	Application module that is being retired/ selected by user.
Retention	Retention period specified by user.
Access	Role which is been provided while associating a policy.
Approved By	Person who approves the retirement procedure.
Approved Date	Date when retirement procedure is been approved.

Click the Application Info tab to view the detailed information of the selected application and click the Source & Target tab to view the configured Source and Target details. Detailed information is displayed by clicking each of the three sections.

Confirming Retirement Run

The last step in an Retirement project is to confirm a retirement run. Internally, during a retirement, data is not directly retired from data source to data target. Instead a staging schema and interim (temporary) tables are used to ensure that retirement data and associated table structures are sufficiently validated during the application retirement.

The following steps are involved in the process:

1. Generate Candidates. Generates Interim tables based on Entities and constraints specified in the previous step.

2. Validate Destination. Validates table structure and Data in the Target repository to generate DML for modifying Table structure and/or adding Rows of data.

3. Copy to Destination. Copies data to Data Destination.

4. Purge Staging. Deletes Interim tables from staging schema.

After the retirement project completes successfully, you cannot edit or re-run the project.

Gaps in Execution of Steps

After each stage in the retirement process, Data Archive for Application Retirement provides the user with an ability to pause. Relevant checkboxes (for that step) can be selected under the Pause After column.

Generating Row Count Reports

To generate ROWCOUNT reports for certain steps in the retirement cycle, select relevant checkboxes under the column Row Count Report.

This report gives information on affected rows for all selected stages in the process. For example, ROWCOUNTs are logged for the steps Generate Candidates and Copy to Target for this archive job execution.

Running Scripts

One can also specify JavaScripts or Procedures to run before or after a particular stage under the Run Before and Run After columns.

Notification Emails

On selecting a check box under the Notify column, a notification email (with relevant status information) is sent to the user when the archive process is aborted due to an error or termination event.

When either Publish or Save Draft is clicked, the project gets listed in the Manage Retirement Projects page. From a user's perspective, the former indicates that the project is ready for scheduling. The latter means modifications are still required.

Copying a Retirement Project

After a retirement project successfully completes, you cannot edit or rerun the project. If you need to edit or rerun the project, you can create a copy of the completed project and edit the copied project.

1. Click **Workbench** > **Manage Retirement Projects**.

   The **Manage Retirement Projects** window opens.

2. Select the **Copy** check box next to the retirement project that you want to copy.

The **Create or Edit a Retirement Project** window opens. Data Archive pre-populates all of the project properties as they existed in the original project.

3. To edit the retirement project, change any properties and click **Next** to continue through the project.

4. When you have made the required changes to the project, schedule the project to run or save the project to run at a later date.

Integrated Validation Overview

When you retire an application to the Data Vault, you can enable integrated validation to ensure the integrity of the data that you retired. Integrated validation helps you to identify data that might have been deleted, truncated, or otherwise changed during the retirement process.

You can enable integrated validation when you create the retirement project, before you schedule it to run. Before the retirement job copies the tables to Data Vault, the integrated validation process uses an algorithm to calculate a checksum value for each row and column in the entity. After the retirement job copies the tables to Data Vault, the validation process calculates another checksum value for the retired rows and columns in Data Vault. The process then compares the original checksum value of each row and column, to the checksum value of each corresponding retired row and column.

When the comparison is complete, you can review the details of the validation. If a deviation occurs in the number of rows, the value of a row, or the value of a column, between the original checksum value and the retired checksum value, you can review the details of the deviation. You can select a status for each deviated table in the retirement project. You must provide a justification for the status that you select.

After you review the validation, you can generate a validation report that includes the details of each deviation. The validation report also contains a record of any status that you selected when you reviewed the deviation, and any justifications that you entered.

Integrated validation lengthens the retirement process by 40% to 60%.

You can enable integrated validation for retirement projects with the following types of source connections:

- Oracle
- Microsoft DB2 for Linux, UNIX, and Windows
- Microsoft DB2 for z/OS
- Microsoft SQL Server
- IBM System i
- Informatica PowerExchange

Integrated Validation Process

1. When you define the application for a retirement project, enable the Integrated Validation option. You must enable the option to run the validation with the retirement job. You must also enable the option if you want to run the validation independent of the retirement process, at a later time. You cannot use integrated validation on a retirement project that has already been run.

2. Complete the remainder of the steps required to run a retirement project, and then schedule the project to run. For more information, see the [Retiring an Application to Data Vault with On-Premise Data Archive](#) H2L.

3. After the retirement project completes, review the details of the validation. For each deviated table, you can select a status of accepted or rejected. If you select rejected for a deviated table, the job remains in a paused state. You can also place a deviated table on hold. When you select a status for a deviated table, you must provide a justification for the status.

4. After you review the validation, you can optionally generate a validation report. The validation report contains the details of each deviation, in addition to any status that you selected and justification that you entered.

Row Checksum

During the copy to destination step of the retirement job, the validation process calculates a checksum value for each row in every table in the entity.

The process writes the row checksum values to a BCP file that appends the values to the end of each row. The process appends the checksum values as a metadata column that is copied to the Data Vault along with the table. When the Data Vault loader is finished, Data Vault calculates a second checksum value for each row in the retired tables.

The validation process then compares the checksum value written to the BCP file to the checksum value calculated by Data Vault. When the checksum comparison is complete, review the validation for any deviations that occur between the two checksum values for each row.

Column Checksum

During the copy to destination step, the validation process calculates a checksum value for each column in every table in the retirement entity.

The validation process stores the column checksum values in the ILM repository. When the Data Vault loader is finished, Data Vault calculates another checksum value for each column and compares it to the checksum value stored in the ILM repository. You can then review the validation for any deviations that occur between the two checksum values for each column.

Validation Review

Review the validation process for details of any deviations.

When the retirement job and checksum comparison is complete, you can review the validation process for details of any deviations that occur. You can select a status for any deviated table and provide a justification for the status. You can also edit a status or justification to change it, for example from rejected to accepted.

If you select a status for a deviated table, it does not change the data in either the source database or the Data Vault. Statuses and justifications are a method of reviewing and commenting on deviations that occur.

To review the validation, you must have the Operator role. To select a status for a deviated table, you must have the Administrator role. For more information on system-defined roles, see the *Informatica Data Archive Administrator Guide*.

Note: For any validated table, if a difference exists between the number of rows in the source and target, you cannot review the table for deviations.

Validation Review User Interface

You can review the validation from the Monitor Jobs page after the retirement job is complete.

The following image shows the **Validation Review** window:
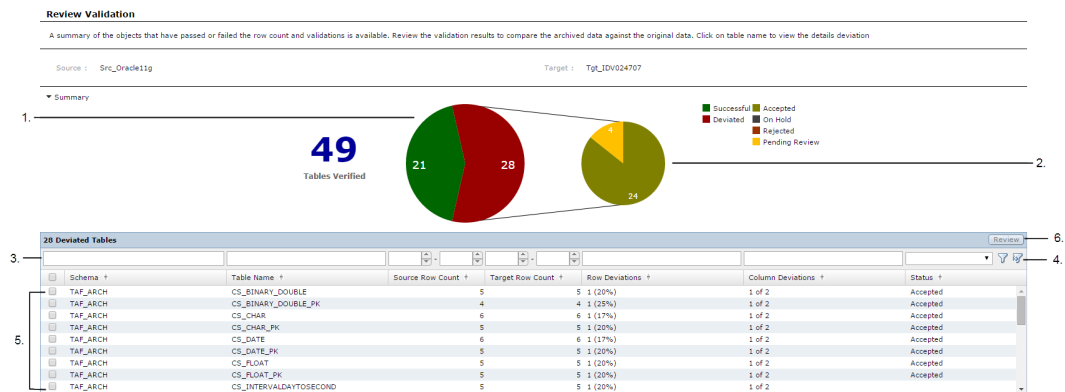
1. Pie chart that shows the number of tables in the retirement project that passed validation, in addition to the number of tables that failed validation.

2. Pie chart that shows the specific actions that you took on the deviated tables. Tables that you have not yet taken an action on appear as "pending review."

3. Filter box. Enter the entire filter or part of the filter value in the filter box above the appropriate column heading, and then click the filter icon.

4. Filter icon and clear filter icon. After you specify filter criteria in the fields above the column names, click the filter icon to display the deviated tables. To clear any filter settings, click the clear filter icon.

5. List of deviated tables. Select the check box on the left side of the row that contains the table you want to review.

6. Review button. After you have selected a table to review, click the review button.

The following image shows the **Review** window:
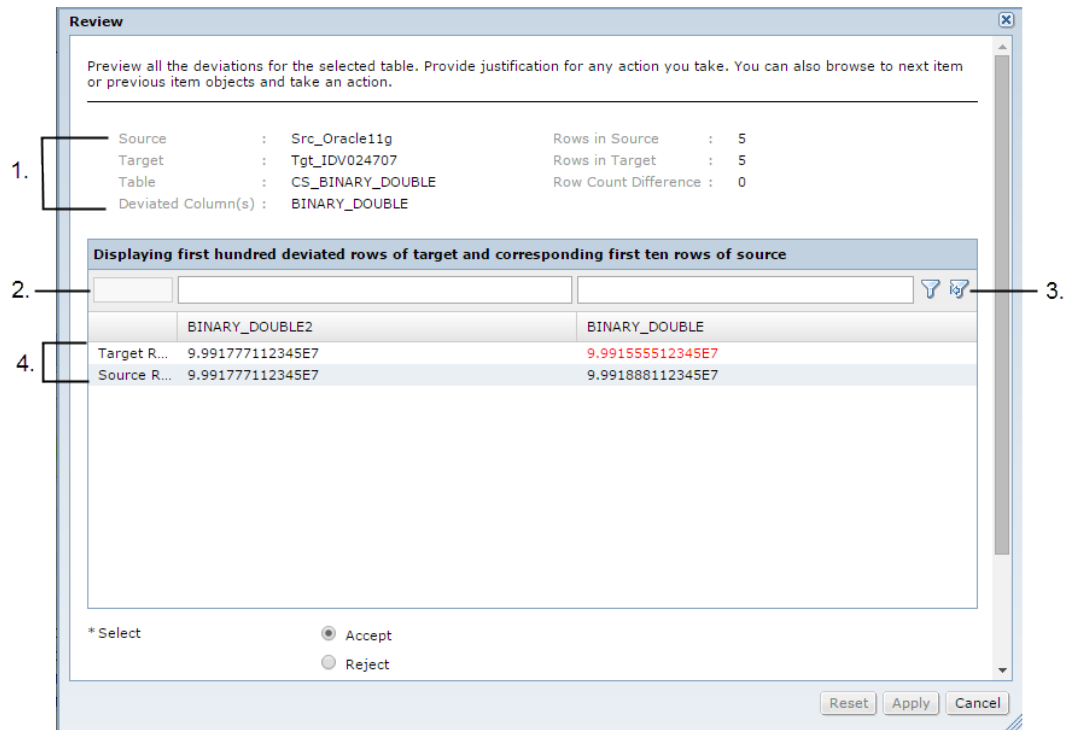
1. Validation details, such as the source connection name, target connection name, and table name.

2. Filter box. Enter the entire filter or part of the filter value in the filter box above the appropriate column heading, and then click the filter icon.

3. Filter icon and clear filter icon. After you specify filter criteria in the fields above the column names, click the filter icon to display the deviated tables. To clear any filter settings, click the clear filter icon.

4. List of deviated rows.

5. Radio buttons that assign a status to the table. Select **Accept**, **Reject**, or **Hold for Clarifications**.

For a deviated table, the **Review** window lists a maximum of the first 100 deviated rows in Data Vault and the first 10 corresponding rows in the source database. For each deviated table, the **Review** window displays a maximum of 10 columns, with a maximum of eight deviated columns and two non-deviated columns.

If all of the columns in a table are deviated, source connection rows are not displayed.

Validation Report Elements

The validation report contains the following elements:

Summary of results

The summary of results section displays the application name, retirement job ID, source database name, and Data Vault archive folder name. The summary of results also displays the login name of the user who ran the retirement job, and the date and time that the user performed the validation. If the retirement job contained any deviations, the summary of results lists the review result as "failed." If the retirement job did not contain any deviated tables, the review result appears as "passed."

Summary of deviations

The summary of deviations section displays pie charts that illustrate the number of deviated tables in the retirement job and the type of status that you selected. The section also contains a list of the

deviated tables, along with details about the tables, for example the row count in the source database and in Data Vault.

Detail deviations

The detail deviations section displays details about the deviations that occur in each table included in the retirement job. For example, the detail deviations section displays the number of row deviations, number of column deviations, and names of the deviated columns on each table. The section also displays the number of rows in both the source database and Data Vault, along with any row count difference.

Appendix

The appendix section contains a list of the tables in the retirement entity that the validation process verified.

Enabling Integrated Validation

You can enable integrated validation when you create the retirement project.

1. Click **Workbench** > **Manage Retirement Projects**.

2. Click **New Retirement Project**.

   The *Create or Edit a Retirement Project* window appears.

3. Enter the application details and select the check box next to **Use Integrated Validation**.

4. Complete the remainder of the steps required to run a retirement project, then schedule the project to run. For more information, see the Retiring an Application to Data Vault with On-Premise Data Archive H2L.

5. Click **Jobs** > **Monitor Jobs** to monitor the retirement job status and review the validation after the process is complete.

Reviewing the Validation

If a "Warning" status appears next to the integrated validation step of the retirement process, at least one deviation has occurred. If no deviations have occurred, the job step displays the "Completed" status. You can review the validation from the **Monitor Jobs** page.

1. Click **Jobs** > **Monitor Jobs**.
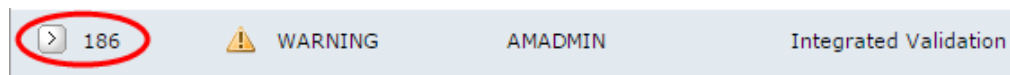
2. Click the arrow next to the integrated validation job step of the retirement project.

   The job step details appear.

3. To expand the job details, click the arrow next to the integrated validation program name.

4. Click the **Review Validation** link.

The *Review Validation* window opens.

5. To review the deviations for each table, click the check box to the left of the table name and then click **Review**.

The **Review** window appears.
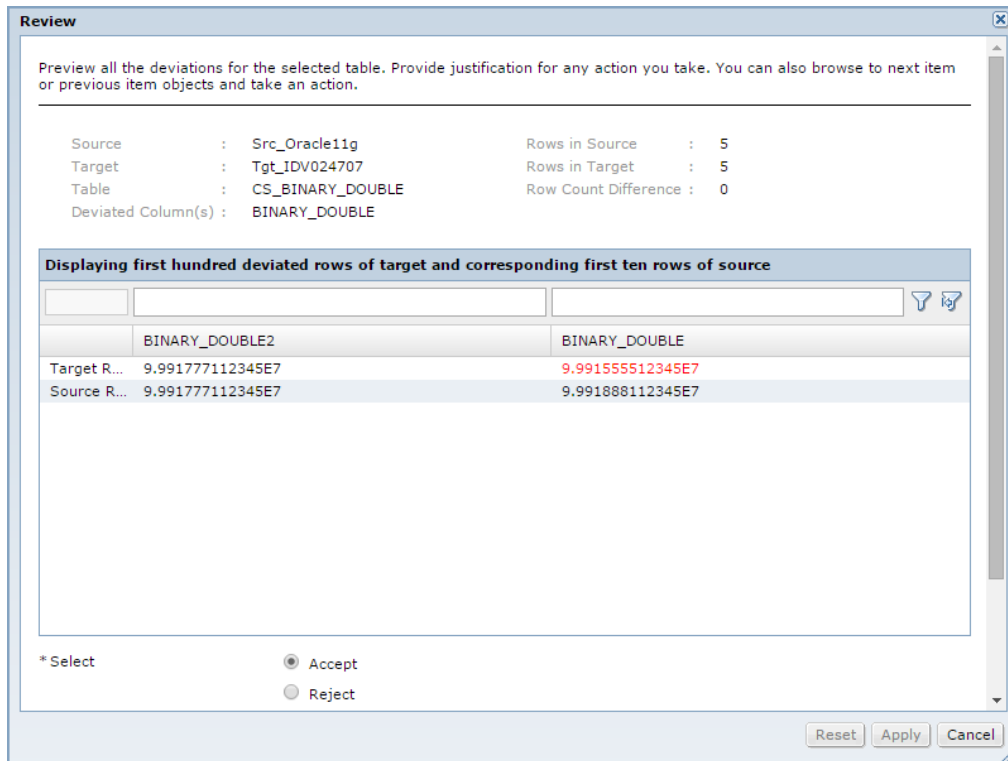
6. To select a status for the deviated table, select the **Accept**, **Reject**, or **Hold for Clarification** option.

7. Enter a justification for the action in the text box.

8. Click **Apply**.

   Note: You can select more than one table at a time and click **Review** to select the same status for all of the selected tables. However, if you select multiple tables you cannot review the deviations for each table.

Running the Validation Report

You can run the validation report from the **Monitor Jobs** page.

1. Click **Jobs** > **Monitor Jobs**.
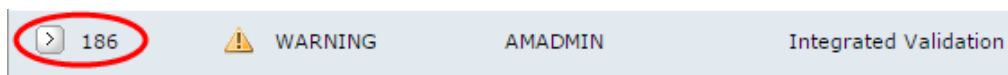
2. Click the arrow next to the integrated validation job step of the retirement project.

The job step details appear.

3. To expand the job details, click the arrow next to the integrated validation program name.

4. Click the **View Validation Report** link.

The validation report opens as a PDF file. You can print or save the validation report.

Integrated Validation Standalone Job

If you enable the integrated validation option when you create the retirement project, you can remove the integrated validation job from the retirement job flow and run the validation at a later date.

To run the integrated validation independent of the retirement flow, delete the integrated validation job step from the retirement project flow before you schedule the retirement project to run. Before you run the job, ensure that the source connection is online. Then, when you are ready to run the validation, run the integrated validation standalone job.

You can run the integrated validation job only once on a retirement definition.

Running the Integrated Validation Standalone Job

To run the integrated validation standalone job, select the retirement definition and schedule the job.

1. Click **Jobs** > **Schedule a Job**.
2. Select **Standalone Programs** and then click **Add Item**.
3. Select the **Integrated Validation** job and click **Select**.
4. Click the list of values button to select the name of the retirement definition. Only retirement projects that have not already run the integrated validation are displayed.
5. Schedule the job to run.

Troubleshooting Retirement Archive Projects

The archive job fails at the Copy to Staging step with the following error: Abnormal end unit of work condition occurred.

The archive job fails if the source database is IBM DB2 and the data contains the Decfloat data type.

To resolve this issue, perform the following tasks:

1. Go to **Workbench** > **Manage Retirement Projects**.
2. Click the edit icon for the retirement project.
3. Set the **Insert Commit Interval** value to 1.
4. Complete the remaining steps in the retirement project to retire the data.

CHAPTER 9

Retention Management

This chapter includes the following topics:

Retention Management Overview

Retention management refers to the process of storing records in the Data Vault and deleting records from the Data Vault. The retention management process allows you to create retention policies for records. You can create retention policies for records as part of a data archive project or a retirement archive project.

A retention policy is a set of rules that determine the retention period for records in an entity. The retention period specifies how long the records must remain in the Data Vault before you can delete them. The retention period can be definite or indefinite. Records with definite retention periods have expiration dates. You can run the Purge Expired Records job to remove expired records from the Data Vault. Records with indefinite retention periods do not expire. You cannot delete them from the Data Vault.

Your organization might require that you enforce different retention rules for different records. For example, you might need to retain insurance policy records for five years after the insurance policy termination date. If messages exist against the insurance policy, you must retain records for five years after the most recent message date in different tables.

Data Archive allows you to create and implement different retention rules for different records. You can create a retention policy to addresses all of these needs.

Retention Management Process

You can create and change retention policies if you have the retention administrator role. You can create retention policies in the workbench and when you create a retirement archive project.

If you create retention policies in the workbench, you can use them in data archive projects that archive data to the Data Vault. You can also use the policies in retirement archive projects. If you create retention policies when you create a retirement archive project, Data Archive saves the policies so that you can use them in other archive projects

After you create retention policies, you can use them in archive projects. When you run the archive job, Data Archive assigns the retention policies to entities and updates the expiration dates for records. You can change the retention periods for archived records through Data Discovery. Change the retention period when a corporate retention policy changes, or when you need to update the expiration date for a subset of records. For example, an archive job sets the expiration date for records in the ORDERS table to five years after the order date. If an order record originates from a particular customer, the record must expire five years after the order date or the last shipment date, whichever is greatest.

To use an entity in a retention policy, ensure that the following conditions are met:

- The entity must have a driving table.
- The driving table must have at least one physical or logical primary key defined.

To create and apply retention policies, you might complete the following tasks:

1. Create one or more retention policies in the workbench that include a retention period.
2. If you want to base the retention period for records on a column date or an expression, associate the policies to entities.
3. Create an archive project and select the entities and retention policies. You can associate one retention policy to each entity.
4. Run the archive job.

   Data Archive assigns the retention policies to entities in the archive project. It sets the expiration date for all records in the project.
5. Optionally, change the retention policy for specific archived records through Data Discovery and run the Update Retention Policy job.

   Data Archive updates the expiration date for the archived records.
6. Run the Purge Expired Records job.

   Data Archive purges expired records from the Data Vault.

Note: If you have the retention viewer role, you cannot create, edit, or assign retention policies. You can perform Data Discovery searches based on retention policy and retention expiration date, and you can view the policy details and the table data. For more information on the differences between the retention administrator and the retention viewer roles, see the *Data Archive Administrator Guide*.

Retention Management Example

You can use Data Archive to enforce different retention rules for different records.

For example, a large insurance organization acquires a small automobile insurance organization. The small organization uses custom ERP applications to monitor policies and claims. The large organization requires

that the small organization retire the custom applications and use its ERP applications instead. After acquisition, records managers use Data Archive to create retirement projects for the custom applications.

The large organization must retain records from the small organization according to the following rules:

- By default, the organization must retain records for five years after the insurance policy termination date.
- If messages exist against the insurance policy, the organization must retain records for five years after the most recent message date in the policy, messages, or claims table.
- If the organization made a bodily injury payment, it must retain records for 10 years after the policy termination date or the claim resolution date, whichever is greatest.
- If a record contains a message and a bodily injury payment, the organization must retain the record according to the bodily injury rule.

The small organization stores policy information in the following tables in the AUTO_POLICIES entity:

POLICY Table

Column	Description
POLICY_NO	Automobile insurance policy number (primary key column)
TERM_DATE	Policy termination date
LAST_TRANS	Date of last message for the policy
…	…

MESSAGES Table

Column	Description
MSG_ID	Message ID (primary key column)
POLICY_NO	Policy number related to the message (foreign key column)
LAST_TRANS_DATE	Message date
…	…

CLAIMS Table

Column	Description
CLAIM_NO	Claim number (primary key column)
POLICY_NO	Policy number related to the claim (foreign key column)
LAST_DATE	Date of last message for the claim
RESOLVE_DATE	Claim resolution date
CLAIM_TYPE	Claim type: property (1) or bodily injury (2)

Column	Description
CLAIM_AMT	Claim damage amount
…	…

As a retention administrator for the merged organization, you must create and apply these rules. To create and apply the retention management rules, you complete multiple tasks.

Task 1. Create Retention Policies in the Workbench

Create retention policies that you can select when you create the retirement project. Create two retention policies. One retention policy has a five year retention period and the other retention policy has a 10 year retention period.

1. Select **Workbench** > **Manage Retention Policies**.

    The **Manage Retention Policies** window appears.

2. Click **New Retention Policy**.

    The **New/Edit Retention Policy** window appears.

3. Enter the following information:

Policy Name	5-Year Policy
Retention Period	5 years

4. Click **Save**.

    Data Archive saves retention policy "5-Year Policy." You can view the retention policy in the **Manage Retention Policies** window.

5.   Repeat steps 2 through 4 and create the following retention policy:

Policy Name	10-Year Policy
Retention Period	10 years

Data Archive saves retention policy "10-Year Policy." You can view the retention policy in the **Manage Retention Policies** window.

Task 2. Associate the Policies to Entities

The default retention rule for insurance policies specifies that the organization must retain records for five years after the insurance policy termination date. To implement this rule, you must associate 5-Year Policy with the column that contains the insurance policy termination date.

1.   Select **Workbench** > **Manage Retention Policies**.

The **Manage Retention Policies** window appears.

2.   Click **Associate to Entity** next to retention policy "5-Year Policy."

The **Associate Retention Policy to Entity** window appears.

3.   Click **Add Entity**.

4.   Select the custom application and the AUTO_POLICIES entity.

5.   Select the entity table "POLICY" and column "TERM_DATE."

6.   Click **Save**.

Data Archive creates a column level association for 5-Year Policy. To indicate that column level retention is associated with 5-Year Policy, Data Archive appends the table and column name to the policy name in selection lists.

Task 3. Create the Retirement Archive Project

When you create the retirement archive project, select 5-Year Policy for the AUTO_POLICIES entity.

1.   Select **Workbench** > **Manage Retirement Projects**.

The **Manage Retirement Projects** window appears.

2.   Click **New Retirement Project**.

The **New/Edit Retirement Project** wizard appears.

3.   Enter the custom application name and the retirement date, and then click **Next**.

The wizard prompts you for source and target information.

4.   Select the application version, the source repository, and the target repository, and then click **Next**.

The wizard prompts you for the entities to include in the project.

5. Click **Add Entity**.

   The **Entity** window appears.

6. Select the AUTO_POLICIES entity and click **Select.**

   The wizard lists the entity that you select.

7. Click **Next**.

   The wizard prompts you for a retention policy and an access role for the entity.

8. Select retention policy "5-Year Policy POLICY TERM_DATE" for the AUTO_POLICIES entity.

   The wizard displays details about the retention policy. The details include the column and table on which to base the retention period.

9. Click **Next**.

   The wizard prompts you for approval information for the retirement project.

10. Optionally, enter approval information for the project, and then click **Next**.

    The wizard displays the steps in the retirement archive job.

11. Click **Publish**.

    Data Archive saves the retirement project. You can view the project in the **Manage Retirement Projects** window.

Task 4. Run the Archive Job

Run the retirement archive job to set the expiration date for all records in the AUTO_POLICIES entity. When the job completes, Data Archive sets the expiration date for all records in the entity to five years after the date in the POLICY.TERM_DATE column.

1. Select **Jobs > Schedule a Job**.

   The **Schedule Job** window appears.

2. In the **Projects/Programs to Run** area, select **Projects**, and then click **Add Item**.

   The **Program** window appears.

3. Select the retirement archive job, and then click **Select**.

   Data Archive adds the retirement job to the list of scheduled projects.

4. In the **Schedule** area, select **Immediately.**

5. Click **Schedule**.

   Data Archive starts the retirement job.

Task 5. Change the Retention Policy for Specific Archived Records

The retirement archive job sets the expiration date for all records in the AUTO_POLICIES entity to five years after the insurance policy termination date. To implement the message and the bodily injury rules, you must change the retention policy for records that have messages or bodily injury payments.

1. Select **Data Discovery** > **Manage Retention**.

    The **Manage Retention** window appears.

2. Click the **Entities** tab.

3. Click **Actions** next to entity "AUTO_POLICIES."

    The **Associate Retention Policy to Entity** window appears.

4. Click **Modify Assigned Retention Policy** next to retention policy "5-Year Policy."

    The **Modify Assigned Retention Policy** window appears. It displays the archive folder, existing retention policy, and entity name.

5. To select records with messages, enter the following condition:

    `MSG_ID Not Null`

6. In the **New Retention Policy** list, select new retention policy "5-Year Policy."

7. In the **Base Retention Period On** area, click **Add Row** and add a row for each of the following columns:

Entity Table	Retention Column
POLICY	LAST_TRANS
MESSAGES	LAST_TRANS_DATE
CLAIMS	LAST_DATE

8. In the **Comments** field, enter comments about the retention policy changes.

9. To view the records affected by your changes, click **View**.

    A list of records that the retention policy changes affect appears. Data Archive lists the current expiration date for each record.

10. To submit the changes, click **Submit**.

    The **Schedule Job** window appears.

11. Schedule the Update Retention Policy job to run immediately.

12. Click **Schedule**.

    Data Archive runs the Update Retention Policy job. For records with messages, it sets the expiration date to five years after the last transaction date across the POLICY, MESSAGES, and CLAIMS tables.

13. Repeat steps 1 through 4 so that you can implement the bodily injury rule.

    The **Modify Assigned Retention Policy** window appears.

14. To select records with bodily injury payments, click **Add Row** and add a row for each of the following conditions:

    - `CLAIM_TYPE Equals 2 AND`
    - `CLAIM_AMT Greater Than 0`

15. In the **New Retention Policy** list, select new retention policy "10-Year Policy."

16. In the **Base Retention Period On** area, click **Add Row** and add a row for each of the following columns:

Entity Table	Retention Column
POLICY	TERM_DATE
CLAIMS	RESOLVE_DATE

17. Repeat steps 8 through 12 to update the expiration date for records with bodily injury payments.

    Data Archive sets the expiration date to 10 years after the policy termination date or the claim resolution date, whichever is greatest.

Note: If a record has a message and a bodily injury payment, Data Archive sets the expiration date according to the bodily injury rule. It uses the bodily injury rule because you ran the Update Retention Policy job for this rule after you ran the Update Retention Policy job for the message rule.

Task 6. Run the Purge Expired Records Job

To delete expired records from the Data Vault, run the Purge Expired Records job.

1. Select **Jobs** > **Schedule a Job**.

   The **Schedule Job** window appears.

2. In the **Projects/Programs to Run** area, select **Standalone Programs** and click **Add Item**.

   The **Program** window appears.

3. Select **Purge Expired Records** and click **Select**.

   The **PURGE_EXPIRED_RECORDS** area appears.

4. Enter the Purge Expired Records job parameters.

5. In the **Schedule** area, select **Immediately**.

6. Optionally, enter notification information.

7. Click **Schedule**.

   Data Archive purges expired records from the Data Vault.

Retention Policies

Retention policies consist of retention periods and optional entity associations that determine which date Data Archive uses to calculate the expiration date for records. You can create different retention policies to address different retention management needs.

Each retention policy must include a retention period in months or years. Data Archive uses the retention period to calculate the expiration date for records to which the policy applies. A record expires on its expiration date at 12:00:00 a.m. The time is local to the machine that hosts the Data Vault Service. The retention period can be indefinite, which means that the records do not expire from the Data Vault.

After you create a retention policy, you can edit it or delete it. You can edit or delete a retention policy if the retention policy is not assigned to an archived record or to an archive job that is running or completed. Data Archive assigns retention policies to records when you run the archive job.

To change the retention period for a subset of records after you run the archive job, you can change the retention period for the records through Data Discovery.

Entity Association

When you create a retention policy, you can associate it to one or more entities. Entity association identifies the records to which the retention policy applies and specifies the rules that Data Archive uses to calculate the expiration date for a record.

When you associate a retention policy to an entity, the policy applies to all records in the entity, but not to reference tables. For example, you associate a 10 year retention policy with records in the EMPLOYEE entity so that employee records expire 10 years after the employee termination date. Each employee record references the DEPT table in another entity. The retention policy does not apply to records in the DEPT table.

For entities with unrelated tables, you can only use an absolute-date based retention policy.

Entities used in a retention policy must have a driving table with at least one physical or logical primary key defined.

When you associate a retention policy to an entity, you can configure the following types of rules that Data Archive uses to calculate expiration dates:

General Retention

Bases the retention period for records on the archive job date. The expiration date for each record in the entity equals the retention period plus the archive job date. For example, records in an entity expire five years after the archive job date.

Column Level Retention

Bases the retention period for records on a column in an entity table. The expiration date for each record in the entity equals the retention period plus the date in a column. For example, records in the EMPLOYEE entity expire 10 years after the date stored in the EMP.TERMINATION_DATE column.

Expression-Based Retention

Bases the retention period for records on a date value that an expression returns. The expiration date for each record in the entity equals the retention period plus the date value from the expression. For example, records in the CUSTOMER table expire five years after the last order date, which is stored as an integer.

General Retention

Configure general retention when you want to base the retention period for records in an entity on the data archive job date or on the retirement archive job date. The expiration date for each record in the entity equals the retention period plus the archive job date.

When you configure general retention, you can select an indefinite retention period. When you select an indefinite retention period, the records do not expire from the Data Vault.

When you configure general retention, all records in the entity to which you associate the policy have the same expiration date. For example, you create a retention policy with a five year retention period. In a retirement archive project, you select this policy for an entity. You run the archive job on January 1, 2011. Data Archive sets the expiration date for the archived records in the entity to January 1, 2016.

You can configure general retention when you select entities in a data archive project or a retirement archive project. Select the entity, and then select the retention policy.

Column Level Retention

Configure column level retention when you want to base the retention period for records in an entity on a date column in an entity table. The expiration date for each record in the entity equals the retention period plus the date in the column.

When you configure column level retention, each record in an entity has a unique expiration date. For example, you create a retention policy with a 10 year retention period and associate it to entity EMPLOYEE, table EMP, and column TERM_DATE. In a data archive project, you select this policy for the EMPLOYEE entity. Each record in the EMPLOYEE entity expires 10 years after the date in the EMP.TERM_DATE column.

You can configure column level retention in the workbench when you manage retention policies or when you change retention policies for archived records in Data Discovery. You can also configure column level retention when you select entities in a data archive project or a retirement archive project. Select the entity, and then select the retention policy. The retention policy list appends the table and column name to the names of retention policies with column level associations. You cannot configure column level retention for retention policies with indefinite retention periods.

Expression-Based Retention

Configure expression-based retention when you want to base the retention period for records in an entity on an expression date. The expiration date for each record in the entity equals the retention period plus the date that the expression returns.

The expression must include statements that evaluate to a column date or to a single date value in DATE datatype. It can include all SQL statements that the Data Vault Service supports. The expression can include up to 4000 characters. If the expression does not return a date or if the expression syntax is not valid, the Update Retention Policy job fails. To prevent the Update Retention Policy job from failing, provide null-handling functions such as IFNULL or COALESCE in your expressions. This ensures that a suitable default value will be applied to records when the expression evaluates to a null or empty value.

The Update Retention Policy job includes the expression in the select statement that the job generates. The job selects the records from the Data Vault. It uses the date that the expression returns to update the retention expiration date for the selected records. When you monitor the Update Retention Policy job status, you can view the expression in the job parameters and in the job log. The job parameter displays up to 250 characters of the expression. The job log displays the full expression.

Use expressions to evaluate any of the following types of table data:

Dates in integer format

Use the TO_DATE function to evaluate columns that store dates as integers. The function converts the integer datatype format to the date datatype format so that the Update Retention Policy job can calculate the retention period based on the converted date. The TO_DATE function selects from the entity driving table.

To select a column from another table in the entity, use the full SQL select statement. For example, you want to set the retention policy to expire records 10 years after the employee termination date. The entity driving table, EMPLOYEE, includes the TERM_DATE column which stores the termination date as an integer, for example, 13052006. Create a retention policy, and set the retention period to 10 years. Add the following expression to the retention policy:

```
TO_DATE(TERM_DATE,'ddmmyyyy')
```

When the Update Retention Policy job runs, the job converts 13052006 to 05-13-2006 and sets the expiration date to 05-13-2016.

Dates or other values in tags

Add tags to archived records to capture dates or other values that might not be stored or available in the Data Vault. Use an expression to evaluate the tag and set the retention period based on the tag value.

For example, you retired a source application that contained products and manufacturing dates. You want to set the expiration date for product records to 10 years after the last manufacturing date. When you retired the application, some products were still in production and did not have a last manufactured date. To add a date to records that were archived without a last manufacturing date, you can add a tag column with dates. Then, define a retention policy and use an expression to evaluate the date tag.

For more information on how to use a tag to set the retention period, see "Using a Tag Column to Set the Retention Period " on page 107.

Data from all tables in the Data Vault archive folder

Use an expression to evaluate table columns across all entities in the Data Vault archive folder. You can use a simple SQL statement to evaluate data from one column, or you can use complex SQL statements to evaluate data from multiple columns.

For example, you retired an application that contained car insurance policies. An insurance policy might have related messages and claims. You want to set the expiration date for insurance policy records to five years after the latest transaction date from the POLICY, MESSAGES, or CLAIMS tables. If the insurance policy has a medical or property damage claim over $100,000, you want to set the expiration date to 10 years.

Note: If you enter a complex SQL statement that evaluates column dates across high-volume tables, you might be able to increase query performance by changing the retention policy for specific records. Change the retention policy for records and then run the Update Retention Policy job instead of entering an expression to evaluate date columns across tables. Data Archive can generate SQL queries which run more quickly than user-entered queries that perform unions of large tables and contain complex grouping.

You can configure expression-based retention in the workbench when you manage retention policies or when you change retention policies for archived records in Data Discovery. The retention policy list in Data Discovery appends "(Expression)" to the names of retention policies with expression-based associations. You cannot configure expression-based retention for retention policies with indefinite retention periods.

Using a Tag Column to Set the Retention Period

You can set the retention period for records based on the value in a tag column. You use a tag column if you archived records without a value, such as a date, that you can base the retention period on.

The following task flow explains how to use a tag column to set an expression-based retention period.

1. Add a tag column to the table that holds the transaction key.

    a. Specify the Date data type for the tag.

    b. Enter a date for the tag value.

    For more information about adding a tag, see "Adding Tags" on page 142.

2. Determine the database name of the tag.

    a. Select **Data Discovery > Browse Data**.

    b. Move all the `ilm_metafield#` columns from the **Available Columns** list to the **Display Columns** list.

    c. Click **Search**.

    d. Browse the results to determine the name of the tag column that contains the value you entered. For example, the name of the tag column is `ilm_metafield1`.

3. Create a retention policy.

    For more information about creating a retention policy, see "Creating Retention Policies" on page 109.

4. Associate the retention policy to the entity.

    a. Select the entity table with the date tag.

    b. Enter an expression such as the following sample expression in the **Expression** dialog box:

        ```
        IFNULL(<name of the date column>, IFNULL(<name of the date tag column>, TODAY()))
        ```

        For example, you set the retention policy to 10 years. You want to base the retention period on the last manufacturing date. If a record does not have the last manufacturing date, you want to base your retention period on the tag date. If a record does not have a tag date, you want to use today's date. Use the following expression:

        ```
        IFNULL(LAST_MANUFACTURING_DATE, IFNULL(ILM_METAFIELD1,TODAY()))
        ```

        In this example, the retention expiration date is set to 10 years after the last manufacturing date. If the record has a null value instead of a date in the last manufacturing date field, then the retention expiration date is set to 10 years after the date in the date tag column. If the record has a null value in the date tag column, the retention expiration date is set to 10 years after today's date.

    For more information about associating a retention policy to an entity, see "Associating Retention Policies to Entities" on page 109.

Retention Policy Properties

Enter retention policy properties when you create a retention policy in the workbench.

The following table describes the retention policy properties:

Property	Description
Policy name	Name of the retention policy. Enter a name of 1 through 30 characters. The name cannot contain the following special characters: , < >
Description	Optional description for the retention policy. The description cannot contain the following special characters: , < >
Retention period	The retention period in months or years. Enter a retention period or enable the **Retain indefinitely** option.
Retain indefinitely	If enabled, the records to which the retention property applies do not expire from the Data Vault. Enter a retention period or enable the **Retain indefinitely** option.
Records management ID	Optional string you can enter to associate the retention policy with the policy ID from a records management system. You can use the records management ID to filter the list of retention policies in Data Discovery. You can enter any special character except the greater than character (>), the less than character (<), or the comma character (,).

Creating Retention Policies

Create retention policies from the workbench. You can also create retention policies when you create a retirement archive project.

1. Select **Workbench** > **Manage Retention Policies**.

   The **Manage Retention Policies** window appears.

2. Click **New Retention Policy**.

   The **New/Edit Retention Policy** window appears.

3. Enter the retention policy properties.

4. Click **Save**.

   Data Archive lists the retention policy in the **Manage Retention Policies** window.

Associating Retention Policies to Entities

Associate retention policies to entities to configure column-level or expression-based retention. You can associate a retention policy to multiple entities, columns, and expressions. You cannot associate a retention policy with an indefinite retention period to an entity.

Note: If you select column-level retention, you must select a table in the entity and a column that contains date values. If you select expression-based retention, you must select a table in the entity and enter an expression.

1. Select **Workbench** > **Manage Retention Policies**.

   The **Manage Retention Policies** window appears.

2. Click **Associate to Entity** next to the retention policy you want to associate.

   The **Associate Retention Policy to Entity** window appears.

3. Click **Add Entity**.

4. Select an application from the **Application Module** list of values.

5. Select the entity from the **Entity Name** list of values.

6. To configure column-level retention, you must select an entity table and a column that contains date values.

   a. Select a table from the **Entity Table** list of values.

   b. Select a column that contains date values from the **Retention Column** list of values.

      The Retention Column list displays date columns that are not null.

   c. Click **Save**.

   Repeat step 6 to associate the retention policy with another column.

7. To configure expression-based retention, you must select an entity table and enter an expression.

   a. Select a table from the **Entity Table** list of values.

   b. Click the icon in the **Expression** column.

   c. Enter the expression in the **Expression** dialog box.

      The **Expression** icon turns green to indicate that expression-based retention is configured for the entity association.

   d. Click **Save**.

   Repeat step 7 to associate the retention policy with another expression.

If the expression needs to evaluate a tag value, you must use the name of the tag column as it appears in the database. The following steps shows how you can determine the name of the tag column:

a. Select **Data Discovery** > **Browse Data**.

b. Select the archive folder, schema, and table.

c. Select all the tag columns, `ilm_metafield1` to `ilm_metafield9`, from the Data Columns section and move them to the Display Columns section using the **>** arrow. Click **Search**.

d. In the Results section, note the name of the tag column that contains the appropriate tag values.

Retention Policy Changes

You can change the retention policy assigned to specific records in an entity after you run a data archive job or a retirement archive job. Change the retention policy when you want to update the expiration date for any record in the entity.

For example, you assign a five-year retention policy to automobile insurance policies in a retirement archive job so that the records expire five years after the policy termination date. However, if messages exist against the insurance policy, you must retain records for five years after the most recent message date in the policy, claims, or client message table.

You change the assigned retention policy through Data Discovery. You cannot edit or delete retention policies in the workbench if the policies are assigned to archived records or to archive jobs that are running or completed.

Before you change a retention policy for any record, you might want to view the records assigned to the archive job. When you view records, you can check the current expiration date for each record.

To change the retention policy for records, you must select the records you want to update and define the new retention period. After you change the retention policy, schedule the Update Retention Policy job to update the expiration date for the records in the Data Vault.

Record Selection for Retention Policy Changes

When you change the retention policy assigned to records in an entity, you must select the records you want to update. Select the records to update when you modify the assigned retention policy in Data Discovery.

To select records to update, you must specify the conditions to which the new retention policy applies. Enter each condition in the format "<Column><Operator><Value>." The **Is Null** and **Not Null** operators do not use the <Value> field.

For example, to select insurance records in an entity that have messages, you might enter the following condition:

```
MSG_ID Not Null
```

You can specify multiple conditions and group them logically. For example, you want to apply a retention policy to records that meet the following criteria:

```
JOB_ID > 0 AND (CLAIM_TYPE = 1 OR CLAIM_NO > 1000 OR COMPANY_ID = 1007)
```

Enter the following conditions:

```
    JOB_ID Greater Than 0 AND
    ( CLAIM_TYPE Equals 1 OR
```

```
CLAIM_NO Greater Than 1000 OR

COMPANY_ID Equals 1007 )
```

Use the following guidelines when you specify the conditions for record selection:

Enclose string and varchar values in single quotes.

For example, to select insurance records with policy numbers that begin with "AU-," enter the following condition:

```
POLICY_NO Starts With 'AU-'
```

Specify case-sensitivity for string comparisons.

By default, Data Archive performs case-sensitive string comparisons. To ignore case, enable the **Case Insensitive** option.

You can specify that a condition is within a range of values.

Use the **In** operator to specify that a condition is within a range of comma-separated values, for example, `CLAIM_CLASS IN BD1, BD2`. Do not enclose the list of values in parentheses.

You can specify multiple conditions.

Use the **Add Row** button to add additional conditions.

You can group conditions.

If you specify multiple conditions, you can group them with multiple levels of nesting, for example, "(A AND (B OR C)) OR D." Use the **Add Parenthesis** and **Remove Parenthesis** buttons to add and remove opening and closing parentheses. The number of opening parentheses must equal the number of closing parentheses.

Retention Period Definition for Retention Policy Changes

When you change the retention policy assigned to records in an entity, you must define a new retention period for the records you want to update. Define the new retention period when you modify the assigned retention policy in Data Discovery.

To define the new retention period, select the new retention policy and enter rules that determine which date Data Archive uses to calculate the expiration date for records. The expiration date for records equals the retention period plus the date value that is determined by the rules you enter.

You can enter the following types of rules:

Relative Date Rule

Bases the retention period for records on a fixed date. For example, you select a retention policy that has a 10 year retention period. You enter January 1, 2011, as the relative date. The new expiration date is January 1, 2021.

Column Level Rule

Bases the retention period for records on a column date. For example, you select a retention policy that has a five year retention period. You base the retention period on the CLAIM_DATE column. Records expire five years after the claim date.

Expression-Based Rule

Bases the retention period for records on the date returned by an expression. For example, you select a retention policy that has a five year retention period. You want to base the retention period on dates in the CLAIM_DATE column, but the column contains integer values. You enter the expression `TO_DATE(CLAIM_DATE,'ddmmyyyy')`. Records expire five years after the claim date.

The rules you enter must evaluate to a single date value or to a column date. If you enter multiple rules, you cannot select a relative date. You must select the maximum date or the minimum date across all rules.

For example, you have a five-year retention policy for insurance records that retires records five years after the last transaction date. You want to select the last transaction date from three tables and apply the new retention policy using that date.

The following entity tables contain the last transaction date:

- Policy.LAST_TRANS
- Messages.LAST_TRANS_DATE
- Claims.LAST_DATE

To select the last transaction date across tables, complete the following tasks:

1. Select the five-year retention policy as the new retention policy.
2. In the **Base Retention Period On** area, add a row for each column:

Entity Table	Retention Column
Policy	LAST_TRANS
Messages	LAST_TRANS_DATE
Claims	LAST_DATE

3. In the **By evaluating the** list, select **Maximum**.

Update Retention Policy Job

The Update Retention Policy job updates the expiration date for records in the Data Vault. Schedule the Update Retention Policy job after you change the retention policy for specific records in Data Discovery.

The Update Retention Policy job selects records from the Data Vault based on the conditions you specify when you change the retention policy. It uses the date value determined by the rules you enter to update the retention expiration date for the selected records. If you change the retention policy for a record using an expression-based date, the Update Retention Policy job includes the expression in the SELECT statement that the job generates.

When you run the Update Retention Policy job, you can monitor the job status in the **Monitor Job** window. You can view the WHERE clause that Data Archive uses to select the records to update. You can also view the old and new retention policies. If you change the retention policy for a record using an expression-based date, you can view the expression in the job parameters and in the job log. The job parameter displays up to 250 characters of the expression. The job log displays the full expression.

When you monitor the Update Retention Policy job status, you can also generate the Retention Modification Report to display the old and new expiration date for each affected record. The report is created with the Arial Unicode MS font. To generate the report, you must have the font file ARIALUNI.TTF saved in the `<Data Archive installation>\webapp\WEB-INF\classes` directory. To view the report, you must have the retention administrator role and permission to view the entities that contain the updated records.

To schedule the Update Retention Policy job, click **Submit** in the **Modify Assigned Retention Policy** window. The **Schedule Job** window appears. You schedule the job from this window.

Viewing Records Assigned to an Archive Job

After you run an archive job, you can view the archived records. You can view all records associated with the archive job. Data Archive displays the current expiration date for each record.

1. Select **Data Discovery** > **Manage Retention**.

   The **Manage Retention** window appears.

2. Click the **Retention Policies** tab or the **Entities** tab.

3. Click **Actions** next to the retention policy name or the entity name.

   The **Assigned Entities** window or the **Assigned Retention Policy** window appears.

4. Click the open button (**>**) next to the entity name or the retention policy name.

   The list of jobs associated with the retention policy and the entity appears.

5. To view archive criteria for the job, click **Archive Criteria** next to the job name.

   The archive criteria window lists the archive criteria for the job. Click **Close** to close the window.

6. To view records assigned to the job, click **View Records** next to the job name.

   Data Archive displays the list of records with the current expiration dates.

Changing Retention Policies for Archived Records

After you run an archive job, you can change the retention policy for any archived record. Change the assigned retention policy through Data Discovery. You cannot change the assigned retention policy in the workbench if the policy is assigned to archived records or to archive jobs that are running or completed.

1. Select **Data Discovery** > **Manage Retention**.

   The **Manage Retention** window appears.

2. Click the **Retention Policies** tab or the **Entities** tab.

3. Click **Actions** next to the retention policy name or the entity name.

   The **Assigned Entities** window or the **Assigned Retention Policy** window appears.

4. To change the retention policy for records that a specific job archived, click the open button (**>**) next to the entity name or the retention policy name.

   The list of jobs associated with the retention policy and the entity appears.

5. Click **Modify Retention Policy** next to the retention policy, entity, or job name.

   The **Modify Assigned Retention Policy** window appears. It displays the archive folder, existing retention policy, and entity name.

6. To select the records that you want to update, enter the conditions to which the new retention policy applies.

7. Select the new retention policy and enter the rules that define the new retention period:

   • To add a relative date rule, select a date from the **Relative to** field.

   • To add a column level rule, add a row in the **Base Retention Period On** area, and select the table and column.

   • To add an expression-based rule, add a row in the **Base Retention Period On** area. Click the **Expression** icon in the row, enter the expression in the **Expression** dialog box, and click **Save**. The **Expression** icon turns green to indicate that an expression is configured.

8. If you enter multiple rules, select the maximum date or minimum date in the **By evaluating the** list.

9. Enter comments about the retention policy changes.

10. To view the records affected by your changes, click **View**.

    A list of records that the retention policy changes affect appears. Data Archive lists the current expiration date for each record.

11. To submit the changes, click **Submit**.

    The **Schedule Job** window appears.

12. Schedule the Update Retention Policy job and select the notification options.

13. Click **Schedule**.

    Data Archive starts the job according to the schedule.

14. To view the Retention Expiration report, click **Jobs** > **Monitor Jobs**, select the Update Retention Policy job, and click **Retention Expiration Report**.

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.

When you run the Purge Expired Records job, by default, the job searches all entities in the Data Vault archive folder for records that are eligible for purge. To narrow the scope of the search, you can specify a single entity. The Purge Expired Records job searches for records in the specified entity alone, which potentially decreases the amount of time in the search stage of the purge process.

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, select a date for Data Archive to base the report on. If you select a past date or the current date, Data Archive generates a list of tables with the number of records that are eligible for purge on that date. You can pause the job to review the report and then schedule the job to purge the records. If you resume the job and continue to the delete step, the job deletes all expired records up to the purge expiration date that you used to generate the report. If you provide a future date, Data Archive generates a list of tables with the number of records that will be eligible for purge by the future date. The job stops after it generates the report.

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.

The reports are created with the Arial Unicode MS font. To generate the reports, you must have the font file ARIALUNI.TTF saved in the `<Data Archive installation>\webapp\WEB-INF\classes` directory.

To purge records, you must enable the purge step through the **Purge Deleted Records** parameter. You must also provide the name of the person who authorized the purge.

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.

Purge Expired Records Job Parameters

The following table describes the parameters for the Purge Expired Records job:

Parameter	Description	Required or Optional
Archive Store	The name of the archive folder in the Data Vault that contains the records that you want to delete. Select a folder from the list.	Required.
Purge Expiry Date	The date that Data Archive uses to generate a list of records that are or will be eligible for delete. Select a past, the current, or a future date. If you select a past date or the current date, Data Archive generates a report with a list of all records eligible for delete on the selected date. You can pause the job to review the report and then schedule the job to purge the records. If you resume the job and continue to the delete step, the job deletes all expired records up to the selected date. If you select a future date, Data Archive generates a report with a list of records that will be eligible for delete by the selected date. However, Data Archive does not give you the option to delete the records.	Required.
Report Type	The type of report to generate when the job starts. Select one of the following options: - **Detail**. Generates the Retention Expiration Detail report. - **None**. Does not generate a report. - **Summary**. Generates the Retention Expiration Summary report.	Required.
Pause After Report	Determines whether the job pauses after Data Archive creates the report. If you pause the job, you must resume it to delete the eligible records. Select **Yes** or **No** from the list.	Required.

Parameter	Description	Required or Optional
Entity	The name of the entity with related or unrelated tables in the Data Vault archive folder. To narrow the scope of the search, select a single entity.	Optional.
Purge Deleted Records	Determines whether to delete the eligible records from the Data Vault. Select **Yes** or **No** from the list.	Required.
Email ID for Report	Email address to which to send the report.	Optional.
Purge Approved By	The name of the person who authorized the purge.	Required if you select a past date or the current date for **Purge Expiry Date**.

Running the Purge Expired Records Job

To purge records in the Data Vault that are eligible for purge, run the Purge Expired Records job.

1. Select **Jobs** > **Schedule a Job**.

   The **Schedule Job** window appears.

2. In the **Projects/Programs to Run** area, select **Standalone Programs** and click **Add Item**.

   The **Program** window appears.

3. Select **Purge Expired Records** and click **Select**.

   The **PURGE_EXPIRED_RECORDS** area appears.

4. Enter the Purge Expired Records job parameters.

5. In the **Schedule** area, select the appropriate options.

6. Optionally, enter notification information.

7. Click **Schedule**.

   Data Archive deletes the expired records from the Data Vault.

Retention Management Reports

You can generate retention management reports when you run the Update Retention Policy job or the Purge Expired Records job. To view retention management reports, you must have the retention administrator role and permission to view the entities.

Note: The reports are created with the Arial Unicode MS font. To generate the reports, you must have the font file ARIALUNI.TTF saved in the `<Data Archive installation>\webapp\WEB-INF\classes` directory.

You can generate the following reports:

Retention Modification Report

The Retention Modification Report displays information about records affected by retention policy changes. The report displays the old and new retention policies, the rules that determine the expiration date for records, and the number of affected records. The report also lists each record affected by the retention policy change.

Generate a Retention Modification report from the **Monitor Jobs** window when you run the Update Retention Policy job.

Retention Expiration Summary Report

The Retention Expiration Summary Report displays information about expired records. The report displays the total number of expired records and the number of records on legal hold within a retention policy. It also displays the number of expired records and the number of records on legal hold for each entity.

Generate a Retention Expiration Summary report from the **Monitor Jobs** window when you run the Purge Expired Records job.

Retention Expiration Detail Report

If your records belong to related entities, the Retention Expiration Detail report lists expired records so you can verify them before you purge expired records. The report displays the total number of expired records and the number of records on legal hold within a retention policy. For each expired record, the report shows either the original expiration date or the updated expiration date. It shows the updated expiration date if you change the retention policy after Data Archive assigns the policy to records.

If your records belong to unrelated entities, the Retention Expiration Detail report displays the total number of expired records and the number of records on legal hold within a retention policy. It does not list individual records.

Generate a Retention Expiration Detail report from the **Monitor Jobs** window when you run the Purge Expired Records job.

Retention Management when Archiving to EMC Centera

Use the informia.defaultArchiveAreaRetention property in the `conf.properties` file to set the default retention period for an archive area. Retention policies are enforced at the record or the transaction level.

You might need to change the default retention period for an archive area for certain Centera versions, such as *Compliance Edition Plus*. Some Centera versions do not allow you to delete a file if it is within an archive area that has indefinite retention. You cannot delete the file even if all records are eligible for expiration.

You can choose immediate expiration or indefinite retention. Specify the retention period in days, months, or years. The retention period determines when the files are eligible for deletion. The default retention period applies to any archive area that you create after you configure the property. When you create an archive area, the Data Vault Service uses the retention period that you configure in the properties file. The Data Vault Service enforces the record level policy based on the entity retention policy that you configure in Data Archive.

CHAPTER 10

External Attachments

This chapter includes the following topics:

External Attachments Overview

External attachments are transaction-related attachments that are stored in the production database. You can archive encrypted or unencrypted external attachments to the target database or the Data Vault. External attachments can be archived with the data you want to archive or retire, or independent of their associated data.

If you want to archive or retire encrypted external attachments, such as Siebel attachments, to the Data Vault, you can use the Data Vault Service for External Attachments (FASA) to convert encrypted attachments from the proprietary format. You can then access the decrypted external attachments through the Data Discovery portal.

Archive External Attachments to a Database

To archive external attachments, you must first configure properties in the source connection. You can then archive the attachments with the associated data or after you have archived the associated data.

You can move external attachments to the target database synchronously or asynchronously. If you archive external attachments synchronously, the ILM Engine archives the external attachments with the associated data when you run an archive job. If you have already archived the associated data, you can move the external attachments asynchronously from source to target with the Move External Attachments job.

You can archive either encrypted or unencrypted attachments to a database. When you archive encrypted external attachments, FASA is not required because the files are stored in tables.

Before you run an archive job or a Move External Attachments job, configure the following properties in the source connection to archive external attachments:

Source/Staging Attachment location

Source location of the external attachments. Enter the directory where the attachments are stored on the production server. You must have read access on the directory.

Target Attachment location

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

Move Attachments in Synchronous Mode

Moves the attachments synchronously with the archive job. If you want to archive the external attachments with the associated data when you run an archive job, select this checkbox.

Archive Job Id

Include the Job Id of the archive or restore job to move external attachments asynchronously.

Archive or Retire External Attachments to the Data Vault

You can archive or retire external attachments to the Data Vault. To archive external attachments to the Data Vault, you must also archive the associated data. You can access external attachments through the Data Discovery Portal after you archive or retire them.

To archive external attachments to the Data Vault, the attachment locations must be available to both the ILM Engine and the Data Vault Service. If the external attachments are encrypted, such as Siebel attachments, you must configure FASA to decrypt the attachments. Before you run the archive or retirement job, configure the source and target connections.

If your environment is configured for live archiving, you must also create and configure an interim table in the Enterprise Data Manager for the attachments.

If you plan to retire all of the external attachments in the directory, you do not need to configure an interim table in the EDM. If you want to retire only some of the external attachments in the directory, you must configure an interim table in the EDM and add a procedure to the Run Procedure in Generate Candidates step of the archive job.

If you configured the source connection to move the attachments synchronously with the archive or retirement job, run the Load External Attachments job after the archive job finishes. The Load External Attachments job loads the attachments to the Data Vault.

If you did not move the attachments synchronously with the archive job, run the Move External Attachments standalone job after you run the archive job. The Move External Attachments job moves the attachments to the interim attachment table. You can then load the attachments to the Data Vault with the Load External Attachments job.

After you archive the external attachments, you can view the attachments in the Data Discovery Portal. To view the archived external attachments, search the Data Vault for the entity that contains the attachments. You cannot search the contents of the external attachments.

Step 1. Configure Data Vault Service for External Attachments

If you want to archive encrypted external attachments to the Data Vault, configure the Data Vault Service for External Attachments (FASA) on the production server.

1. Install `File_Archive_service_for_attachments.zip` on the production server where the source ERP is running.
2. Open the `configurations.properties` file in the installation directory.

3. Configure the following properties:

- ZIP.UTILITY. Zip utility name and location.

    For example, `\\siebsrvr\\BIN\\ssezipXX.exe` where XX is the utility version.

- UNZIP.UTILITY. Unzip utility name and location.

    For example, `\\siebsrvr\\BIN\\sseunzipXX.exe` where XX is the utility version.

- PROD.FILE.LOCATION. Production file system location.

- ARCHIVE.FILE.LOCATION. Archive file system location.

- TEMP.FOLDER.LOCATION. Temporary folder location that is required for storing the `.bat` and `.sh` files in the decryption script.

4. Change the port number in the `conf.properties` file and start FASA.

5. To verify that FASA is running, enter the port number in a browser window. If the FASA server is running, you will receive a confirmation message.

Step 2. Configure the Source Connection

Before you archive or retire external attachments, configure source connection properties.

1. Create or update the source connection.

- To create a connection, click **Administration > New Source Connection**.

- To update a connection, click **Administration > Manage Connections**.

2. In the source connection, configure the following properties:

- **Source/Staging Attachment Location.** Source location of the external attachments. Enter the directory where the attachments are stored on the production server. You must have read access on the directory.

- **Target Attachment Location.** Target location for the external attachments. Enter the archive folder where you want to archive the external attachments. You must have write access on the directory.

3. If the attachments are encrypted, enter the location of the temporary folder you designated for the decryption script in the **Staging Script Location** field.

4. If you want to move the attachments to the interim attachment table during the archive or retirement job run, select **Move Attachments in Synchronous Mode**.

Step 3. Configure the Target Connection

If the attachments are encrypted, configure the target connection for decryption.

1. Create or update the target connection.

- To create a connection, click **Administration > New Target Connection**.

- To update a connection, click **Administration > Manage Connections**.

2. In the **Add On URL** field, enter the FASA URL.

Step 4. Configure the Entity in the Enterprise Data Manager

To perform a live archive, or retire only some of the attachments in the directory, configure the attachment entity in the EDM. **Skip this step if you plan to retire all of the attachments in the directory**.

Create an interim table with additional columns to archive external attachments. Then modify the entity steps to move the attachments. The archive job moves the attachments during the Generate Candidates step.

1. To launch the EDM, click **Accelerators > Enterprise Data Manager**.

2. Right-click the entity that contains the external attachments you want to archive.

3. Select **New Interim Table**.

4. Enter a description for the interim table and click **Finish**.

5. On the **Default Columns** tab of the interim table, add the following columns:

    - ATTACHMENT_NAME. In the **Select clause** column, enter the following string:
      'S_ORDER_ATT' || '_' || Z.ROW_ID || '_' || Z.FILE_REV_NUM || '.' || Z.FILE_EXT

    - ZIPPED_FILE_EXT. In the **Select clause** column, enter "SAF".

    Interim table configuration is complete.

6. Select the entity that contains the external attachments you want to archive.

7. To retire only some of the files in the directory, locate the attachments interim table on the **Steps** tab and select the step called **Run Procedure in Generate Candidates**.

8. Enter the following text in the **Procedure** column: `java://`
    `com.applimation.archive.dao.impl.MoveFileDAOImpl`

Step 5. Archive or Retire External Attachments

To archive or retire the attachments synchronously, run the archive job and then load the external attachments to the Data Vault. To archive the attachments asynchronously, run the archive or retirement job and then move the attachments to the interim attachment table before you archive or retire them.

Archiving or Retiring External Attachments Synchronously

1. Run the archive or retirement job.

    The attachments will be moved from the source directory to the interim attachment table during the job run.

2. Run a Load External Attachments job to archive the attachments to the Data Vault. When you schedule the Load External Attachments job, configure the following parameters:

   Directory

   The source location of the attachments. If the attachments are not encrypted, this directory must be accessible to both the ILM Engine and the Data Vault Service.

   Attachment Entity Name

   The name of the attachment entity.

   Target Archive Store

   The target connection used to archive the application data. The Target Archive Store parameter ensures that the attachments are archived to the same Data Vault archive folder as the application data.

   Purge After Load

   If you want the job to delete the attachments you are loading to the Data Vault, select **Yes**. The job deletes the contents of the directory that you specified in the Directory parameter.

Archiving or Retiring External Attachments Asynchronously

1. Run the archive or retirement job.

2. Run a Move External Attachments standalone job. To move external attachments asynchronously, configure the following parameters:

   Add on URL

   FASA URL.

   Archive Job Id

   Job ID of the completed archive or retirement job.

   The Move External Attachments job moves the attachments from the source directory to the interim directory in the attachment entity.

3. To move the attachments to the Data Vault, run a Load External Attachments job.

Step 6. View Attachments in the Data Discovery Portal

To confirm the attachments were retired, view them in the Data Discovery Portal.

1. Click **Data Discovery** > **Search Data Vault.**

2. In the **Entity** field, select the entity containing the attachments you retired and click **View**.

   The **Results** table appears.

3. Click the icon beneath **View Archived Data**.

CHAPTER 11

Data Archive Restore

This chapter includes the following topics:

Data Archive Restore Overview

If you need to modify an archived transaction, or if a cycle was archived in error, you can restore the data from the archive. When you restore data, Data Archive reverses the actions that it used to archive the data.

You can restore data from the Data Vault or a database archive. You can restore data from the Data Vault to an empty or existing production database or to a database archive. You can restore data from a database archive to an empty or existing database, including the original source.

You can also restore external attachments that have been archived to a database archive or the Data Vault. Data Archive restores external attachments as part of a restore job from either a database or the Data Vault.

By default, when you restore data from a database archive or the Data Vault, Data Archive deletes the data from the archive location. If you restore a cycle or transaction, you can choose to keep the restored data at the archive location.

Restore Methods

You can run a transaction or cycle restore from either the Data Vault or a database archive. You can only perform a full restore from a database archive.

You can perform a restore job using the following methods:

Full restore

A full restore job restores the entire contents of a database archive. You might perform a full restore if you archived an entire database that must be returned to production. For optimal performance, conduct a full restore to an empty database. Performance is slow if you run a full restore to the original database.

Data Archive deletes the restored data from the database archive as part of the full restore job.

Cycle restore

A cycle restore job restores an archive cycle. You might restore an archive cycle if a cycle was archived in error, or if you must change numerous transactions in a cycle. To restore a cycle you must know the cycle ID. A cycle restore can also restore a specific table. To restore a specific table, use the Enterprise Data Manager to create an entity containing the table you want to restore.

By default, Data Archive deletes the restored data from the database archive or the Data Vault as part of the cycle restore job. If you do not want the restore job to delete the data from the archive location, you can specify that the restore job skip the delete step.

Transaction restore

A transaction restore job restores a single transaction, such as a purchase order or invoice. To restore a transaction, you must search for it by the entity name and selection criteria. If your search returns multiple transactions, you can select multiple transactions to restore at the same time.

By default, Data Archive deletes the restored data from the database archive or the Data Vault as part of the transaction restore job. If you do not want the restore job to delete the data from the archive location, you can specify that the restore job skip the delete step.

Data Vault Restore

You can restore data from the Data Vault to a database archive or a production database. A production database can be either empty or existing. Data can be restored as either an archive cycle, a single transaction, or multiple transactions. You can also restore a specific table with the cycle restore feature.

You cannot restore archive only cycles from the Data Vault. Only archive and purge cycles can be restored from the Data Vault.

When you restore from the Data Vault, determine the type of target for the restore job:

- Restore to an empty production database or database archive.
- Restore to an existing production database without schema differences.
- Restore to an existing production database with schema differences.

Data Vault Restore to an Empty Database or Database Archive

Before you restore data to an empty production database or database archive, you must run the metadata DDL on the target database. You can then run either a cycle or transaction restore.

Restoring to an Empty Database or Database Archive

Verify that a target connection exists and that you have access to the connection.

If you restore to an empty database, you must first export the metadata associated with the records to be restored and run the metadata DDL on the database.

1. In the Data Discovery Portal, search for the records that you want to restore.
2. Export the metadata associated with the records.
3. Run the metadata DDL on the target database.
4. Run the transaction or cycle restore job from the Data Vault to the target database.

Data Vault Restore to an Existing Production Database

When you restore data to an existing production database, you must determine whether the schema on the original production database has been updated since the data was archived.

If the production database schema has been updated, you must synchronize schemas between the existing database and the Data Vault. You can then run either a cycle or transaction restore.

Note: When you restore archived data and the records exist in the target database, the restore job will fail if key target constraints are enabled. If target constraints are disabled, the operation might result in duplicate records.

Restoring to an Existing Database Without Schema Differences

If the Data Vault schema is identical to the existing database schema, you can search for the transaction records and then run the restore job.

Verify that the target connection exists and that you have access to the connection.

1. In the Data Discovery Portal, search for the records that you want to restore.
2. Run the restore job from the Data Vault to the target database.

Schema Synchronization

When you restore data from the Data Vault to an existing database, you might need to synchronize schema differences between the Data Vault and the existing database.

To synchronize schemas, you can generate metadata in the Enterprise Data Manager and run the Data Vault Loader to update the Data Vault schema, or you can use a staging database.

Update the Data Vault schema with the Data Vault Loader when changes to the target table are additive, including extended field sizes or new columns with a null default value. New columns with a non-null default value require manually running an ALTER TABLE SQL command. If you removed columns or reduced the size of columns on the target table, you must synchronize through a staging database.

Synchronizing Directly to the Target Database

When you restore data from the Data Vault to an existing database with a different schema, you must first generate new metadata. Then run the Data Vault Loader standalone job to update the Data Vault schema.

Verify that the target connection exists and that you have access to the connection.

1. Define and run an archive job that generates metadata for the new schema, but does not archive any records.
2. From the **Schedule Jobs** page, run a Data Vault Loader job to load the metadata and update the schema.
3. Verify that the job was successful on the **Monitor Jobs** page.
4. Run the restore job from the Data Vault to the target database.

Synchronizing through a Staging Database

If you are unable to update the Data Vault by running the Data Vault Loader, you can use a staging database on which you run the updated metadata. You can then manually conduct the restore from the staging database to the existing database.

Verify that the target connection exists and that you have access to the connection.

1. In the Data Discovery Portal, search for the records that you want to restore.
2. Export the metadata associated with the records.
3. Run the metadata DDL on the staging database.
4. Run the restore job from the Data Vault to the staging database.
5. Alter the table in staging to synchronize with the target database schema.
6. Using dlink and SQL scripts, manually restore the records from the staging database to the target database.

Running a Cycle or Transaction Restore Job

Once schemas are synchronized, you can run a cycle or transaction restore from the Data Vault.

Verify that a target connection exists and that you have access to the connection.

1. Click **Workbench > Restore File Archive**.
2. Select the **Cycle Restore** or **Transaction Restore** tab.
   - To restore an archive cycle, select the cycle ID.
   - To restore a transaction, enter the entity name and selection criteria.
3. Click **Next.**

   The **Manage Restore Steps** page appears.
4. 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 Data Vault, enable the **Skip** check box for the **Delete from Archive** step.
5. Click **Schedule** to schedule the job.
6. Monitor the restore job from the **Jobs > Monitor Jobs** page.

Configuring Data Vault Restore Steps

When you restore data from the Data Vault, Data Archive performs a series of steps similar to an archive cycle. After each step in the restore process you can choose to pause the job or receive a notification.

1. Generate Candidates. Generates a row count report for the data being restored.
2. Validate Destination. Validates the destination by checking table structures.
3. Copy to Destination. Restores the data from the Data Vault to the target. You have the option of generating a row count report after this step.
4. Delete from Archive. Deletes the restored data from the Data Vault if it is not on legal hold. This step is optional.

Restoring a Specific Table

To restore a specific table, you must create a new entity and run a cycle restore.

1. In the Enterprise Data Manager, create a new entity containing the specific table you want to restore.
2. Run a cycle restore and specify the entity you created.

Database Archive Restore

You can restore data that you have archived to a database archive. When you restore data from a database archive, Data Archive performs the same steps in the same sequence as an archive cycle.

When you restore data from the database archive, the history application user functions like the production application user during an archive cycle. The history application user requires the SELECT ANY TABLE privilege to run a restore job.

You can conduct a full, cycle, or transaction restore from a database archive. If you are restoring data from a database archive that was previously restored from the Data Vault to a database archive, you can only run a cycle restore. You can only run a cycle restore because the restore job ID is not populated in the interim tables during the initial archive cycle.

The target database for a database archive restore job can be either an empty or existing production database.

When you run a full restore, the system validates that the source and target connections point to different databases. If you use the same database for the source and target connections, the source data may be deleted.

Prerequisites

Verify that the target connection for the restore exists and that you have access to the connection.

If you are running a restore job from an IBM DB2 database back to another IBM DB2 database, an ILM Administrator may have to complete prerequisite steps for you. Prerequisite steps are required when your organization's policies do not allow external applications to create objects in IBM DB2 databases. The prerequisite steps include running a script on the staging schema of the restore source connection.

Note: When you restore archived data and the records exist in the target database, the restore job will fail if key target constraints are enabled. If target constraints are disabled, the operation might result in duplicate records.

Running the Create Cycle Index Job

Before you conduct a restore from a database archive, run the Create Cycle Index job to increase restore job performance. The Create Cycle Index job creates database indexes based on Job IDs to optimize performance during the restore.

1. Click **Jobs > Schedule a Job**.

2. Choose **Standalone Programs** and click **Add Item**.

   The **Select Definitions** window appears.

3. Select the **Create Cycle Index** job.

4. Select the target repository and the entity where indexes need to be created.

5. Schedule the job to run immediately or on a certain day and time.

6. Enter an email address to receive a notification when the completes, terminates, or returns an error.

7. Click **Schedule.**

Schema Synchronization

When you restore from a database archive to an existing database, the Validate Destination step in the restore cycle will synchronize schema differences when the changes are additive, such as extending field sizes or adding new columns. If changes are not additive, you must create and run a SQL script on the database archive to manually synchronize with the production database.

Restoring from a Database Archive

You can run a cycle, full, or transaction restore job from a database archive.

1. Click **Workbench > Restore Database Archive**.

2. Select the **Cycle Restore**, **Full Restore**, or **Transaction Restore** tab.

   - To restore an archive cycle, select the cycle ID.

   - To restore a full database archive, select the source and target connections.

   - To restore a transaction, enter the entity name and selection criteria.

3. Click **Next**.

   The **Manage Restore Steps** page appears.

4. 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, enable the **Skip** check box for the **Delete from Source** step.

5. Click **Schedule** to schedule the job.

6. Monitor the restore job from the **Jobs > Monitor Jobs** page.

Configuring Database Archive Restore Steps

When you restore data from a database archive, Data Archive performs a series of steps similar to an archive cycle. After each step in the restore process you can choose to pause the job or receive a notification.

1. Generate Candidates. Generates interim tables and a row count report for the data being restored.

2. Build Staging. Builds the staging table structures for the data you want to restore.

3. Copy to Staging. Copies the data from the database archive to your staging tablespace. You have the option of generating a row count report after this step.

4. Validate Destination. Validates the destination by checking table structures.

5. Copy to Destination. Restores the data from the database archive to the target. You have the option of generating a row count report after this step.

6. Delete from Source. Deletes the restored data from the database archive. This step is optional.

7. Purge staging. Deletes interim tables from the staging schema.

Restoring a Specific Table

To restore a specific table, you must create a new entity and run a cycle restore.

1. In the Enterprise Data Manager, create a new entity containing the specific table you want to restore.

2. Run a cycle restore and specify the entity you created.

External Attachments Restore

External attachments are restored as part of a regular Data Vault or database archive restore job. Attachments restored from a database archive are identified in interim tables. Attachments restored from the Data Vault are identified in the attachments.properties file.

External Attachments from the Data Vault

When you run a restore job from the Data Vault, the ILM engine retrieves unencrypted attachments through the Data Vault Service and places the attachments in a staging attachment directory.

For encrypted attachments, when you run a restore job the ILM engine generates an encryption script in the staging attachment directory. The engine invokes the Data Vault Service for External Attachments (FASA) to create an encryption utility that reencrypts the attachments in the staging directory and then restores them to the target directory.

The following table lists the directories that FASA and the ILM engine need to access:

Format	Scripts Directory	Decrypted Directory	BCP Directory/Attachment Target Directory
Regular	Not required	ILM engine	ILM engine
Encrypted	FASA and ILM engine (shared mount)	FASA and ILM engine (shared mount)	FASA

External Attachments from a Database Archive

The following table lists the directories that the ILM engine needs to access:

Format	Attachment Source Directory	Scripts Directory	Attachment Target Directory
Regular/Encrypted	ILM engine (mount point)	ILM engine (local)	ILM engine (mount point)

CHAPTER 12

Data Discovery Portal

This chapter includes the following topics:

Data Discovery Portal Overview

Use the Data Discovery portal to search data that is archived to the Data Vault. The Data Discovery portal includes the following search tools:

Search Data Vault

Use a keyword, term, or specific string to search for records across applications in Data Vault. Records that match the keyword appear in the search results. A keyword or string search searches both structured and unstructured data files, for example .docx, .pdf or .xml files. You can narrow the search results by filtering by application, entity, and table. After you search the Data Vault, you can export the search results to a .pdf file.

Search Within an Entity

Use Search Within an Entity to search and examine application data from specific entities in the Data Vault. The Data Discovery portal uses the search criteria that you specify and generates a SQL query to access data from the Data Vault. After you search the Data Vault, you can export the search results to a CSV or PDF file.

Browse Data

Use Browse Data to search and examine application data from specific tables in the Data Vault. The Data Discovery portal uses the search criteria that you specify and generates a SQL query to access data from the Data Vault. After you search the Data Vault, you can export the search results to a .csv or .pdf file.

You can also use the Data Discovery portal to apply retention policies, legal holds, and tags to archived records. If data-masking is enabled, then sensitive information is masked or withheld from the search results.

Search Data Vault

You can enter a keyword or set of words, add wildcards or boolean operators to search for records in Data Vault.

Data Archive searches across applications in Data Vault for records that match your search criteria. You can narrow your search results by filtering by application, entity, and table. Then export your results to a PDF file.

You access Search Data Vault from the **Data Discovery** menu.

The following figure shows the **Search Data Vault** page:
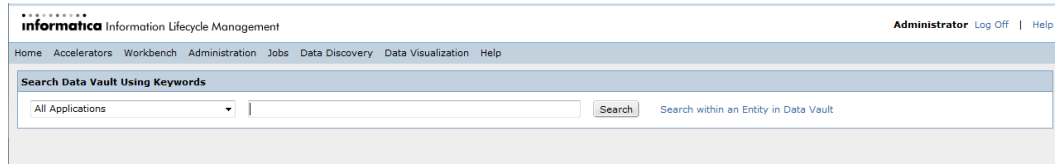
Records that match your search query display on the **Results** page. The following figure shows the **Results** page.
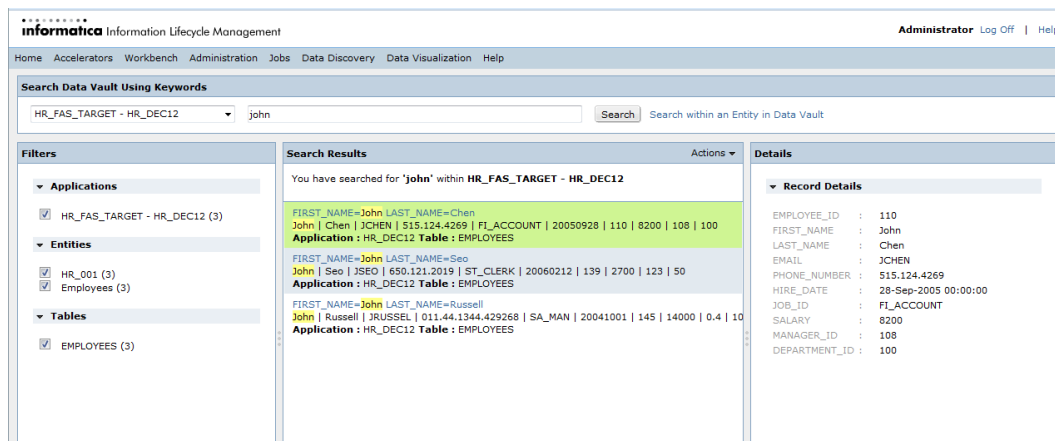
1. The Filters panel contains sections for Application, Entities, and Tables filters. Clear the appropriate filter boxes to narrow your search results.

2. The Search Results panel displays records that match your search criteria. Click the record header to view technical details. Click the record summary to view the record's details in the Details panel.

3. The Actions menu contains options to export results to a PDF file. You can export all results or results from the current page.

4. The Details panel displays a record's column names and values.

Examples of Search Queries

You can search for records archived to the Data Vault in a variety of ways to make your search results more relevant. You can use a word or phrase, add boolean operators or wildcard characters, or specify a range of values.

The following list describes some of the search queries you can use:

Single Word

Enter a single word in the search field. For example, enter `Anita`. Records with the word Anita appear in the results.

Phrase

Enter a phrase surrounded by double quotes. For example, enter `"exceeding quota"`. Records with the term exceeding quota appear in the results.

Replace Characters

Use ? and * as wildcard characters to replace letters in a single word search. You cannot use wildcards if your search query has more than one word. You cannot replace the first character of the word with a wildcard character.

To replace a single character in a word, use ?. For example, enter `te?t`. Records with the word tent, test, or text appear in the results.

To replace multiple characters in a word, use *. For example, enter `te**`. Records with the word teak, teal, team, tear, teen, tent, tern, test, or text appear in the results.

Similar Spelling

Use the ~ character to search for a word with a similar spelling. For example, enter `roam~`. Records with the word foam or roams appear in the results.

Range Searches

You can search for records with a range of values. Use a range search for numbers or words.

Enter the high and low value of a range separated by TO. The TO must be in uppercase.

To search by range, use the following formats:

- To include the high and low values, surround the range with square brackets:

      `<column name>: [<high> TO <low>]`

  For example, enter `name: [Anita TO Ella]`

  Records with the word Anita, Ella, and words listed alphabetically between Anita and Ella in the name column, appear in the results.

- To exclude the high and low values, surround the range with curly brackets:

      `<column name>: {<high> TO <low>}`

  For example, enter `name: {Anita TO Ella}`

  Records with the words listed alphabetically between Anita and Ella in the name column appear in the results. Records with the word Anita or Ella do not appear in the results.

- To search for a date range, use the format yyyymmdd. Use square brackets to include the high and low values. Use curly brackets to exclude the high and low values.

      `<column name>: [<yyyymmdd> TO <yyyymmdd>]`

  For example, enter `hire_date: [20040423 TO 20121231]`

  Records with dates from April 23, 2004 to December 31, 2012 in the hire_date column appear in the results.

AND Operator

To search for records that contain all the terms in the search query, use AND between terms. The AND must be in uppercase. You can replace AND with the characters &&.

For example, enter `"Santa Clara CA" AND Platinum`. Records with the phrase Santa Clara CA and the word Platinum appear in the results.

+ Operator

To specify that a term in the search query must exist in the results, enter the + character before the term. For example, enter `+semiconductor automobile`. Records with the word semiconductor appear in the results. These records may or may not contain the word automobile.

NOT Operator

To search for records that exclude a term in the search query, enter NOT before the term. You can replace NOT with either the ! or - character.

For example, enter `Member NOT Bronze`. Records that contain the word Member but not the word Bronze appear in the results.

Searching Across Applications in Data Vault

Search for records across applications in Data Vault. Enter a search criteria and specify the applications to search from. Narrow your search results by application, entity, and table. Then export results to a PDF file.

1. Click **Data Discovery** > **Search Data Vault**.

   The **Search Data Vault Using Keywords** page appears.

2. Click the **Choose an Application** list and select one or more applications.

3. Enter a search query and click **Search**.

   Records that match your search criteria display on the **Search Data Vault Results** page.

4. On the **Filters** panel, check the appropriate application, entity or tables to filter your results.

5. On the **Search Results** panel, click on a record.

   The column names and values appear on the **Details** panel.

6. On the **Search Results** panel, click the record header.

   The technical view appears with the details and referential data details of the record.

7. To export results, click **Actions**.

   - To export results from the current page, click **Export Current Page as PDF**.
   - To export all results, click **Export All Results as PDF**.

   A PDF file with the results appear.

Search Within an Entity in Data Vault

The Search Within an Entity enables you to use the Data Discovery portal to search and examine application data from specific entities in the Data Vault. The Data Discovery portal uses the search criteria that you specify and generates an SQL query to access data from the Data Vault. After you search the Data Vault, you can export the search results to a file.

When you search the Data Vault, you specify the archive folder and the entity that you want to search in. If the archive folder includes multiple schemas, the query uses the mined schema in the Enterprise Data Manager for the corresponding entity. The search only returns records from the corresponding schema and table combination. If the archive folder only includes one schema, the query uses the `dbo` schema.

Use table columns to specify additional search criteria to limit the amount of records in the search results. The columns that you can search on depend on what columns are configured as searchable in the Data

Discovery portal search options. Based on your system-defined role, you can also search the Data Vault based on tags, legal hold groups, retention expiration date, and effective retention policy.

You can add multiple search conditions. Use AND or OR statements to specify how the search handles the conditions. You can also specify the maximum number of records you wish to see in the results.

Search Within an Entity Parameters

Specify the search parameters when you search the Data Vault by entity.

Search Within an Entity includes the following search parameters:
Archive Folder

Archive folder that contains the entity of the archived data that you want to search. The archive folder limits the amount of entities shown in the list of values.

The archive folder list of values shows archive folders that include at least one entity that you have access to. You have access to an entity if your user is assigned to the same role that is assigned to the entity.

Entity

Entity that contains the table that you want to view archived data from.

The entity list of values shows all entities that you have access to within the application version of the related archive folder.

Maximum Number of Records in Results

The maximum number of records that you want to view in the search results.

This parameter does not affect the export functionality. When you export results, Data Archive exports all records that meet the search criteria regardless of the value you specify in the Maximum Number of Records in Results parameter.

Enter a number greater than zero.

Searching Within an Entity

Use the Data Discovery portal to search and examine application data from the Data Vault. When you search the Data Vault, you search archived data from one entity.

1. Click **Data Discovery** > **Search Within an Entity in Data Vault**.
2. Select the archive folder and entity.

   After you choose the entity, the system displays the search elements that are defined for the entity.

   If you want to filter the search by metadata columns such as Legal Hold, Effective Retention Policy, and up to nine tag columns, contact the Data Archive administrator for permissions.
3. Optionally, specify the maximum number of records you wish to view in the search results.
4. Enter a condition for at least one of the search elements.

   Use the **Add (+)** or **Remove (-)** icons to add or remove multiple search conditions.

   Use % as a wildcard to replace one or more characters for the **Starts With**, **Does Not Start With**, **Ends With**, **Does Not End With**, **Contains**, and **Does Not Contain** search options.
5. Click **View**.

Search Within an Entity Results

When you search the Data Vault by entity, the Data Discovery portal generates two patterns of search results:

- Transaction View. The transaction view displays basic transaction details of the entity.
- Technical View. The technical view displays transaction details along with referential data details of the entity.

Search Within an Entity Export Results

When you search the Data Vault by entity, you can export the search results to a file. The Data Discovery portal search options determine which columns are exported.

The export file includes columns that are defined as display columns. The export file does not include the legal hold, tagging, or effective retention policy columns. The export file does not contain the retention expiration date column if you export selected records.

You can export the search results to one of the following file types:

- XML file
- PDF file
- Insert statements
- Delimited text file

Based on the file type that you download to, you can choose one of the following download options:

- Individual files. If multiple transactions are selected, each transaction has a separate file.
- Compressed file. All the selected transactions are downloaded in a compressed file.

You determine which records from the search results are included in the export file. Before you export, you can select individual records, select all records on the current page, or select all records from all pages.

When you export the search results, the job exports the data and creates the .pdf file in the temporary directory of the ILM application server. The job uses the following naming convention to create the file:

```
<table>_<number>.pdf
```

The job creates a hyperlink to the file that is stored in the temporary directory. Use the link to view or save the .pdf file when you view the job status. By default, the files are stored in the temporary directory for one day.

Delimited Text Files

When you export to a delimited text file, the system exports data to a .csv or .txt file.

When you export to a delimited text file, specify the following parameters:

Column and Row Separators

When you search the Data Vault by entity and select delimited text file as the export file type, the separators include default values. The following table describes the default values for the separators:

Separator	Default Value
Column Separator	,
Row Separator	\n

The values for the column and row separator parameters determine the type of delimited text file the system exports to. If you use the default values for the column and row separator parameters, then the system exports data to a .csv file. If you use values other than the default values, then the system exports data to a .txt file.

Put Values in Quotes

The Put Values in Quotes parameter is for .csv export. Enable the parameter if the source data includes new line characters. The parameter determines whether the system puts the source values in quotes internally when the system creates the .csv file. The quotes ensure that source data with new line characters is inserted into one column as compared to multiple columns.

For example, the source data contains the following information:

Name	Address
John Doe	555 Main Street Apartment 10

The address column contains a new line character to separate the street and the apartment number. If you enable the parameter, the exported file includes the address data in one column. If you disable the parameter, the exported file splits the address data into two columns.

Only use this parameter if you use the default values for the separators. If you enable this parameter and use values other than the default values for the separators, then the system puts the values in quotes in the exported data.

Exporting Selected Records from the Search Within an Entity Results

After you run a search, you can export selected records to a file.

1. From the search results, select the rows that you want to export.

   Tip: You can manually select the rows or use the column header checkbox to select all records on the page.

2. Click **Export Data**.

   The **Export Transaction Data** dialog box appears.

3. Select the export data file type.

   Note: User authorizations control whether you can choose the file type to export to. Depending on your user role assignment, the system may automatically export to a .pdf file.

4. If you select Delimited Text File, enter the additional parameters.

5. Select the download option.

6. Click **Fetch Data**.

   The **Files to Download** dialog box appears.

7. View or save the file.

   - To view the file, click the hyperlink.

   - To save the file, right-click the hyperlink, specify the file location, and save.

Browse Data

Browse Data enables you to use the Data Discovery portal to search and examine application data from specific tables in the Data Vault. The Data Discovery portal uses the search criteria that you specify and generates a SQL query to access data from the Data Vault. After you search the Data Vault, you can export the search results to a CSV or PDF file.

When you specify the search criteria, you choose which table columns to display in the search results. You can display all records in the table or use a where by clause to limit the search results. You can also specify the sort order of the search results. Before you run the search, you can preview the SQL statement that the Data Discovery portal uses to access the data.

You can also use Browse Data to view external attachments for entities that do not have a stylesheet. You can view external attachments after the Load External Attachments job loads the attachments to the Data Vault. You cannot search the contents of the external attachments.

When you search for external attachments, enter the entity that the attachments are associated to. The search results show all attachments from the AM_ATTACHMENTS table. The table includes all attachments that you archived from the main directory and also from any subfolders in that main directory. The search results show a hyperlink to open the attachment, the original attachment directory location, and the attachment file name. To view the attachment, click the hyperlink in the corresponding ATTACHMENT_DATA column. When you click the link, a dialog box appears that prompts you to either open or save the file.

Browse Data Search Parameters

Specify the search parameters when you browse data from the Data Vault.

The Browse Data search includes the following search parameters:

Archive Folder

Archive folder that contains the entity of the archived data that you want to search.

The archive folder list of values shows archive folders that include at least one entity that you have access to. You have access to an entity if your user is assigned to the same role that is assigned to the entity.

You must select an archive folder to limit the amount of tables shown in the list of values.

Entity

Entity that contains the table that you want to view archived data from.

The entity list of values shows all entities that you have access to within the application version of the related archive folder.

The entity is optional. Use the entity to limit the amount of tables shown in the list of values.

Schema

Schema that contains the table that you want to view. Required if the target connection for the archive folder is enabled to maintain the source schema structure. You must enter the schema because the archive folder may include multiple schemas with the same table name.

Table

Table that includes the archived data that you want to view.

For Browse Data, the table list of values shows all tables that you have access to within the application version of the related archive folder.

No. of Rows

Maximum amount of rows in the search results. Default is 100.

Searching with Browse Data

Use the Browse Data search tool to search data from the Data Vault.

1. Click **Data Discovery** > **Browse Data**.
2. Specify the search parameters.

   After you select the table, the system populates the **Available Columns** list.

   If you want to filter the search by metadata columns such as Hold, Effective Retention Policy, and up to nine tag columns, contact the Data Archive administrator for permissions.

   Tag columns are named *ilm_metafield1* through *ilm_metafield9*.
3. Use the arrows to move the columns that you want to display to the **Display Columns** list.
4. Optionally, specify the **Where Clause** and **Order By** to further filter and sort the search results.
5. Click **Preview SQL** to view the generated SQL query.
6. Click **Search**.

   The records that match the search criteria appear. A **Click to View** link appears if the data in a record contains more than 25 characters or the record has an attached file.

Export Data

You can export the search results to a .csv or a .pdf file. System-defined role assignments determine which file type users can export data to. The exported file includes the columns that you selected as the display columns for the browse data search. Note that the exported file does not include the legal hold and tagging columns.

When you export the search results, you schedule the Browse Data Export job. The job exports the data and creates a .pdf file in the temporary directory of the ILM application server. The job uses the following naming convention to create the file:

```
<table>_<number>.pdf
```

The job creates a hyperlink to the file that is stored in the temporary directory. Use the hyperlink to view or save the .pdf file when you view the job status.

Exporting from Browse Data

After you run a search, you can export the search results. You can export all columns except for columns that include large datatype objects such as BLOB or CLOB datatypes.

1. After the search results appear, select **CSV** or **PDF** and click **Export Data**.

   If you export data from a table that includes a large datatype object, a message appears. The message indicates which columns contain the large object datatypes and asks if you want to continue to export the data. If you continue, all of columns except for the columns with the large datatypes are exported.

   Note: User authorizations control whether you can choose the file type to export to. Depending on your user role assignment, the system may automatically export to a .pdf file.

   The **Schedule Jobs** screen appears.
2. Schedule the Browse Data Export job.
3. Access **Jobs** > **Monitor Jobs**.

The **Monitor Jobs** screen appears if you schedule the job to run immediately.

4. Once the job completes, expand the Job ID.

5. Expand the BROWSE_DATA_EXPORT_SERVICE job.

6. Click **View Export Data**.

   The **File Download** dialog box appears.

7. Open or save the file.

Legal Hold

You can put records from related entities on legal hold to ensure that records cannot be deleted or purged from the Data Vault. You can also put a database schema or an entire application on legal hold. A record, schema, or application can include more than one legal hold.

Records from related entities, when put on legal hold, are still available in Data Vault search results. Legal hold at the application or schema does not depend on the relationship between tables.

You can apply a legal hold at the schema level to one or more schemas within an archive folder. A schema might include an entity that contains tables that are also a part of a different schema. If you try to apply a legal hold on a schema that contains tables that have a relationship to another schema, you receive a warning message. The warning includes the name of the schema that contains the related tables.

You can continue to apply the legal hold on the selected schemas, or you can cancel to return to the selection page and add the schema that contains the related tables. If you do not include the schema that contains the related tables, orphan data is created in the related tables when you purge the tables that you placed on legal hold.

A legal hold overrides any retention policy. You cannot purge records on legal hold. However, you can still update the retention policy for records on legal hold. When you remove the legal hold, you can purge the records that expired while they were flagged on legal hold. After you remove the legal hold, run the Purge Expired Records job to delete the expired records from the Data Vault.

To apply a legal hold, you must have one of the following system-defined roles:

- Administrator
- Legal Hold User

Legal Hold Groups and Assignments

You create and delete legal hold groups and apply legal hold assignments from the Manage Legal Hold Groups menu.

When you access the menu, you see a results list of all legal hold groups. The results list includes the legal hold group name and description. You can use the legal hold name or description to filter the list. From the results list, you can select a legal hold group and apply the legal hold to records or an application as a whole in the Data Vault. You can delete a legal hold group to remove the legal hold from records or an entire application in the Data Vault.

Expand the legal hold group to view all records that are assigned to the legal hold group. The details section includes the following information:

Archive Folder and Entities

Records that are on legal hold within the legal hold group are grouped by the archive folders and entities. The archive folder is the target archive area that the entity was archived to. User privileges restrict the entities shown.

Comments

Use the comments hyperlink to display the legal hold comments. The comments provide a long text description to capture information related to the legal hold group, such as the reason why you put the records on hold. You add comments when you apply a legal hold for the legal hold group.

View Icon

You can view the list of records that are on legal hold for each entity within the group. When you view the records in an entity, the Data Discovery portal search results page appears. You can view each record individually. Or, you can export the records to an XML file or a delimited text file. You must manually select the records that you want to export. You can export up to 100 records at a time.

Print Icon

You can create a PDF report that lists all the records that are on legal hold for each entity within a legal hold group. Each entity within the legal hold group has a print PDF icon. When you click the icon, a PDF report opens in a new window. The PDF report displays the entity, the legal hold name, and all of the columns defined in the search options for the entity. You can download or print the report.

Creating Legal Hold Groups

You must create a legal hold group before you can apply a legal hold to records in the Data Vault. Legal hold groups are stored in the ILM repository.

1. Click **Data Discovery** > **Manage Legal Hold Groups**.

2. Click **New Legal Hold Group**.

   The **New Legal Hold Group** dialog box appears.

3. Enter the legal hold group name and description.

   The description can be up to 200 characters.

Apply the legal hold group to records in the Data Vault.

Applying a Legal Hold

When you apply a legal hold, you search the Data Vault for records or an entire application that you want to apply the legal hold to. Then, you schedule the Add Legal Hold job. After you run the job, the system adds the legal hold to the selected records.

1. Click **Data Discovery** > **Manage Legal Hold Groups**.

2. Click the apply legal hold icon from the corresponding legal hold group.

3. Choose the archive folder.

4. To apply legal hold on all the tables in an application check **All Tables** and skip to step 11.

5. To apply a legal hold to a schema, select **Schema** and then the schema name from the menu. To apply the legal hold to all available schemas within the archive folder, select **All**. Then skip to step 11.

   Note: If you do not have access to a table or tables in a schema, that schema does not appear in the list of schemas. If you select the **All** option, only schemas to which you have access are selected.

6. To apply legal hold on a selected group of records, select the entity from the list of values.

The system displays the search elements that are defined for the entity.

7. Enter a condition for at least one of the search elements.

    Use the **Add (+)** or **Remove (-)** icons to add or remove multiple search conditions.

8. Click **View**.

    The search results appear.

9. Select records from the search results in one of the following ways:

    - Select individual records.

    - Enable the select-all box to select all records. Click the drop-down icon and select **All Records** for all records or **Visible Records** for records on the current page.

10. Optionally, click **Export** to export the selected records. You can export to an XML, PDF, or delimited text file.

11. Click **Apply Legal Hold**.

    The **Apply Legal Hold** screen appears.

12. In the comments field, enter a description that you can use for audit purposes. The description can include alphanumeric, hyphen (-), and underscore (_) characters. Any other special characters are not supported.

    For example, enter why you are adding the legal hold. The comments appear in the job summary when you monitor the Apply Legal Hold job.

13. Click **Apply Legal Hold**.

    The Schedule Job screen appears.

14. Schedule when you want to run the Apply Legal Hold job.

    When the job runs, the system applies the legal hold to all tables in an application or all records that you selected in the search results.

Deleting Legal Hold Groups

When you delete a legal hold group, the system navigates you to schedule the Remove Legal Hold job. The job removes the legal hold from all records that are associated to the legal hold group across all archive folders. Then, the system deletes the legal hold group.

To delete a legal hold group, you must have access to at least one entity that includes records that have the legal hold group assignment. You have access to an entity if your user has the same Data Vault access role as the entity. After you delete a legal hold group, you must run the Purge Expired Records job to delete any expired records from the Data Vault.

1. Click **Data Discovery** > **Manage Legal Hold Groups**.

2. Click the delete icon from the corresponding legal hold group.

    The **Remove Legal Hold** dialog box appears.

3. In the comments field, enter a description that you can use for audit purposes.

    For example, enter why you are removing the legal hold. The comments appear in the job summary when you monitor the Remove Legal Hold job.

4. Click **Delete Legal Hold**.

    The **Schedule Job** screen appears.

5. Schedule when you want to run the Remove Legal Hold job.

    When the job runs, the system removes the legal hold from all records that are associated to the legal hold group and deletes the legal hold group.

Tags

You can tag a set of records in the Data Vault with a user-defined value and later retrieve the records based on the tag name. You can also base a retention policy on a tag value. A single record in the Data Vault can be part of more than one tag. A tag can be of date, numeric, or string datatype.

You can define a maximum of four date datatype tags, a maximum of four numeric datatype tags, and one string datatype tag in the Data Vault. You can use a maximum of 4056 characters for the string datatype tag. A single tag can have different tag values for different records. When you tag a record, you specify a value based on the datatype of the tag. You can update the value of a tag and remove the tags.

Adding Tags

Use the Data Discovery portal to add tags to your records.

1. Click **Data Discovery** > **Manage Tags**.
2. Select the archive folder and entity.

   The system displays the search elements that are defined for the entity.
3. Enter a condition for at least one of the search elements.

   Use the **Add (+)** or **Remove (-)** icons to add or remove multiple search conditions.
4. Click **View**.

   The records that match the search criteria appear.
5. Use the checkbox on the left to select the records you want to tag.
6. Click **Add Tag**.
7. Select a tag or choose to create one.

   - To select an existing tag, choose a name from the **Tag Name** drop-down list and enter a tag value.
   - To create a tag, enter a tag name, select the datatype, and enter a tag value.

8. Click **Add Tag**.

   The **Schedule Job** page for the Add_Tagging_Records job appears.
9. Select the appropriate schedule options and click **Schedule**.

   The **Monitor Jobs** page appears.
   After the Add Tagging Records job completes, you can click the expand option to view log details.

Updating Tags

Use the Data Discovery portal to update the tag value of a record.

1. Click **Data Discovery** > **Manage Tags**.
2. Select the archive folder and entity.

   The system displays the search elements that are defined for the entity.
3. Enter a condition for at least one of the search elements.

   Use the **Add (+)** or **Remove (-)** icons to add or remove multiple search conditions.
4. Click **View**.

   The records that match the search criteria appear.
5. Use the checkbox on the left to select one or more records with the tag you want to update.

6. Click **Update Tag**.

7. Select a tag from the **Tag Name** drop-down list.

8. Enter a value.

9. Click **Update Tag**.

   The **Schedule Job** page for the Update_Tagging_Records job appears.

10. Select the appropriate schedule options and click **Schedule**.

    The **Monitor Jobs** page appears.
    After the Update Tagging Records job completes, you can click the expand option to view log details.

Removing Tags

Use the Data Discovery portal to remove a tag from a record.

1. Click **Data Discovery** > **Manage Tags**.

2. Select the archive folder and entity.

   The system displays the search elements that are defined for the entity.

3. Enter a condition for at least one of the search elements.

   Use the **Add (+)** or **Remove (-)** icons to add or remove multiple search conditions.

4. Click **View**.

   The records that match the search criteria appear.

5. Use the checkbox on the left to select one or more records with the tag you want to remove.

6. Click **Remove Tag**.

7. Select a tag from **Tag Name** drop-down list.

8. Click **Remove Tag**.

   The **Schedule Job** page for the Remove_Tagging_Records job appears.

9. Select the appropriate schedule options and click **Schedule**.

   The **Monitor Jobs** page appears.
   After the Remove Tagging Records job completes, you can click the expand option to view log details.

Searching Data Vault Using Tags

Use tag values as filter options when you search for records from the Data Vault.

1. Click **Data Discovery** > **Search Data Vault**.

   The **Search Data Vault** page appears.

2. Select the archive folder and entity.

   The system displays the search elements that are defined for the entity.

3. Select a tag element from the drop-down list of search elements and enter a condition.

   If you do not see tag columns in the drop-down list of search elements, contact the Data Archive administrator for permissions.

   Use the **Add (+)** or **Remove (-)** icons to add or remove multiple search conditions.

4. Click **View**.

    The records that match the search criteria appear.

    Tag columns are identified by a yellow icon in the column heading. You can export all or a select set of records.

Browsing Data Using Tags

Use tag values as filter and sort options when you browse for records in the database.

1. Click **Data Discovery** > **Browse Data**.

    The **Browse Data** page appears.

2. Select the archive folder, schema, and table. You may also select the entity and the number of rows to display.

    The system populates the **Available Columns** list.

    Tag columns appear as ilm_metafield1 to ilm_metafield9.

    If you do not see the tag columns, contact the Data Archive administrator for permissions.

3. Use the arrows to move the tag and any other column to the **Display Columns** list.

4. Optionally, specify the **Where Clause** and **Order By** to further filter and sort the search results.

5. Click **Preview SQL** to view the generated SQL query.

6. Click **Search**.

    The records that match the search criteria appear. Tag columns are identified by a yellow icon in the column heading.

    A **Click to View** link appears if the data in a record has more than 25 characters or if the record has an attached file.

Troubleshooting the Data Discovery Portal

The View button in the Search Data Vault window is always disabled.

The **View** button might become permanently disabled because of an enhanced security setting in the web browser. To enable the button, close Data Archive, disable enhanced security configuration for the web browser, and restart Data Archive.

For example, to disable enhanced security configuration for Microsoft Internet Explorer on Windows Server 2008, perform the following steps:

1. Log in to the computer with a user account that is a member of the local Administrators group.

2. Click **Start** > **Administrative Tools** > **Server Manager**.

3. If the **User Account Control** dialog box appears, confirm that the action it displays is what you want, and then click **Continue**.

4. Under **Security Information**, click **Configure IE ESC**.

5. Under **Administrators**, click **Off.**

6. Under **Users**, click **Off.**

7. Click **OK**.

8.   Restart Microsoft Internet Explorer.

For more information about browser enhanced security configuration, see the browser documentation.

I cannot view a logical entity in the Data Discovery portal.

To resolve the issue, verify the following items:

- The Data Vault Service is running.

- The entity is assigned to a Data Vault access role and your user is assigned to the same role. If your user is not assigned to the same role as the entity, you will not see the entity in the Data Discovery portal.

- The driving table is set for the entity.

- The metadata source for the Data Discovery portal searches is set to the correct source. In the `conf.properties` file, the informia.dataDiscoveryMetadataQuery property is set to `AMHOME`.

- For tables with exotic columns, the primary key constraint for the table is defined.

The technical view does not show child tables.

To resolve the issue, verify the following items:

- All of the constraints are imported for the entity in the Enterprise Data Manager.

  Tip: View the entity constraints in the **Data Vault** tab.

- The metadata source for the Data Discovery portal searches is set to the correct source. In the `conf.properties` file, verify that the informia.dataDiscoveryMetadataQuery property is set to `AMHOME`.

Non-English characters do not display correctly from the Data Discovery portal or any 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. To verify what type of encoding the operating system has, run the following command:

```
local -a
```

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`.

The technical view for an entity returns a class cast exception.

Technical view is not available for entities with binary or varbinary data types as primary, foreign, or unique key constraints.

CHAPTER 13

Data Visualization

This chapter includes the following topics:

Data Visualization Overview

Data visualization is a self-service reporting tool available on the Data Archive user interface.

Use data visualization to create, view, copy, run, and delete reports from data that is archived to the Data Vault. You can select data from related or unrelated tables for a report. You can also create relationships between tables and view a diagram of these relationships. You can then lay out, style, and format the report with the design tools available on the data visualization interface.

When you access the **Reports and Dashboards** page, you can also view any pre-packaged reports that were installed by an accelerator. Accelerator reports have the type "TEMPLATE_REPORT" in the **Reports and Dashboards** window. You can perform only copy and grant operations on the TEMPLATE_REPORT type of report. You must copy the accelerator reports to an archive folder that corresponds to the Data Vault connection where the retired tables on which the reports are built are located. For example, when you copy the reports from the Application Retirement for Healthcare accelerator, select an archive folder where the healthcare-related tables are archived.

To create highly designed, pixel-perfect reports, install and use the Designer application on your desktop. The Designer application contains an advanced design tool box. After you create a report in Designer, you can publish it to the Data Archive server to view the report on the data visualization interface.

Data Visualization User Interface

You access data visualization from the **Data Visualization** menu option. The **Reports and Dashboards** window displays a list of reports with sort and filter options.

The following figures show the **Reports and Dashboards** window:

1. List of available reports. Click a row to select the report you plan to run. Click one or more checkboxes to select reports you want to delete.

2. Filter box. Enter the entire or part of the value in the filter box above the appropriate column heading and click the filter icon.

3. Actions drop-down menu. Contains options to run, create, edit, and copy reports. Also contains the options to add permissions or revoke permissions for selected reports.

4. Filter icon. After you specify filter criteria in the fields above the column names, click the filter icon to display the reports.

5. Clear filter icon. Click this icon to clear any filter settings.

6. Sort icon. Click this icon to sort columns alphabetically or chronologically.

7. Resizing bar. Click and drag the bar up and down to resize the sections.

8. Properties section. Displays report details.

9. Permissions icon. Click the Permissions icon to switch to the permissions section from the properties section.

10. Permissions section. The permissions section displays which users and access roles have permission to take an action on a report. This information is only visible to you if you have the "grant" permission on your login user or access role for the selected report. If you do not have the grant permission, you will only see the permissions information for your own user and access role.

You can also launch the Data Visualization Advanced Reporting interface from the **Data Visualization** menu. You can use the Advanced Reporting interface to run reports. To access the Advanced Reporting interface, you must run the chosen report with a user that has the Report Admin role.

Note: You can see the **Data Visualization** menu option if your product license includes data visualization.

Data Visualization Process

Complete the following tasks to create a report:

1. Run a retirement job or an archive job so that data exists in the Data Vault.

2. Create a report by either entering an SQL query or selecting tables from the report wizard.

3. Optionally, create relationships between tables.

4. Design the report layout and style.

5. Run the report.

Report Creation

You can create reports with tables, cross tabs, and charts. Then add parameter filters, page controls, labels, images, and videos to the report.

For example, you might create a tabular report that displays the customer ID, name, and last order date for all retired customer records. You can add filters to allow the user to select customer-status levels. Or, you might create a pie chart that shows the percentage of archived customer records that have the platinum, gold, or silver status level.

You create a report by specifying the data you want in the report. You can specify the data for the report in one of the following ways:

- Select the tables from the report-creation wizard. If required, create relationships between fields within and between tables.

- Enter an SQL query with the schema and table details.

After you specify the data for a report, design, save, and run the report.

Table and Column Names

When a table or column name begins with a number, Data Archive appends the prefix INFA to the name when the name appears in the report creation wizard.

For example, if the column name is `34_REFTABLE_IVARCHAR`, then the column name appears as `INFA_34_REFTABLE_IVARCHAR` in the report creation wizard.

When a column name contains special characters, Data Archive replaces the special characters with an underscore when the column name appears in the report creation wizard.

For example, if the column name is `Order$Rate`, then the column name appears as `Order_Rate` in the report creation wizard.

Creating a Report from Tables

To specify data for a report, select the tables in an archive folder to include in the report. Each report you create can contain data from one archive folder.

1. Click **Data Visualization** > **Reports and Dashboards**.

   The **Reports and Dashboards** window appears.

2. Click **Actions** and select **New** > **Report**.

   The **New Report** tab with the **Create Report: Step 1 of 3 Step(s)** wizard appears.

3. Select a target connection from the **Archive Folder Name** drop-down list.

   The archive folder list displays archive folders with the same file access role that you have.

4. Select **Table(s)** and click **Next**.

   The **Add Tables** page appears.

5. Select the tables containing the data that you want in the report. Optionally, filter and sort schema, table, or data object name to find the tables. Click **Next**.

The **Create Report: Step 2 of 3 Steps(s)** page appears with a list of the selected tables.

6. Enter a name in the **Data Object Name** field.

7. Click **Add More** or **Remove Selected** to add or remove tables from the list.

8. Click **Add constraints**.

    The **Relationship Map** page appears. If necessary, click and drag the tables to get a clearer view of the relationship arrows. You can also zoom in and out for clarity.

9. Click **Add Relationship**.

    A list of the column names in each table you selected, appears in a child table section and in a parent table section. Arrows indicate the relationships between columns.

10. To create a relationship, select a column from Child Table Columns list and a column from the Parent

    Table Columns list. Click the **Join** button .

Links appear between the table columns to show relationships.

11.
To remove a relationship, click the relationship arrow and click the **Remove Join** button .

12. Click **OK**.

The **Add Tables** page appears.

13. Click **Add Constraints** to view the updated relationship map. Click **Close**.

The **Add Tables** page appears.

14. Click **Next**.

The **Create Report: Step 3 of 3 Step(s)** window appears. This is the design phase of the report-creation process.

Creating a Report From an SQL Query

To specify data for a new report you can write an SQL query that contains the archive folder and tables to include in the report. Each report you create can contain data from one archive folder.

1. Click **Data Visualization** > **Reports and Dashboards**.

The **Reports and Dashboards** window appears.

2. Click **Actions** and select **New** > **Report**.

The **New Report** tab with the **Create Report: Step 1 of 3 Step(s)** wizard appears.

3. Select a target connection from the **Archive Folder Name** drop-down list.

   The archive folder list displays archive folders with the same file access role that you have.

4. Select **SQL Query** and click **Next**.

5. Enter a name for the SQL query in the **Query Name** field.

6. Enter a query and click **Validate**.

   If you see a validation error, correct the SQL query and click **Validate**.

7. Click **Next**.

   The **Create Report: Step 3 of 3 Step(s)** window appears. This is the design phase of the report-creation process.

Parameter Filters

You can add parameter filters to a report to make it more interactive. Use a parameter filter to show records with the same parameter value. To add parameter filters on a report, you must insert a prompt in the SQL query you use to create the report.

Add a parameter prompt for each parameter you want to filter on. For example, you want to create a report based on employee details that includes the manager and department ID. You want to give users the flexibility to display a list of employees reporting to any given manager or department. To add filters for manager and department IDs, you must use placeholder parameters in the SQL query for both the manager and department ID variables.

The placeholder parameter must be in the following format: `@P_<DataType>_<Prompt>`

\<DataType\>

Specify a datatype. Use one of the following values:

- Integer
- Number
- String
- Decimal

If you do not specify a datatype, the parameter will be assigned the default datatype, string.

Note: Date, time, and datetime data types must be converted to string datatypes in the SQL query. Convert data with a date, time, or datetime datatype to a string datatype in one of the following ways:

- Use the format `@P_<Prompt>` where the `<DataType>` value is not specified. The parameter values will be treated as string by default.
- Add the `char()>` function to convert the parameter to string. The complete format is: `char(<column name>)>@P_<DataType>_<Prompt>`

\<Prompt\>

Represents the label for the parameter filter. Do not use spaces or special characters in the `<Prompt>` value.

Example

The SQL query to add parameter filters for manager and department IDs has the following format: `select * from employees where manager_id =@P_Integer_MgrId and Department_id =@P_Integer_DeptId`

When you run a report, you can choose to filter data by manager ID and department ID. The following image shows the filter lists for **MgrId** and **DeptId**:

The report displays details of employees associated with the manager and department you selected.

You will also see the **MgrId** and **DeptId** filters in the Parameters section on the **Report** page.

LIKE Operators

For information about using the LIKE operator as a parameter filter, see the *Informatica Data Vault SQL Reference*.

Example SQL Query to Create a Report

The following example shows an sql query to create a report:

```
SELECT hp.party_name
      ,hc.cust_account_id
      ,hc.account_number         customer_number
      ,amount_due_remaining       amount_due_remaining
      ,aps.amount_due_original   amount_due_original
      ,ra.trx_number
  FROM apps.ra_customer_trx_all       ra,
       apps.ra_customer_trx_lines_all rl,
       apps.ar_payment_schedules_all  aps,
       apps.ra_cust_trx_types_all     rt,
       apps.hz_cust_accounts          hc,
       apps.hz_parties                hp,
       apps.hz_cust_acct_sites_all    hcasa_bill,
       apps.hz_cust_site_uses_all     hcsua_bill,
       apps.hz_party_sites            hps_bill,
       apps.ra_cust_trx_line_gl_dist_all rct
 WHERE ra.customer_trx_id       = rl.customer_trx_id
   AND ra.customer_trx_id       = aps.customer_trx_id
   AND ra.org_id                = aps.org_id
   AND rct.customer_trx_id      = aps.customer_trx_id
   AND rct.customer_trx_id      = ra.customer_trx_id
```

```
AND rct.customer_trx_id          = rl.customer_trx_id
AND rct.customer_trx_line_id     = rl.customer_trx_line_id
AND ra.complete_flag             = 'Y'
AND rl.line_type                 IN ('FREIGHT', 'LINE')
AND ra.cust_trx_type_id          = rt.cust_trx_type_id
AND ra.bill_to_customer_id       = hc.cust_account_id
AND hc.status                    = 'A'
AND hp.party_id                  = hc.party_id
AND hcasa_bill.cust_account_id   = ra.bill_to_customer_id
AND hcasa_bill.cust_acct_site_id = hcsua_bill.cust_acct_site_id
AND hcsua_bill.site_use_code     = 'BILL_TO'
AND hcsua_bill.site_use_id       = ra.bill_to_site_use_id
AND hps_bill.party_site_id       = hcasa_bill.party_site_id
AND hcasa_bill.status            = 'A'
AND hcsua_bill.status            = 'A'
AND aps.amount_due_remaining     <> 0
AND aps.status                   = 'OP'
```

Designing the Report

After you select the data for the report, continue with the Create Report wizard to select the page type, layout, and style for your report. For more information on designing a report, click the help button (🔵) on the top right corner of the wizard to open online help.

1. Select a predefined page template on the **Page** page of the wizard. To customize page size and orientation, click **Page Setup** and enter specifications. Click **Next**.

    The **Layout** page appears.

2. Select a layout for your report. You can customize the layout in the following ways:

    - To split a cell, select the cell and click either the **Horizontal Split** and **Vertical Split**.

- To merge cells, select two or more adjacent cells while holding down the **Ctrl** key and click **Merge**.
- To align the components in each cell, select the cell and click **Align**.
- To resize a cell, hold and drag the cell's borders. The width and height of each cell is displayed below.

3. Select a component to add to each cell.

   Component options include table, crosstab, or chart. You can also choose to leave a cell blank.

4. Click **Next**.

   The **Bind Data** page appears.

5. Specify the data source and fields for the first component in the report.

6. Click **Next** to specify the data in the next component.

   After you specify data for all the components you selected on the **Layout** page, the **Style** page appears.

7. Select a style template.

8. Click **Save**.

   The **Save As** page appears.
   When you save the report, you must save it in the same archive folder that you selected when you created the report. If you select a different archive folder than the one you used to create the report, you receive an internal error when you try to run the report.

9. Complete the fields on the **Save As** page and click **Save**.

   The report is saved and displayed on the **Reports and Dashboards** window.

10. Click **Run** to run the report.

Report Permissions

Before you can run, delete, or copy a report, you must have the corresponding permission. If you have the grant permission, you can also grant the ability to run, copy, or delete a report to other users or access roles.

The run, copy, delete, and grant permissions are separate permissions. Every report, whether created in Data Visualization or imported from an accelerator, has permissions that can be granted or revoked for only that report. Each permission can be granted to a specific user or to a Data Vault access role. Each user or access role assigned to a report can have different permissions.

The run, delete, and copy permissions allow a user or access role to take those actions on an individual report or multiple reports at one time. If your user or access role has the grant permission, you can also assign the run, copy, and delete permissions to multiple users or access roles for a single report or multiple reports.

You can add or delete permissions for one report at a time, or you can add or delete permissions for multiple reports at one time.

By default, the Report Admin user has permission to run, copy, delete, and grant permissions for all reports. If you do not have any permissions on a report, then the report is not visible to you in the Reports and Dashboards window.

Assigning a User or Access Role to a Report

To grant a user or access role permissions on a report, assign the user or access role to the report. When you assign the user or access role to the report, you also specify which permissions the user or access role has on the report. To be able to assign a user to a report and grant permissions to the user, you must have the grant permission applied to your login user or access role.

1. Click **Data Visualization** > **Reports and Dashboards**.

    The **Reports and Dashboards** window appears.

2. In the upper pane, select the name of the report that you want to assign a user or access role to.

    Do not select the check box next to the report name. Select the name of the report to highlight the report row.

3. In the lower pane, click **Permissions**.

4. To assign an access role to the report, click **Edit** in the **Access Roles** tab. To assign a user to the report, click **Edit** in the **Users** tab.

    The **Edit Report Permissions** window appears. This window lists the users and access roles that are already assigned to the report.

5. Click **Add**.

    The **Add Access Roles** or **Add Users** window appears.

6. From the list of available users or access roles, select the check box next to the user or access role that you want to add. Users and roles that are already assigned to the report do not appear.

7. In the **Permissions** box, select the permissions that you want to grant to the user or access role.

8. Click **OK**.

    The user or access role, along with the permissions you granted, appear in the lower pane of the **Reports and Dashboards** screen.

Editing Report Permissions for a User or Access Role

If you have the grant permission assigned to your user or access role, you can edit the permissions for other users and access roles.

1. Click **Data Visualization** > **Reports and Dashboards**.

    The **Reports and Dashboards** window appears.

2. In the upper pane, select the name of the report that you want to edit permissions for.

    Do not select the check box next to the report name. Select the name of the report to highlight the report row.

3. In the lower pane, click **Permissions**.

4. To edit permissions for an access role, click **Edit** in the **Access Roles** tab. To edit permissions for a user, click **Edit** in the **Users** tab.

    The **Edit Report Permissions** window appears.

5. From the list of assigned users or access roles, select the user or access role that you want to edit permissions for.

6. Clear or select the check box next to each permission that you want to edit.

To remove the delete and copy permissions, you must first remove the grant permission. To remove the run permission, you must first remove the grant, delete, and copy permissions.

7. To save your selections for an individual user or access role, click the **Checkmark** icon. To cancel without changing any permissions, click the **X** icon.

8. To save all of the permission changes, click **Save** at the bottom of the window.

The edited permissions appear next to the user or access role in the lower pane of the **Reports and Dashboards** window.

Deleting a User or Access Role From a Report

If you have the grant permission assigned to your user or access role, you can delete a user or access role from a report to remove all of the permissions that are granted to the user or access role.

1. Click **Data Visualization** > **Reports and Dashboards**.

The **Reports and Dashboards** window appears.

2. In the upper pane, select the name of the report that you want to delete a user or access role from.

Do not select the check box next to the report name. Select the name of the report to highlight the report row.

3. In the lower pane, click **Permissions**.

4. To delete an access role from the report, click **Edit** in the **Access Roles** table. To delete a user from the report, click **Edit** in the **Users** table.

The **Edit Report Permissions** window appears.

5. Select the check box next to the user or access role that you want to delete.

6. Click **Delete**.

The user or access role disappears from the list of assigned users and access roles on the **Reports and Dashboards** page.

Assigning Permissions for Multiple Reports

You can assign the permissions for multiple reports at one time. When you assign the permissions for multiple reports, you select a user or access role and then select the permissions for that user or access role. Your selections are applied to all of the reports. To grant permissions to multiple reports, you must have the grant permission applied to your login user or access role for all of the selected reports.

1. Click **Data Visualization** > **Reports and Dashboards**.

The **Reports and Dashboards** window appears.

2. In the upper pane, select the check box next to each report that you want to edit permissions for. Alternatively, you can select the check box next to the name of an archive folder to select all of the reports in the folder.

3. Click **Actions** > **Add Permissions**.

The **Manage Permissions on Reports** window appears.

4. To assign permissions for an access role, select the **Access Roles** tab. To assign permissions for a user, select the **Users** tab.

5. Select the check box next to the user name or access role that you want to assign permissions to.

6. In the **Permissions** box, select the permissions that you want to grant to the user or access role.

   The permissions that you apply are in addition to any existing permissions for the user or access role. For example, if access role "Role1" previously had the run and copy permissions, and you select only the delete permission in the Permissions box, Role1 retains the run and copy permissions and gains the delete permission.

7. Click **OK**.

   The user or access role along with the associated permissions appear in the lower pane of the **Reports and Dashboards** window for each report that you selected.

Deleting Permissions for Multiple Reports

You can delete permissions for multiple reports at one time. When you delete permissions on multiple reports, you remove all users and access roles assigned to the reports. Only the Report Admin role can delete permissions for multiple reports at one time.

1. Click **Data Visualization** > **Reports and Dashboards**.

   The **Reports and Dashboards** window appears.

2. In the upper pane, select the check box next to each report that you want to edit permissions for. Alternatively, you can select the check box next to the name of an archive folder to select all of the reports in the folder.

3. Click **Actions** > **Revoke All Permissions**.

   A warning appears to ask if you are certain that you want to revoke all permissions on the selected reports.

4. To revoke all permissions on the reports, select **Yes**. To cancel, select **No**.

Running and Exporting a Report

You can run a report after you create it. You can also export the report to a file such as a PDF or Microsoft Excel file.

1. Click **Data Visualization** > **Reports and Dashboards**.

   The **Reports and Dashboards** window appears with a list of reports sorted by report name.

2. Optionally, sort reports by type, description, creator, created date, or the last modification date to locate the report.

3. Optionally, filter on one or more columns to limit the list of reports displayed.

4. Click the row of the report you want to run.

   Do not use the checkbox to select a report to run.

The row is highlighted.

5. Click **Actions** > **Run Report**.

    The report appears in view mode.

6. You can complete the following tasks in view mode:

    a. Click the **Export** button to export the report to a file.

        You can export the report to one of multiple file formats such as PDF, HTML, Microsoft Excel, Text, RTF, or XML.

    b. Click the **Print** button.

    c. Click the **Filter** button and select or exclude variables or certain values in a variable for each component in the report.

    d. Click the **Save** button to save any changes you made.

    e. From the **Parameters** view, select values to display in the report.

        You will see the **Parameters** view if the report was created with an SQL query that specified parameter filters.

    f. From the **Filter** view, click the **+** button. Select fields to filter data in the report.

7. Click **Edit Mode** to edit the report.

The report opens in edit mode.

8. You can complete the following tasks in edit mode:

a. All tasks allowed in view mode.

b. From the **Resources** view, expand **Data Source** > report name > column name. Drag a column name to the appropriate area to add a new field to a table, crosstab, or chart.

c. From the **Resources** view, expand **Data Source** > report name > **Dynamic Resource** > **Aggregations**. Click **<Add Aggregations...>** and select an option such as count, sum, or average to add to a table or chart.

d. From the **Resources** view, expand **Data Source** > report name > **Dynamic Resource** > **Formulas** > **Add Formula** and add a customized formula to your table or chart.

e. From the **Filter** view, click the **+** button. Select fields to filter data in the report.

f. From the **Components** view, drag a component to an area in the report.

You can add components such as a table, crosstab, chart, label, image, and video.

For more information on how to edit your report, click **Menu** > **Help** > **User's Guide**.

Deleting a Report

If you have the delete permission, you can delete reports from the data visualization user interface.

1. From the main menu, select **Data Visualization** > **Reports and Dashboards**.

   The **Reports and Dashboards** window appears with a list of reports.

2. Select one or more reports to delete.

3. From the **Actions** drop-down list, select **Delete Reports**.

   A confirmation message appears. Click **Yes** to delete the reports.

Copying Reports

If you have the copy permission, you can copy the reports in a folder to another folder.

To copy a report folder from one archive folder to another, the target archive folder must meet the following requirements:

- The target archive folder must contain all the tables required for the report.
- The tables for the report must be in the same schema in the target archive folder.

The copy functionality works at the folder level. Select all the reports in the folder or select the folder to enable the copy action. You cannot select an individual report to copy if the folder contains other reports.

1. From the main menu, select **Data Visualization** > **Reports and Dashboards**.

The **Reports and Dashboards** page appears with a list of reports within each folder.

2. Select a report folder that contains the reports that you want to copy.

3. Click **Actions** > **Save As**.

The **Save As** dialog box appears.

4. Select an archive folder to copy the report folder to.

5. Enter a name for the new folder.

6. Click **OK**.

    The new folder with the copied reports appears on the **Reports and Dashboards** page.

Copying SAP Application Retirement Reports

To copy SAP application reports, you save the folder containing the reports to a new location. Then, publish details for the new schema from Data Visualization Designer to Data Archive.

Before you copy reports based on a retired SAP application, perform the following tasks:

-   Install the SAP application accelerator.

-   Install Data Visualization Designer.

1. From the main menu, select **Data Visualization** > **Reports and Dashboards**.

    The **Reports and Dashboards** window appears with a list of reports within each folder.

2. Enable the check box for the SAP report folder.

Data Archive selects all the reports in the SAP folder.

3. Click **Actions** > **Save As**.

The **Save As** dialog box appears.

4. Enter values for the following parameters:

Current Folder

Name of the report folder.

Save to Archive Folder

Target folder where the tables for the reports reside.

New Folder Name

Name for the folder that you want to copy the reports to.

5. Click **OK**.

Data Archive copies the reports to the new folder.

6. From Data Archive, click **Data Visualization** > **Reports and Dashboards**.

7. Expand the folder that contains the copied SAP reports.

8. Run the file `Home_Page.cls`.

The **Run Report:Home_Page.cls** window appears.

The `Home_Page.cls` file runs all the reports in the folder.

Troubleshooting

The following list describes some common issues you might encounter when you use Data Visualization.

You want to create a report but do not see the tables that you need in the selection list.

Contact your administrator and ensure that you have the same file access role as the data that you require for the report.

You encountered a problem when creating or running a report and want to view the log file.

The name of the log file is `applimation.log`. To access the log file, go to **Jobs > Job Log** on the Data Archive user interface.

When you click any button on the Data Visualization page, you see the error message, "HTTP Status 404."

You cannot access Data Visualization if you installed Data Visualization on one or more machines but did not configure the machine to load it. Contact your Data Archive administrator.

You created a new user with the Report Designer role, but when you access Reports and Dashboards you receive an HTTP error.

Data Archive updates new users according to the value of the user cache timeout parameter in the Data Archive system profile. The default value of this parameter is five minutes. You can wait five minutes for the user cache to refresh and then access the **Reports and Dashboards** page. If you do not want to wait five minutes, or if a user assignment has not updated after five minutes, review the user cache timeout

parameter in the Data Visualization section of the system profile. For more information about the Data Archive system profile, see the *Informatica Data Archive Administrator Guide*.

A user can run a report even if the user does not have the Data Archive access roles for the tables in the report.

Run the report as a user with a Report Designer role. When you run the report with the Report Designer role, Data Archive activates the security levels.

1. Log in to Data Archive as a user with the Report Designer role.

2. Click **Data Visualization** > **Reports and Dashboard**.

3. Select the report.

4. Click **Actions** > **Run Report**.

Data in a report does not appear masked although the data appears masked in Data Discovery search results.

Run the report as a user with a Report Designer role. When you run the report with the Report Designer role, Data Archive activates the security levels.

1. Log in to Data Archive as a user with the Report Designer role.

2. Click **Data Visualization** > **Reports and Dashboard**.

3. Select the report.

4. Click **Actions** > **Run Report**.

You cannot create a Data Visualization report for entity reference tables.

Data Visualization reports for entity reference tables are not supported.

You cannot copy a report folder to another archive folder.

You might not be able to copy a report folder from one archive folder to another because of one of the following reasons:

*   The target archive folder does not contain one or more tables required for the report.

*   The tables required for the report are not in the same schema.

Select another archive folder to save the report folder to.

The following table describes the outcome when you copy a report folder from one archive folder to another under different scenarios:

Scenario	Tables in the Archive Folder You are Copying From	Tables in the Archive Folder You are Copying To	Result
1	UF2.MARA UF2.MARC UF2.BKPF	SAP.MARA SAP.MARC SAP.BKPF	Success. Data Archive copies the report folder to the target archive folder.
2	UF2.MARA UF2.MARC UF2.BKPF	SAP.MARA SAP.MARC	Error. Data Archive cannot copy the report folder because the target archive folder does not contain the table UF2.BKPF.
3	UF2.MARA UF2.MARC UF2.BKPF	SAP.MARA SAP.MARC DBO.BKPF	Error. Data Archive cannot copy the report folder because the tables in the target archive folder are spread across two different schemas, the SAP and the DBO schemas.

Designer Application

Data Visualization Designer is a stand-alone application that you can install on your desktop. Use the Designer application to create pixel-perfect page and web reports and then publish these reports to the Data Archive server. Reports on the Data Archive server appear on the data visualization interface and can be viewed and executed by you or other users.

Reports created in the Designer application can be edited only in Designer. Reports created on the data visualization interface can be edited on the interface or saved to a local machine and edited in Designer.

For information on how to use Designer, see the *Data Archive Data Visualization Designer User Guide*.

Oracle E-Business Suite Retirement Reports

This chapter includes the following topics:

Retirement Reports Overview

If you have a Data Visualization license, you can generate reports for certain application modules after you have retired the module data to the Data Vault. Run the reports to review details about the application data that you retired.

You can generate the reports in the Data Visualization area of Data Archive. Each report contains multiple search parameters, so that you can search for specific data within the retired entities. For example, you want to review payment activities related to a specific invoice for a certain supplier and supplier site. You can run the Invoice History Details report and select a specific value for the supplier and supplier type.

When you install the accelerator, the installer imports metadata for the reports as an entity within the appropriate application module. Each report entity contains an interim table and potentially multiple tables and views. You can view the report tables, views, and their columns and constraints in the Enterprise Data Manager. You can also use the report entity for Data Discovery.

Some reports also contain user-defined views. The user-defined views are required for some reports that include tables that have existed in different schemas in different versions of Oracle E-Business Suite. Some of the user-defined views are required because of the complexity of the queries that generate a single report. Finally, some of the user-defined views are required because some of the queries that generate the reports refer to package functions that Data Vault does not support.

You can run a script to create a user-defined view that is required for a report. Contact Informatica Global Customer Service to acquire the scripts for the user-defined views in each report.

Prerequisites

Before you can run a Data Visualization report, you must first retire the Oracle E-Business Suite application.

To retire a Oracle E-Business Suite application, perform the following high-level tasks:

1. In the Enterprise Data Manager, create a customer-defined application version.

2. Import metadata from the Oracle E-Business Suite application to the customer-defined application version, along with user-defined views.

3. Create retirement entities. Use the generate retirement entity wizard to automatically generate the retirement entities.

4. In the Data Archive user interface, create a source connection. For the connection type, choose the database in which the Oracle E-Business Suite application is installed.

5. Create a target connection to the Data Vault.

6. Define a new data type mapping to handle VARCHAR2(0) length columns that exist in the views.

    a. In the Data Archive user interface, click **Administration** > **Data Type Mappings for File Archive**.

    b. Click **New Data Type Mapping**.

    c. For Connection Type, select All Connections.

    d. For Data Type, select VARCHAR2.

    e. For Operator, select Equals.

    f. For Length, enter 0.

    g. For Precision and Scale, enter 0,0.

    h. For File Archive Data Type, select VARCHAR.

    i. For Source Length, enter 40.

    j. Click **Save**.

7. Run the create archive folder standalone job to create archive folders in the Data Vault.

8. Create and run a retirement project.

9. In the Enterprise Data Manager, create constraints for the imported source metadata in the ILM repository. Constraints are required for Data Discovery portal searches. The Oracle E-Business Suite application retirement accelerator includes constraints for some tables in the Oracle E-Business Suite system. You might need to define additional constraints for tables that are not included in the accelerator.

10. Copy the entities from the pre-packaged Oracle E-Business Suite application version to the customer-defined application version. To copy the entities, perform the following high-level steps:

    a. Right-click the customer-defined application version where the metadata is imported and select **Copy entities from Application Version**. The **Copy Entities from Application Version** window appears.

    b. Select the application version from where you need to copy, which is typically Oracle Applications Retirement 1.0.

    c. Provide the prefix that will be appended before the entity while copying. The Copy All Entities option is not required.

    d. Click **OK**.

    e. After you submit the background job, you can view the job status from the Data Archive user interface. Once the job successfully completes, you can view the copied entities in your customer-defined application version. You can use these entities for Data Discovery.

11. Before you run the reports, copy the report templates to the folder that corresponds to the Data Vault connection. You can use these pre-packaged Data Visualization reports to access retired data in the Data Vault. When you copy the reports, the system updates the report schemas based on the connection details.

12. Optionally, validate the retired data.

13. After you retire and validate the application, use the Data Discovery Portal, Data Visualization reports, or third-party query tools to access the retired data.

Accounts Receivable Reports

In the Accounts Receivable module, you can run the following reports:

- Customer Listing Detail
- Adjustment Register

Customer Listing Detail Report

Use the Customer Listing Detail report to review summary information about your customers. You can view customer name, customer number, status, and any addresses or site uses you entered for a customer.

Report Parameters

The following table describes the report input parameters:

Parameter	Default Value	Mandatory
Set of Books	<Any>	Yes
Operating Unit	<Any>	Yes
Order By	Customer Name	No
Customer Name	<Any>	No
Customer Number Low	<Any>	No
Customer Number High	<Any>	No

Report Tables and Views

The following tables and views are included in the report:

Tables

- HZ_CUSTOMER_PROFILES
- HZ_CUST_ACCOUNTS
- HZ_CUST_ACCT_SITES_ALL
- HZ_CUST_PROFILE_CLASSES

- HZ_PARTIES
- HZ_PERSON_PROFILES

Views

- OE_LOOKUPS_115
- OE_PRICE_LISTS_115_VL
- OE_TRANSACTION_TYPES_VL
- ORG_FREIGHT_VL
- ORG_ORGANIZATION_DEFINITIONS
- AR_LOOKUPS
- RA_SALESREPS

User-Defined Views

- V_REPORT_CU_ADDRESS
- V_REPORT_CU_AD_CONTACTS
- V_REPORT_CU_AD_PHONE
- V_REPORT_CU_AD_BUSPURPOSE
- V_REPORT_CU_AD_CON_ROLE
- V_REPORT_CU_AD_CON_PHONE
- V_REPORT_CU_BANKACCOUNTS
- V_REPORT_CU_BUS_BANKACCOUNT
- V_REPORT_CU_BUS_PYMNT_MTHD
- V_REPORT_CU_CONTACTS
- V_REPORT_CU_CONTACT_PHONE
- V_REPORT_CU_CONTACT_ROLES
- V_REPORT_CU_PHONES
- V_REPORT_CU_PYMNT_MTHDS
- V_REPORT_CU_RELATIONS
- V_REPORT_GL_SETS_OF_BOOKS

Adjustment Register Report

Use the Adjustment Register report to review approved adjustments by document number. Adjustments include manual adjustments, automatic adjustments, invoices applied to commitments, and credit memos applied to commitment-related invoices. This report groups and displays transactions by currency, postable status, document sequence name, and balancing segment.

Report Parameters

The following table describes the report input parameters:

Parameter	Default Value	Mandatory
Chart of Accounts	-	Yes
Operating Unit	<Any>	Yes
Set of Books	<Any>	Yes
Company Segment Value	<Any>	No

Report Tables and Views

The following tables and views are included in the report:

Tables

- GL_CODE_COMBINATIONS

Views

- ORG_ORGANIZATION_DEFINITIONS
- FND_FLEX_VALUES_VL
- FND_ID_FLEX_SEGMENTS_VL

User-Defined Views

- V_REPORT_GL_SETS_OF_BOOKS
- V_REPORT_ADJMNT_REGISTER

AR Transaction Detail Report

Use the AR Transaction Detail report to review all of the information you entered for invoices, credit memos, debit memos, chargebacks, guarantees, and deposits.

Report Parameters

The following table describes the report input parameters:

Parameter	Default Value	Mandatory
Operating Unit	<Any>	Yes
Set of Books	<Any>	Yes
Transaction Class	<Any>	No
Transaction Number Low	-	Yes
Transaction Number High	-	Yes

Report Views

The following views are included in the report:

Views

- ORG_ORGANIZATION_DEFINITIONS
- AR_LOOKUPS

User-Defined Views

- V_REPORT_GL_SETS_OF_BOOKS
- V_REPORT_AR_ACCTSETS
- V_REPORT_AR_FRTLINES
- V_REPORT_AR_LINES
- V_REPORT_AR_RELTRX
- V_REPORT_AR_REVACCTS
- V_REPORT_AR_SALESREPS
- V_REPORT_AR_TAXLINES
- V_REPORT_AR_TRNFLEX
- V_REPORT_TRANSACTIONS

Accounts Payable Reports

In the Accounts Payable module, you can run the following reports:

- Supplier Details
- Supplier History Payment Details
- Invoice History Details

Supplier Details Report

The Supplier Details report contains detailed information about supplier records. You can use this report to verify the accuracy of your current supplier information and to manage your master listing of supplier records.

The report displays detailed information for each supplier, and optionally, supplier site, including the user who created the supplier/site, creation date, pay group, payment terms, bank information, and other supplier or site information. You can sort the report by suppliers in alphabetic order, by supplier number, by the user who last updated the supplier record, or by the user who created the supplier record.

Report Parameters

The following table describes the report input parameters:

Parameter	Default Value	Mandatory
Set of Books	<Any>	Yes
Operating Unit	<Any>	Yes
Supplier Type	<Any>	No
Supplier Name	<Any>	No
Include Site Information	Yes	No
Include Contact Information	Yes	No
Include Bank Account Information	Yes	No
Order By	Supplier Name	No

Report Tables and Views

The following tables and views are included in the report:

Tables

- AP_BANK_ACCOUNTS_ALL
- AP_BANK_ACCOUNT_USES_ALL
- AP_TERMS_TL
- FND_USER
- AP_TOLERANCE_TEMPLATES

Views

- ORG_ORGANIZATION_DEFINITIONS
- PO_LOOKUP_CODES
- AP_LOOKUP_CODES
- FND_LOOKUP_VALUES_VL
- FND_TERRITORIES_VL

User-Defined Views

- V_REPORT_PO_VENDORS
- V_REPORT_PO_VENDOR_SITES_ALL
- V_REPORT_GL_SETS_OF_BOOKS
- V_REPORT_PO_VENDOR_CONTACTS

Supplier Payment History Details Report

Run the Supplier Payment History report to review the payment history for a supplier or a group of suppliers with the same supplier type.

You can submit this report by supplier or supplier type to review the payments that you made during a specified time range. The report displays totals of the payments made to each supplier, each supplier site, and all suppliers included in the report. If you choose to include invoice details, the report displays the invoice number, date, invoice amount, and amount paid.

The Supplier Payment History Details report also displays the void payments for a supplier site. The report does not include the amount of the void payment in the payment total for that supplier site. The report lists supplier payments alphabetically by supplier and site. You can order the report by payment amount, payment date, or payment number. The report displays payment amounts in the payment currency.

Report Parameters

The following table describes the report input parameters:

Parameter	Default Value	Mandatory
Set of Books	<Any>	Yes
Operating Unit	<Any>	Yes
Supplier Type	<Any>	No
Supplier Name	<Any>	No
Start Payment Date	System Date	Yes
End Payment Date	System Date	Yes
Invoice Detail	Yes	No
Order By Option	Payment Date	No

Report Tables and Views

The following tables and views are included in the report:

Tables

- AP_CHECKS_ALL
- AP_INVOICES_ALL
- AP_INVOICE_PAYMENTS_ALL
- FND_LOOKUP_VALUES

Views

- ORG_ORGANIZATION_DEFINITIONS

User-Defined Views

- V_REPORT_PO_VENDORS
- V_REPORT_PO_VENDOR_SITES_ALL

- V_REPORT_GL_SETS_OF_BOOKS

Invoice History Details Report

The Invoice History Details report includes information to support the balance due on an invoice.

The report generates a detailed list of all payment activities that are related to a specific invoice, such as gains, losses, and discounts. The report displays amounts in the payment currency.

Important: Payments must be accounted before the associated payment activities appear on the report.

Report Parameters

The following table describes the report input parameters:

Parameter	Default Value	Mandatory
Set of Books	<Any>	Yes
Operating Unit	<Any>	Yes
Supplier	<Any>	No
Supplier Site	<Any>	No
Pre-payments Only	YES	No
Invoice Number From	-	No
Invoice Number To	-	No
Invoice Date From	-	No
Invoice Date To	-	No

Report Views

The following views are included in the report:

Views

- ORG_ORGANIZATION_DEFINITIONS

User-Defined Views

- V_INVOICE_HIST_HDR
- V_INVOICE_HIST_DETAIL
- V_REPORT_PO_VENDORS
- V_REPORT_PO_VENDOR_SITES_ALL
- V_REPORT_GL_SETS_OF_BOOKS

Print Invoice Notice Report

Use the print invoice notice report to generate a standard invoice notice to send to a supplier to inform them about one or more invoices you have entered. The notice informs the supplier of outstanding credit or debit memos that you will apply to future invoices.

Report Parameters

The following table describes the report input parameters:

Parameter	Default Value	Mandatory
Invoice Date From	-	Yes
Supplier Type	\<Any\>	Yes
Supplier Name	\<Any\>	No
Pay Group	\<Any\>	No
Invoice Type	\<Any\>	No
Invoice Number	-	Yes
Name of Sender	-	No
Title of Sender	-	No
Phone of Sender	-	No

Report Tables and Views

The following tables and views are included in the report:

Tables

- PO_USAGES
- PO_LINE_LOCATIONS_ALL
- PO_LINES_ALL
- PO_HEADERS_ALL
- PO_DISTRIBUTIONS_ALL
- FND_LOOKUP_VALUES
- FND_DOCUMENT_CATEGORIES
- FND_DOCUMENTS_TL
- FND_DOCUMENTS_SHORT_TEXT
- FND_DOCUMENTS_LONG_TEXT
- FND_DOCUMENTS
- FND_ATTACHED_DOCUMENTS
- AP_INVOICES_ALL
- AP_INVOICE_DISTRIBUTIONS_ALL

Views

- PO_LOOKUP_CODES

User-Defined Views

- V_REPORT_PO_VENDORS
- V_REPORT_PO_VENDOR_SITES_ALL

Purchasing Reports

In the Purchasing module, you can run the following report:

- Purchase Order Details

Purchase Order Details Report

The Purchase Order Detail report lists all, specific standard, or planned purchase orders.

The report displays the quantity that you ordered and received, so that you can monitor the status of your purchase orders. You can also review the open purchase orders to determine how much you still have to receive and how much your supplier has already billed to you.

Report Parameters

The following table describes the report input parameters:

Parameter	Default Value	Mandatory
Set of Books	<Any>	Yes
Operating Unit	<Any>	Yes
Vendor From	<Any>	No
To	<Any>	No
PO Number From	-	No
To	-	No
Buyer Name	<Any>	No
Status	<Any>	No

Report Tables and Views

The following tables and views are included in the report:

Tables

- FND_CURRENCIES

- MTL_CATEGORIES_B
- MTL_SYSTEM_ITEMS_B
- PO_DOCUMENT_TYPES_ALL_TL
- PO_HEADERS_ALL
- PO_LINES_ALL
- PO_LINE_LOCATIONS_ALL

Views

- ORG_ORGANIZATION_DEFINITIONS
- PO_LOOKUP_CODES
- PER_PEOPLE_F

User-Defined Views

- V_PO_HDR_LINES
- V_PO_HDR_LKP_CODE
- V_REPORT_GL_SETS_OF_BOOKS
- V_REPORT_PO_VENDORS
- V_REPORT_PO_VENDOR_SITES_ALL

General Ledger Reports

In the General Ledger module, you can run the following reports:

- Chart of Accounts - Detail Listing
- Journal Batch Summary

Chart of Accounts - Detail Listing Report

Use the Chart of Accounts - Detail Listing report to review the detail and summary accounts defined for the chart of accounts.

This report has three sections. The first section is for enabled detail accounts, followed by disabled accounts, and then summary accounts. Each section is ordered by the balancing segment value. You can specify a range of accounts to include in your report. You can also sort your accounts by another segment in addition to your balancing segment.

Report Parameters

The following tables describes the report input parameters:

Parameter	Default Value	Mandatory
Chart of Accounts	US Federal Grade	Yes
Account From - Company	-	Yes

Parameter	Default Value	Mandatory
Account From - Department	-	Yes
Account From - Account	-	Yes
Account From - Sub-account	-	Yes
Account From - Product	-	Yes
Account To - Company	-	Yes
Account To - Department	-	Yes
Account To - Account	-	Yes
Account To - Sub-account	-	Yes
Account To - Product	-	Yes
Order By	Account	Yes

Report Tables and Views

The following tables and views are included in the report:

Tables

- GL_CODE_COMBINATIONS

Views

- FND_FLEX_VALUES_VL
- FND_ID_FLEX_SEGMENTS_VL
- FND_ID_FLEX_STRUCTURES_VL
- GL_LOOKUPS

Journal Batch Summary Report

The Journal Batch Summary report displays information on actual balances for journal batches, source, batch and posting dates, and total entered debits and credits.

Use the report to review posted journal batches for a particular ledger, balancing segment value, currency, and date range. The report sorts the information by journal batch within each journal entry category. In addition, the report displays totals for each journal category and a grand total for each ledger and balancing segment value combination. This report does not report on budget or encumbrance balances.

Report Parameters

The following table describes the report input parameters:

Parameter	Default Value	Mandatory
Chart of Accounts	<Any>	Yes
Ledger/Ledger Set	<Any>	Yes
Currency	<Any>	Yes
Balancing Statement	<Any>	Yes
Start Date	-	Yes
End Date	-	Yes
Adjustment Periods	Yes	Yes

Report Tables and Views

The following tables and views are included in the report:

Tables

- GL_DAILY_CONVERSION_TYPES
- GL_JE_BATCHES
- GL_CODE_COMBINATIONS
- GL_JE_HEADERS
- GL_JE_LINES
- GL_PERIODS
- GL_PERIOD_STATUSES

Views

- FND_CURRENCIES_VL
- FND_FLEX_VALUES_VL
- FND_ID_FLEX_SEGMENTS_VL
- GL_JE_CATEGORIES_VL
- GL_JE_SOURCES_VL

User-Defined Views

- V_REPORT_GL_SETS_OF_BOOKS

Saving the Reports to the Archive Folder

Before you can run a report, you must save the imported reports to the Data Vault archive folder where you retired the application data.

1. Click **Data Visualization** > **Reports and Dashboards**.
2. Click the check box next to **ILM_OracleApps** to select all of the imported reports.
3. Click **Actions** > **Save As**.

   The **Save As** window appears.
4. Select the Data Vault archive folder where you retired the application data.
5. Enter a name for the reports folder.
6. Click **OK**.

   Data Archive creates the reports folder within the archive folder.

Running a Report

You can run a report through the Data Visualization area of Data Archive. Each report has two versions. To run the report, use the version with "Search Form" in the title, for example "Invoice History Details Report - Search Form.cls."

To ensure that authorized users can run the report, review the user and role assignments for each report in Data Archive.

1. Click **Data Visualization** > **Reports and Dashboards**.
2. Select the check box next to the report that you want to run.
3. Click **Actions** > **Run Report**.

   The **Report Search Form** window appears.
4. Enter or select the values for each search parameter that you want to use to generate the report.
5. Click **Apply**.

CHAPTER 15

JD Edwards Enterprise Retirement Reports

This chapter includes the following topics:

Retirement Reports Overview

If you have a Data Visualization license, you can generate reports for certain application modules after you have retired the module data to the Data Vault. Run the reports to review details about the application data that you retired.

You can generate the reports in the Data Visualization area of Data Archive. Each report contains multiple search parameters, so that you can search for specific data within the retired entities. For example, you want to review purchase orders with a particular order type and status code. You can run the Print Purchase Order report and select a specific value for the order type and status code.

When you install the accelerator, the installer imports metadata for the reports as an entity within the appropriate application module. Each report entity contains an interim table and multiple report-related tables. You can view the report tables and the table columns and constraints in the Enterprise Data Manager. You can also use the report entity for Data Discovery.

Prerequisites

Before you can run a Data Visualization report, you must first retire the JDEdwards Enterprise application.

To retire a JD Edwards Enterprise application, perform the following high-level tasks:

1. In the Enterprise Data Manager, create a customer-defined application version.

2. Import metadata from the JD Edwards application to the customer-defined application version.

3. Create retirement entities. Use the generate retirement entity wizard to automatically generate the retirement entities.

4. In the Data Archive user interface, create a source connection. For the connection type, choose the database in which the JD Edwards application is installed.

5. Configure the JD Edwards application login properties.

6. Create a target connection to the Data Vault.

7. Run the create archive folder standalone job to create archive folders in the Data Vault.

8. Create and run a retirement project.

9. In the Enterprise Data Manager, create constraints for the imported source metadata in the ILM repository. Constraints are required for Data Discovery Portal searches. The JD Edwards application retirement accelerator includes constraints for some tables in the JD Edwards system. You might need to define additional constraints for tables that are not included in the accelerator.

10. Copy the entities from the pre-packaged JD Edwards application version to the customer-defined application version. To copy the entities, perform the following high-level steps:

    After you install the JD Edwards retirement report accelerator, the report entities and application modules appear underneath the JD Edwards Retirement application 1.0.

    - **If you imported the application metadata from a single schema** to the customer-defined application version:

    a. Right click on the customer-defined application version where you imported the metadata and select **Copy Entities from Application Version**.

        The **Copy Entities from Application Version** window appears.

    b. Select the application version from where you need to copy, which is typically JD Edwards Retirement 1.0.

    c. Provide the prefix that will be appended before the entity while copying. The Copy All Entities option is not required.

    d. Click **OK**.

    e. After you submit the background job, you can view the job status from the Data Archive user interface. Once the job successfully completes, you can view the copied entities in your customer-defined application version. You can use these entities for Data Discovery.

    - **If you imported the application metadata from multiple schemas** to the customer-defined application version, you must march the schema names and the application module names in the ILM repository to the pre-packaged JD Edwards application version names. To map these names, you must manually update the ILM repository. After you copy the entities, you will revert back to the original names.

The pre-packaged JD Edwards application version has two schemas: TESTDTA and TESTCTL. The TESTCTL schema has one table called "F0005." The remaining tables are from TESTDTA.

a. Run the following update statements:

```
/* identify your custom-defined product family version id*/
SELECT * FROM AM_PRODUCT_FAMILY_VERSIONS;
/* identify schema id's of custom-defined product family version*/
SELECT * FROM AM_META_SCHEMAS WHERE PRODUCT_FAMILY_VERSION_ID= < pfv id>;
UPDATE AM_META_SCHEMAS SET META_SCHEMA_NAME='TESTCTL' WHERE
META_SCHEMA_ID=<Schema id of F0005 table>;
UPDATE AM_META_SCHEMAS SET META_SCHEMA_NAME='TESTDTA'  WHERE
META_SCHEMA_ID=<Schema id of remaining tables>;
```

b. Create the metadata application modules to map to the JD Edwards retirement application modules. Right-click on the application version and select **New Application Module**.

The **Application Module Wizard** window appears.

c. Enter the application module name to match the source name.

d. Repeat the naming process for all of the modules.

e. Right-click the product family version where the tables are imported and retired.

f. Select **Copy Entities from Application Version**.

The **Copy Entities from Application Version** window appears.

g. Select the application version from where you need to copy, typically JD Edwards Retirement 1.0.

h. Provide the prefix that will be appended before the entity while copying. The Copy All Entities option is not required.

i. Click **OK**.

j. After you submit the background job, you can view the job status from the Data Archive user interface. Once the job successfully completes, you can view the copied entities in your customer-defined application version. You can use these entities for Data Discovery and retention.

k. Once the copy process is complete, revert the updates you made in step a.

In JD Edwards applications, dates are stored in Julian format which is 5 or 6 digits. When you search on columns with date values in Data Discovery, select data conversion from Julian data to Gregorian data in the available search options.

To update a retention policy, use the following conversion from Julian to Gregorian as an expression:

```
(CASE WHEN length(char(dec(F43199.OLUPMJ))) = 5 THEN date('1900-01-01') +
int(left(char(int(F43199.OLUPMJ)),2)) years + int(right(char(int(F43199.OLUPMJ)),
3))-1 days ELSE date('1900-01-01') + int(left(char(int(F43199.OLUPMJ)),3)) years +
int(right(char(int(F43199.OLUPMJ)),3))-1 days END)
```

11. Before you run the reports, copy the report templates to the folder that corresponds to the Data Vault connection. You can use these pre-packaged Data Visualization reports to access retired data in the Data Vault. When you copy the reports, the system updates the report schemas based on the connection details.

12. Optionally, validate the retired data.

13. After you retire and validate the application, use the Data Discovery Portal, Data Visualization reports, or third-party query tools to access the retired data.

Accounts Payable Reports

In the Accounts Payable module, you can run the following reports:

- AP Payment History Details
- Open AP Summary
- Manual Payment
- Voucher Journal

AP Payment History Details Report

The AP Payment History Details report shows voucher details for all payments from a selected supplier. Use this report to determine which vouchers have been paid.

Report Parameters

The following table describes the report input parameters:

Parameter	Default Value	Mandatory
Supplier	%	Yes
Start Payment Date	-	No
End Payment Date	-	No

Report Tables

The following tables are included in the report:

- F0005
- F0010
- F0101
- F0111
- F0411
- F0413
- F0414

Open AP Summary Report

The Open AP Summary report displays the total open payable amounts for each supplier. Use the report to review summary information about open balances and aging information.

Report Parameters

The following table describes the report input parameters:

Parameter	Default Value	Mandatory
Aging Date	System Date	Yes
Date Type	G - G/L Date	No
Aging Credits	0	No
Company Name	%	Yes
Order Type	%	No
Supplier Number	%	Yes

Report Tables

The following tables are included in the report:

- F0005
- F0010
- F0101
- F0013
- F0115
- F0411

Manual Payment Report

Use this report to review manual payment journal information in a printed format, instead of viewing the information online.

Report Parameters

The following table describes the report input parameters:

Parameter	Default Value	Mandatory
Document Company	%	Yes
Document Type	%	No
Document Number	-	No

Report Tables

The following tables are included in the report:

- F0005

- F0010
- F0013
- F0111
- F0411
- F0413
- F0414
- F0901
- F0911

Voucher Journal Report

Use this report to review voucher journal information in a printed format, instead of viewing the information online.

Report Parameters

The following table describes the report input parameters:

Parameter	Default Value	Mandatory
Document Company	%	Yes
Document Type	%	No
Document Number	-	No
Batch Number	-	No

Report Tables

The following tables are included in the report:

- F0005
- F0010
- F0013
- F0111
- F0411
- F0901
- F0911

Procurement Management Reports

In the Procurement Management module, you can run the following reports:

- Print Purchase Order

- Purchase Order Detail Print
- Purchase Order Revisions

Print Purchase Order Report

The Print Purchase Order report prints the purchase orders that you have created. Use the reports to review the orders and send them to the appropriate suppliers.

Report Parameters

The following table describes the report input parameters:

Parameter	Default Value	Mandatory
Order Number From	-	No
Order Number To	-	No
Order Type	%	No
Company	%	Yes
Supplier	%	No
Status Code	%	Yes

Report Tables

The following tables are included in the report:

- F0005
- F0101
- F0010
- F0111
- F0116
- F4301
- F4311

Purchase Order Detail Print Report

The Purchase Order Detail Print report displays the system print information about subcontracts, for example associated text and tax information.

Report Parameters

The following table describes the report input parameters:

Parameter	Default Value	Mandatory
Order Number From	-	No
Order Number To	-	No
Order Type	%	No
Company	%	Yes
Supplier	%	Yes

Report Tables

The following tables are included in the report:

- F0005
- F0101
- F0010
- F0006
- F4301
- F4316
- F4311

Purchase Order Revisions History Report

Use the Purchase Order Revisions History report to review information about order revisions, for example the number of revisions to each detail line and the history of all detail line revisions.

Report Parameters

The following table describes the report input parameters:

Parameter	Default Value	Mandatory
Order Number From	-	No
Order Number To	-	No
Order Type	%	No
Order Company	%	Yes
Supplier	%	Yes

Report Tables

The following tables are included in the report:

- F0005
- F0101
- F0010
- F0013
- F40205
- F4301
- F4311
- F4316
- F43199

General Accounting Reports

In the General Accounting module, you can run the following reports:

- General Journal by Account
- Supplier Customer Totals by Account
- General Journal Batch
- GL Chart of Accounts
- GL Trial Balance

General Journal by Account Report

Use the General Journal by Account report to review posted and unposted transactions by account. The report displays totals by the account number.

Report Parameters

The following table describes the report input parameters:

Parameter	Default Value	Mandatory
Document Number	-	No
Document Type	%	Yes
Account Number	%	No

Report Tables

The following tables are included in the report:

- F0005
- F0010

- F0013
- F0901
- F0911

Supplier Customer Totals by Account Report

Report Parameters

The following table describes the report input parameters:

Parameter	Default Value	Mandatory
Account Number	%	No
Address Number	%	No
Company	%	Yes

Report Tables

The following tables are included in the report:

- F0006
- F0010
- F0013
- F0101
- F0111
- F0116
- F0901
- F0911

General Journal Batch Report

The General Journal Batch report prints the debit and credit amounts that amount to balanced entries for invoices and vouchers.

Report Parameters

The following table describes the report input parameters:

Parameter	Default Value	Mandatory
Batch Number	-	No
Batch Type	%	No
Document Number From	-	No

Parameter	Default Value	Mandatory
Document Number To	-	No
Document Type	%	Yes
Ledger Type	%	Yes

Report Tables

The following tables are included in the report:

- F0005
- F0901
- F0911

GL Trial Balance Report

Use the Trial Balance report to review account balances across all business units. You can review similar object accounts, such as all cash accounts. You can also view account totals for each group of accounts.

Report Parameters

The following table describes the report input parameters:

Parameter	Default Value	Mandatory
Fiscal Year (yyyy)	Current Fiscal Year	Yes
Period	1	Yes
Company Name	%	Yes
Ledger Type	AA%	No
Sub Ledger	*	No
Sub Ledger Type	%	No
Currency Code	%	No

Report Tables

The following tables are included in the report:

- F0012
- F0013
- F0902
- F0901
- F0008

- F0005
- F0010

GL Chart of Accounts Report

Use the Chart of Accounts report to review the updated chart of accounts.

Report Parameters

The following table describes the report input parameters:

Parameter	Default Value	Mandatory
Posting Edit Code	%	No
Company	%	Yes

Report Tables

The following tables are included in the report:

- F0909
- F0005
- F0010

Accounts Receivable Reports

In the Accounts Receivable module, you can run the following reports:

- Open Accounts Receivable Summary
- AR Print Invoice
- Invoice Journal

Open Accounts Receivable Summary Report

Use the Open Accounts Receivable Summary report to review summary information about open voucher balances in addition to aging information.

Report Parameters

The following table describes the report input parameters:

Parameter	Default Value	Mandatory
Aging Date	System date	No
Print Parent Number	-	No
Date to Aging Accounts From	D-DUEDATE	No
Age Credits	0	No
Company	%	Yes
Address Number	%	Yes

Report Tables

The following tables are included in the report:

- F03B11
- F0010
- F0013
- F0101

Invoice Journal Report

Use the Invoice Journal report to print invoice journal information. The system selects transactions from the customer ledger (F03B11) and account ledger (F0911) tables.

Report Parameters

The following table describes the report input parameters:

Parameter	Default Value	Mandatory
Batch Type	%	No
Batch Number	-	No
Document Type	%	No
Document Number	-	No
Company	%	Yes

Report Tables

The following tables are included in the report:

- F0005

- F0010
- F0013
- F03B11
- F0901
- F0911

AR Print Invoice Report

Use this report to print the invoices that you created during invoice entry or receipts entry.

Report Parameters

The following tables describes the report input parameters:

Parameter	Default Value	Mandatory
Invoice Print Date	System date	No
Print Currency	-	No
Print Tax Amounts	-	No

Report Tables

The following tables are included in the report:

- F0005
- F0010
- F0013
- F0101
- F0111
- F0116
- F03B11

Sales Order Management Reports

In the Sales Order Management module, you can run the following reports:

- Print Open Sales Order
- Sales Ledger
- Print Held Sales Order

Print Open Sales Order Report

Use the Print Open Sales Order report to review orders that have been picked but not shipped, or picked but not billed. You can also review orders that exceed the customer's requested ship date.

Report Parameters

The following table describes the report input parameters:

Parameter	Default Value	Mandatory
Branch Number	%	No
From Date	1900-01-01	No
Thru Date	System Date	No
Order Type	%	No
Company	%	Yes
Status Code	%	No
Currency Print Option	-	No

Report Tables

The following tables are included in the report:

- F0005
- F0006
- F0013
- F0101
- F4211
- F0010

Sales Ledger Report

Use the Sales Ledger report to analyze and review sales history based on the specific customer number and branch.

Report Parameters

The following table describes the report input parameters:

Parameter	Default Value	Mandatory
From Date	1900-01-01	No
Thru Date	System date	No

Parameter	Default Value	Mandatory
Customer Number	%	No
Branch Number	%	No

Report Tables

The following tables are included in the report:

- F0006
- F0013
- F0101
- F4101
- F42199
- F0010

Print Held Sales Order Report

Use this report to review a list of all sales orders that are on hold for credit, profit margin, partial order hold, and price review.

Report Parameters

The following table describes the report input parameters:

Parameter	Default Value	Mandatory
Hold Code	%	No

Report Tables

The following tables are included in the report:

- F0005
- F0101
- F4209

Inventory Reports

In the Inventory module, you can run the following report:

- Item Master Directory

Item Master Directory Report

Use this report to print a list of item master records.

Report Parameters

The following table describes the report input parameters:

Parameter	Default Value	Mandatory
Item Number	%	No

Report Tables

The following tables are included in the report:

- F4101
- F0005

Address Book Reports

In the Address Book module, you can run the following reports:

- One Line Per Address

One Line Per Address Report

Use the One Line Per Address report to print a list of all addresses that contain one line of detail for each address number.

Report Parameters

The following table describes the report input parameters:

Parameter	Default Value	Mandatory
Address Number	%	No

Report Tables

The following tables are included in the report:

- F0005
- F0101
- F0111
- F0115
- F0116

Saving the Reports to the Archive Folder

Before you can run a report, you must save the imported reports to the Data Vault archive folder where you retired the application data.

1. Click **Data Visualization** > **Reports and Dashboards**.
2. Click the check box next to **ILM_JDEdwards** to select all of the imported reports.
3. Click **Actions** > **Save As**.

    The **Save As** window appears.
4. Select the Data Vault archive folder where you retired the application data.
5. Enter a name for the reports folder.
6. Click **OK**.

    Data Archive creates the reports folder within the archive folder.

Running a Report

You can run a report through the Data Visualization area of Data Archive. Each report has two versions. To run the report, use the version with "Search Form" in the title, for example "Invoice Journal Report - Search Form.cls."

To ensure that authorized users can run the report, review the user and role assignments for each report in Data Archive.

1. Click **Data Visualization** > **Reports and Dashboards**.
2. Select the check box next to the report that you want to run.
3. Click **Actions** > **Run Report**.

    The **Report Search Form** window appears.
4. Enter or select the values for each search parameter that you want to use to generate the report.
5. Click **Apply**.

Oracle PeopleSoft Applications Retirement Reports

This chapter includes the following topics:

Retirement Reports Overview

If you have a Data Visualization license, you can generate reports for certain application modules after you have retired the module data to the Data Vault. Run the reports to review details about the application data that you retired.

You can generate the reports in the Data Visualization area of Data Archive. Each report contains multiple search parameters, so that you can search for specific data within the retired entities. For example, you want to review purchase orders with a particular order type and status code. You can run the Print Purchase Order report and select a specific value for the order type and status code.

When you install the accelerator, the installer imports metadata for the reports as an entity within the appropriate application module. Each report entity contains an interim table and multiple report-related tables. You can view the report tables and the table columns and constraints in the Enterprise Data Manager. You can also use the report entity for Data Discovery.

Prerequisites

Before you can run a Data Visualization report, you must first retire the Oracle PeopleSoft application.

To retire a PeopleSoft application, perform the following high-level tasks:

1. In the Enterprise Data Manager, create a customer-defined application version.
2. Import metadata from the PeopleSoft application to the customer-defined application version, along with user-defined views.
3. Create retirement entities. Use the generate retirement entity wizard to automatically generate the retirement entities.
4. In the Data Archive user interface, create a source connection. For the connection type, choose the database in which the PeopleSoft application is installed.
5. Create a target connection to the Data Vault.
6. Run the create archive folder standalone job to create archive folders in the Data Vault.
7. Create and run a retirement project.
8. In the Enterprise Data Manager, create constraints for the imported source metadata in the ILM repository. Constraints are required for Data Discovery portal searches. The PeopleSoft application retirement accelerator includes constraints for some tables in the PeopleSoft system. You might need to define additional constraints for tables that are not included in the accelerator.
9. Copy the entities from the pre-packaged PeopleSoft application version to the customer-defined application version. To copy the entities, perform the following high-level steps:
   a. Right-click the customer-defined application version where the metadata is imported and select **Copy entities from Application Version**. The **Copy Entities from Application Version** window appears.
   b. Select the application version from where you need to copy, which is typically PeopleSoft Retirement 1.0.
   c. Provide the prefix that will be appended before the entity while copying. The Copy All Entities option is not required.
   d. Click **OK**.
   e. After you submit the background job, you can view the job status from the Data Archive user interface. Once the job successfully completes, you can view the copied entities in your customer-defined application version. You can use these entities for Data Discovery and retention policies.
10. Before you run the reports, copy the report templates to the folder that corresponds to the Data Vault connection. You can use these pre-packaged Data Visualization reports to access retired data in the Data Vault. When you copy the reports, the system updates the report schemas based on the connection details.
11. Optionally, validate the retired data.
12. After you retire and validate the application, use the Data Discovery Portal, Data Visualization reports, or third-party query tools to access the retired data.

Accounts Payable Reports

In the Accounts Payable module, you can run the following reports:

- AP Outstanding Balance by Supplier

- AP Payment History by Payment Method
- AP Posted Voucher Listing

AP Outstanding Balance by Supplier Report

The AP Outstanding Balance by Supplier report lists the gross amount of all the outstanding vouchers for the specified supplier or vendor.

Report Parameters

The following table describes the report input parameters:

Parameter	Default Value	Mandatory
Supplier Set ID	<Any>	Yes
Supplier ID	<Any>	No

Report Tables

The report contains the following tables:

- PS_BUS_UNIT_TBL_AP
- PS_BUS_UNIT_TBL_FS
- PS_INSTALLATION_FS
- PS_PYMNT_VCHR_XREF
- PS_PYMT_TRMS_HDR
- PS_SEC_BU_CLS
- PS_SEC_BU_OPR
- PS_SETID_TBL
- PS_SET_CNTRL_REC
- PS_VCHR_VNDR_INFO
- PS_VENDOR
- PS_VOUCHER

AP Payment History by Payment Method Report

The AP Payment History by Payment Method report provides a history of payments in detail by different payment methods and for different suppliers.

Report Parameters

The following table describes the report input parameters:

Parameter	Default Value	Mandatory
From Date	1900-01-01	Yes
To Date	System Date	Yes
SetID	<Any>	Yes
Currency	USD -US Dollar	No
Run Option	All Banks	No
Payment Method	ACH : Automated Cleaning House	No

Report Tables

The following tables are included in the report:

- PSXLATDEFN
- PSXLATITEM
- PS_BANK_ACCT_CPTY
- PS_BANK_ACCT_DEFN
- PS_BANK_BRANCH_TBL
- PS_BANK_CD_TBL
- PS_CURRENCY_CD_TBL
- PS_PAYMENT_TBL
- PS_SETID_TBL

AP Posted Voucher Listing Report

The AP Posted Voucher Listing report lists all of the posted vouchers for a given business unit and date range. For each voucher ID, the report lists document type, document date, document sequence, ledger, accounting date, application journal, dist type, vchr line, dist line, GL unit, account, dept, product, project, debit amount, credit amount, currency code, and unpost seq#.

Report Parameters

The following table describes the report input parameters:

Parameter	Default Value	Mandatory
From Date (yyyy-MM-dd)	System Date	Yes
Through Date(yyyy-MM-dd)	System Date	Yes

Parameter	Default Value	Mandatory
Business Unit	<Any>	Yes
Vendor/Supplier	Select All Vendors/Suppliers	No

Report Tables

The following tables are included in the report:

- PS_BUS_UNIT_TBL_AP
- PS_BUS_UNIT_TBL_FS
- PS_VCHR_ACCTG_LINE
- PS_VCHR_VNDR_INFO
- PS_VENDOR
- PS_VOUCHER

Purchase Order Reports

In the Purchase Order module, you can run the following reports:

- PO Detail Listing by PO Date
- Non-Owned Purchase History
- PO Order Status by Vendor

PO Detail Listing by PO Date Report

The PO Detail Listing by PO Date report provides detailed purchase order information sorted by purchase order date, list PO ID, Vendor ID, Item ID, Item Description, and more.

Report Parameters

The following table describes the report input parameters:

Parameter	Default Value	Mandatory
From Date	System Date	No
Through Date	System Date	No
Business Unit	<Any>	No
Vendor SetID	<Any>	No

Parameter	Default Value	Mandatory
Vendor ID	<Any>	No
Buyer	<Any>	No

Report Tables

The following tables are included in the report:

- PS_BUS_UNIT_TBL_FS
- PS_BUS_UNIT_TBL_PM
- PS_CNTRCT_CONTROL
- PS_OPR_DEF_TBL_FS
- PS_OPR_DEF_TBL_PM
- PS_PO_HDR
- PS_PO_LINE
- PS_PO_LINE_SHIP
- PS_SETID_TBL
- PS_VENDOR
- PSOPRDEFN

Non-Owned Purchase History Report

The Non-Owned Purchase History report lists utilization information of no-stock items for a business unit or departments. The report information can help to evaluate vendor contracts, establish and maintain budgets, and anticipate supply needs for a company.

Report Parameters

The following table describes the report input parameters:

Parameter	Default Value	Mandatory
Business Unit	<Any>	Yes
Select Date Option	As Of Date	No
As Of Date (yyyy-MM-dd)	-	No
DateRange: From (yyyy-MM-dd) To (yyyy-MM-dd)	-	No
For Department Range: From To	-	No

Parameter	Default Value	Mandatory
All/Specific Item ID	<Any>	No
For Item ID Range: From To	-	No

Report Tables

The following tables are included in the report:

- PS_BUS_UNIT_TBL_FS
- PS_DEPT_TBL
- PS_MASTER_ITEM_TBL
- PS_PO_HDR
- PS_PO_LINE
- PS_PO_LINE_SHIP
- PS_VENDOR
- PS_PO_LINE_DISTRIB

PO Order Status by Vendor Report

This PO Order Status by Vendor report provides purchase order status information sorted by vendor or supplier.

Report Parameters

The following table describes the report input parameters:

Parameter	Default Value	Mandatory
Business Unit	<Any>	Yes
From Date (yyyy-MM-dd)	-	No
To Date (yyyy-MM-dd)	-	No

Report Tables

The report includes the following tables:

- PS_BUS_UNIT_TBL_FS
- PS_BUS_UNIT_TBL_PM
- PSOPRDEFN
- PS_PO_HDR
- PS_PO_LINE

- PS_PO_LINE_SHIP
- PS_VENDOR
- PSXLATITEM

Accounts Receivable Reports

In the Accounts Receivable module, you can run the following reports:

- AR Deposit Summary
- AR Payment Details
- AR Payment Summary

AR Deposit Summary Report

The AR Deposit Summary report lists detailed information for deposits, in either the business unit base currency or the entry currency. It displays the entry date, deposit ID, bank account, post status, control amount, entered amount, posted amount, and more.

Report Parameters

The following table describes the report input parameters:

Parameter	Default Value	Mandatory
From Date	1900-01-01	Yes
To Date	System Date	Yes
Unit	<Any>	Yes
Amount Type	Bass Curr	No
UserID	<Any>	No
Deposit Type	<Any>	No
Bank Code	<Any>	No
Bank Account	<Any>	No
Posting Status	All	No

Report Tables

The following tables are included in the report:

- PSXLATITEM
- PS_BANK_ACCT_CPTY
- PS_BANK_ACCT_DEFN

- PS_BANK_BRANCH_TBL

- PS_BANK_CD_TBL

- PS_BUS_UNIT_TBL_AR

- PS_BUS_UNIT_TBL_GL

- PS_DEPOST_TYPE_TBL

- PS_OPR_DEF_TBL_AR

- PS_OPR_DEF_TBL_FS

- PS_PAYMENT

- PS_SET_CNTRL_REC

- PS_DEPOSIT_CONTROL

- PS_BUS_UNIT_TBL_FS

AR Payment Detail Report

The AR Payment Detail report lists detailed information for all of the payments within a deposit, in either the business unit base currency or the entry currency. The payment details are unavailable if the posting status is "not posted."

Report Parameters

The following table describes the report input parameters:

Parameter	Default Value	Mandatory
From Date	1900-01-01	Yes
Thru Date	System Date	Yes
Deposit Unit	<Any>	Yes
Amount Type	Base Curr	No
User ID	<Any>	No
Deposit ID	<Any>	No
Posting Status	All	No

Report Tables

The following tables are included in the report:

- PS_DEPOSIT_CONTROL

- PS_PAYMENT

- PS_OPR_DEF_TBL_AR

- PS_OPR_DEF_TBL_FS

- PS_PENDING_ITEM

- PS_GROUP_CONTROL

- PS_ENTRY_TYPE_TBL
- PS_CUSTOMER
- PS_SET_CNTRL_REC
- PS_BUS_UNIT_TBL_AR
- PS_BUS_UNIT_TBL_FS

AR Payment Summary Report

The AR Payment Summary report lists the status for all payments within a deposit, in either the business unit base currency or the entry currency. The report lists the entry date, deposit ID, post status, entered amount, payment status, and more.

Report Parameters

The following table describes the report parameters:

Parameter	Default Value	Mandatory
From Date	1900-01-01	Yes
Thru Date	System Date	Yes
Deposit Unit	<Any>	Yes
Amount Type	Base Curr	No
User ID	<Any>	No
Deposit ID	<Any>	No
Posting Status	All	No

Report Tables

The following tables are included in the report:

- PS_BUS_UNIT_TBL_AR
- PSXLATITEM
- PS_BUS_UNIT_TBL_FS
- PS_OPR_DEF_TBL_AR
- PS_OPR_DEF_TBL_FS
- PS_PAYMENT
- PS_DEPOSIT_CONTROL

General Ledger Reports

In the General Ledger module, you can run the following reports:

- GL Trial Balance
- GL Journal Activity
- GL Journal Entry Detail

GL Trial Balance Report

The GL Trial Balance report combines detail and summary balance information. The report displays the ending ledger balances for the specified year and period by chatfield combination, and prints a final total for debits and credits.

Report Parameters

The following table describes the report input parameters:

Parameter	Default Value	Mandatory
Ledger	<Any>	Yes
Business Unit	<Any>	Yes
Fiscal Year	-	Yes
Currency Option	<Any>	No
For Account Range: From To	-	No
For Account Range: From To	-	No

Report Tables

The following tables are included in the report:

- PS_BUS_UNIT_TBL_GL
- PS_BUS_UNIT_TBL_FS
- PS_CAL_DETP_TBL
- PS_CURRENCY_CD_TBL
- PS_GL_ACCOUNT_TBL
- PS_LEDGER
- PS_LED_DEFN_TBL
- PS_PRODUCT_TBL

GL Journal Activity Report

The GL Journal Activity report displays information about journals. It lists the journal date, journal ID, description, and debit and credit amount totals for each journal.

Report Parameters

The following table describes the report input parameters:

Parameter	Default Value	Mandatory
Ledger	<Any>	Yes
Business Unit	<Any>	Yes
Ledger Group	<Any>	Yes
Date or Period	Date	Yes
For Date: From Date (yyyy-MM-dd) To Date (yyyy-MM-dd)	-	No
Date Option	Journal Date	No
For Period: From To Fiscal Year	-	No
Detail/Summary	Detail	No
Journal Status	ALL	No
Report By	Account	No

Report Tables

The following tables are included in the report:

- PS_BUS_UNIT_TBL_GL
- PS_BUS_UNIT_TBL_FS
- PSXLATITEM
- PS_JRNL_HEADER
- PS_JRNL_LN
- PS_LED_DEFN_TBL
- PS_LED_GRP_TBL
- PS_SOURCE_TBL

GL Journal Entry Detail Report

The GL Journal Entry Detail report shows journal entry detail information by business unit, journal ID, ledger and ledger group, and source. The report lists like accounts and products based on the chart fields selected. It also displays debits, credits, and totals for the journal.

Report Parameters

The following table describes the report input parameters:

Parameter	Default Value	Mandatory
From Date (yyyy-MM-dd)	-	No
To Date (yyyy-MM-dd)	-	No
Journal ID	ALL	No
Journal Status	ALL	No
Business Unit	<Any>	Yes
Ledger Group	<Any>	Yes
Ledger	<Any>	Yes
Source	ALL	No
Account	<Any>	No
Product	<Any>	No

Report Tables

The following tables are included in the report:

- PS_BUS_UNIT_TBL_GL
- PS_BUS_UNIT_TBL_FS
- PSXLATITEM
- PS_JRNL_HEADER
- PS_JRNL_LN
- PS_LED_DEFN_TBL
- PS_LED_GRP_TBL
- PS_SOURCE_TBL
- PS_GL_ACCOUNT_TBL
- PS_OPEN_ITEM_GL
- PS_PRODUCT_TBL

Saving the Reports to the Archive Folder

Before you can run a report, you must save the imported reports to the Data Vault archive folder where you retired the application data.

1. Click **Data Visualization** > **Reports and Dashboards**.
2. Click the check box next to the imported reports to select all of the reports.
3. Click **Actions** > **Save As**.

   The **Save As** window appears.
4. Select the Data Vault archive folder where you retired the application data.
5. Enter a name for the reports folder.
6. Click **OK**.

   Data Archive creates the reports folder within the archive folder.

Running a Report

You can run a report through the Data Visualization area of Data Archive. Each report has two versions. To run the report, use the version with "Search Form" in the title, for example "Invoice Journal Report - Search Form.cls."

To ensure that authorized users can run the report, review the user and role assignments for each report in Data Archive.

1. Click **Data Visualization** > **Reports and Dashboards**.
2. Select the check box next to the report that you want to run.
3. Click **Actions** > **Run Report**.

   The **Report Search Form** window appears.
4. Enter or select the values for each search parameter that you want to use to generate the report.
5. Click **Apply**.

CHAPTER 17

Smart Partitioning

This chapter includes the following topics:

Smart Partitioning Overview

Smart partitioning is a process that divides application data into segments based on business rules and dimensions that you configure. Smart partitioning can improve application performance and help you manage application data.

As an application database grows, application performance declines. Smart partitioning uses native database partitioning methods to create segments that increase application performance and help you manage application data growth.

Segments are sets of data that you create with smart partitioning to optimize application performance. You can query and manage segments independently, which increases application response time and simplifies processes such as database compression. You can also restrict access to segments based on application, database, or operating system users.

You create segments based on dimensions that you define according to your organization's business practices and the applications that you want to manage. A dimension is an attribute that defines the criteria to create segments.

Some examples of dimensions include:

- Time. You can create segments that contain data for a certain time period, such as a year or quarter.
- Business unit. You can create segments that contain the data from different business units in your organization.
- Geographic location. You can create segments that contain data for employees based on the country they live in.

Smart Partitioning Process

When you implement smart partitioning, you first create dimensions and segmentation groups with the Enterprise Data Manager. Then, in the Data Archive user interface, you create a data classification and a

segmentation policy. After you run the segmentation policy to create segments, and can create access policies.

When you use smart partitioning, you complete the following tasks:

1. Create dimensions. Before you can create segments, you must create at least one dimension. A dimension adds a business definition to the segments that you create. You create dimensions in the Enterprise Data Manager.

2. Create a segmentation group. A segmentation group defines database and application relationships. You create dimensions in the Enterprise Data Manager.

3. Create a data classification. A data classification defines the data contained in the segments you want to create. Data classifications apply criteria such as dimensions and dimension slices.

4. Create and run a segmentation policy. A segmentation policy defines the data classification that you apply to a segmentation group.

5. Create access policies. An access policy determines the segments that an individual user or program can access.

After you create segments, you can optionally run management operations on them. You can compress segments, make segments read-only, or move segments to another storage classification.

As time passes and additional transactions enter the segmented tables, you can create more segments.

Dimensions

A dimension determines the method by which the segmentation process creates segments, such as by time or business unit. Dimensions add a business definition to data, so that you can manage it easily and access it quickly. You create dimensions during smart partitioning implementation.

The Data Archive metadata includes a dimension called time. You can use the time dimension to classify the data in a segment by date or time. You can also create custom dimensions based on business needs, such as business unit, product line, or region.

Segmentation Groups

A segmentation group is a group of tables that defines database and application relationships. A segmentation group represents a business function, such as order management, accounts payable, or human resources.

The ILM Engine creates segments by dividing each table consistently across the segmentation group. This process organizes transactions in the segmentation group by the dimensions you associated with the group in the data classification. This method of partitioning maintains the referential integrity of each transaction.

If you purchased an application accelerator such as the Oracle E-Business Suite or PeopleSoft accelerator, some segmentation groups are predefined for you. If you need to create your own segmentation groups, you must select the related tables that you want to include in the group. After you select the segmentation group tables you must mark the driving tables, enter constraints, associate dimensions with the segmentation group, and define business rules.

Data Classifications

A data classification is a policy composed of segmentation criteria such as dimensions and dimension slices. You create data classifications to apply consistent segmentation criteria to a segmentation group.

Before you can create segments, you must create a data classification and assign a segmentation group to the data classification. Depending on how you want to create segments, you add one or more dimensions to

a data classification. Then you configure dimension slices to specify how the data for each dimension is organized. When you run a segmentation policy to create segments, the ILM Engine uses native database partitioning methods to create a segment for each dimension slice in the data classification.

You can assign multiple segmentation groups to the same data classification. If the segmentation groups are related, such as order management and purchasing, you might want to apply the same segmentation criteria to both groups. To ensure that the ILM Engine applies the same partitioning method to both segmentation groups, assign the groups to the same data classification.

Segmentation Policies

A segmentation policy defines the data classification that you want to apply to a segmentation group. When you run the segmentation policy, the ILM Engine creates a segment for each dimension slice you configured in the data classification.

After you run a segmentation policy, choose a method of periodic segment creation for the policy. The method of periodic segment creation determines how the ILM Engine creates segments for the segmentation group in the future, as the default segment becomes too large.

Access Policies

An access policy establishes rules that limit user or program access to specified segments. Access policies increase application performance because they limit the number of segments that a user or program can query.

Access policies reduce the amount of data that the database retrieves, which improves both query performance and application response time. When a user or program queries the database, the database performs operations on only the segments specified in the user or program access policy.

Before you create access policies based on program or user, create a default access policy. The default access policy applies to anything that can query the database. Access policies that you create for specific programs or users override the default access policy.

Smart Partitioning Example

Your organization needs to increase the response time of an financial application used by different types of business users in your organization.

To increase application response time, you want to divide the transactions in a high-volume general ledger application module so that application users can query specific segments. You decide to create segments based on calendar year. In the Enterprise Data Manager you create a time dimension that you configure as a date range. Then you create a general ledger segmentation group that contains related general ledger tables. In the Data Archive user interface you create a data classification and configure the dimension slices to create segments for all general ledger transactions by year, from 2010 through the current year, 2013.

When you run the segmentation policy, the ILM Engine divides each table across the general ledger segmentation group based on the dimension slices you defined. The ILM Engine creates four segments, one for each year of general ledger transactions from 2010-2012, plus a default segment. The default segment contains transactions from the current year, 2013, plus transactions that do not meet business rule requirements for the other segments. As users enter new transactions in the general ledger, the application inserts the transactions in the default segment where they remain until you move them to another segment or create a new default segment.

After you create segments, you create and apply access policies to the segmentation group. The default access policy you apply allows all users and programs access to one year of application data. You create another access policy for business analysts who need access to all four years of application data. The access policy you create for the business analysts overrides the default policy.

CHAPTER 18

Smart Partitioning Data Classifications

This chapter includes the following topics:

Smart Partitioning Data Classifications Overview

A data classification is a policy composed of segmentation criteria such as dimensions. You create data classifications to apply consistent segmentation criteria to a segmentation group.

Before you run the segmentation policy, you must create a data classification and assign a segmentation group to the data classification. Depending on how you want to create segments, you add one or more dimensions to a data classification.

After you add dimensions to a data classification, create dimension slices for each dimension in the data classification. Dimension slices specify how the data for each dimension is organized. When you run the segmentation policy, the ILM Engine uses native database partitioning methods to create a segment for each dimension slice in the data classification.

You can assign multiple segmentation groups to the same data classification. If the segmentation groups are related, such as order management and purchasing, you might want to apply the same segmentation criteria to both groups. To ensure that the ILM Engine applies the same partitioning method to both segmentation groups, assign the groups to the same data classification.

Data classifications are specific to a source connection.

Single-dimensional Data Classification

A single-dimensional data classification uses one dimension to divide the application data into segments.

Use a single-dimensional data classification when you want to create segments for application data in one way, such as by year or quarter. Single-dimensional data classifications often use the time dimension.

For example, you manage three years of application data that you want to divide into segments by quarter. You create a time dimension and select range as the dimension type and date as the datatype. When you run the segmentation policy the ILM Engine creates 13 segments, one for each quarter and one default segment. The data in the default segment includes all of the data that does not meet the requirements for a segment, such as transactions that are still active. The default segment also includes new transactional data.

Multidimensional Data Classification

A multidimensional data classification uses more than one dimension to divide the application data into segments.

You can include more than one dimension in a data classification, so that Data Archive uses all the dimensions to create the segments. When you apply multidimensional data classifications, the ILM Engine creates segments for each combination of values.

For example, you want to create segments for a segmentation group by both a date range and sales office region. The application has three years of data that you want to create segments for. You choose the time dimension to create segments by year. You also create a custom dimension called region. You configure the region dimension to create segments based on whether each sales office is in the Eastern or Western sales territory.

When you run the segmentation policy, the ILM Engine creates seven segments, one for each year of data from the Western region, one for each year of data from the Eastern region, and a default segment. Each non-default segment contains the combination of one year of data from either the East or West sales offices. All remaining data is placed in the default segment. The remaining data includes data that does not meet the requirements for a segment, such as transactions that are still active, plus all new transactions.

Dimension Slices

Create dimension slices to specify how the data for each dimension is organized. When you run the segmentation policy, the ILM Engine creates a segment for each dimension slice in the data classification.

Dimension slices define the value of the dimensions you include in a data classification. Each dimension slice you create corresponds to a segment. You might create a dimension slice for each year or quarter of a time dimension. When you create the dimension slice you enter the value of the slice.

Every dimension slice has a corresponding sequence number. When you create a dimension slice, the ILM Engine populates the sequence number field. Sequence numbers indicate the order in which you created the segments. The slice with the highest sequence number should correspond to the most recent data. If the slice with the highest sequence number does not correspond to the segment with the most chronologically recent data, change the sequence number of the slice.

Dimension Slice Example

You need to save space on a production database. Most application users in your organization need access to only the current fiscal year of transactions in a general ledger application module. You decide to create segments for every fiscal year of transactions in the GL module and then compress the segments that contain data from previous years.

First you create a single-dimensional data classification with the time dimension. Then you add three dimension slices to the time dimension, one for each year of transactions. When you run the segmentation policy, the ILM Engine creates a segment for each dimension slice in the classification.

Time Dimension Formulas

Use formulas to populate the next logical dimension slice for a time dimension. The value of the dimension slice begins where the value of the previous dimension slice ended.

The initial segmentation process creates a segment for each dimension slice in the data classification, plus a default segment for new transactions and transactions that do not meet the business rule criteria of other segments. Over time, the default segment becomes larger and database performance degrades. To avoid performance degradation you must periodically create a new default segment, split the current default segment, or create a new non-default segment to move data to. Before you redefine the default segment or create a new one, you must create a dimension slice that specifies the values of the new or redefined segment.

If you plan to use the time dimension to create segments by year or quarter, you can use a formula to simplify the process of redefining or creating a new default segment. Before you run a segmentation policy for the first time, select a formula in the data classification that you plan to apply to the segmentation group. The formula creates and determines the value of the next logical dimension slice that Data Archive uses to redefine the default segment or create a new one.

Formulas use a specific naming convention. If you use a formula in a data classification, you must name the dimension slices you create according to the formula naming convention.

The following table describes the formula naming conventions:

Naming Convention	Formula Description
YYYY	Creates a dimension slice with a value of one year, such as January 01, 2013-December 31, 2013. If you use the YYYY formula, you must name the dimension slices for the dimension in the YYYY format, such as "2013."
QQYY	Creates a dimension slice with a value of one quarter, such as January 01, 2013-March 31, 2013. If you use the QQYY formula, you must name the dimension slices for the dimension in the QQYY format, such as "0113" for the first quarter of 2013.
YYYYQQ	Creates a dimension slice with a value of one quarter, such as January 01, 2013-March 31, 2013. If you use the YYYYQQ formula, you must name the dimension slices for the dimension in the YYYYQQ format, such as "201301" for the first quarter of 2013.
YYYY_QQ	Creates a dimension slice with a value of one quarter, such as January 01, 2013-March 31, 2013. If you use the YYYY_QQ formula, you must name the dimension slices for the dimension in the YYYY_QQ format, such as "2013_01" for the first quarter of 2013.

Creating a Data Classification

Before initial segmentation, create a data classification that defines the dimensions you want to use when you run the segmentation policy.

1. Click **Workbench** > **Manage Segmentation**.
2. Click the **Data Classification** tab.
3. Click the **Actions** tab and select **New**.

   The **New Data Classification** page appears.
4. Enter a name for the data classification.

   Data classification names do not support special characters.
5. From the **Dimension** drop-down menu, choose a dimension to add to the data classification.
6. Select the **Use Formula** checkbox if you want to use a formula to populate new dimension slices.
7. To use a formula, select the format from the **Formula** drop-down menu.
8. To add another dimension to the data classification, click the **Add** button.
9. When you are finished adding dimensions, click the **Save** button.

   The data classification appears in the list of classifications on the main data classification page.

Creating a Dimension Slice

Create dimension slices for each dimension in a data classification. The ILM Engine creates a segment for each dimension slice in a data classification.

1. Click **Workbench** > **Manage Segmentation**.
2. Click the **Data Classification** tab.
3. Under the desired data classification, select the dimension that you want to create slices for.
4. Click the **Add** button on the lower pane of the data classification window.
5. Enter a name for the dimension slice. If you chose to use a formula in the data classification, you must name the dimension slice according to the formula naming convention. For example, enter "2013" for the YYYY formula.
6. Enter a value for the dimension slice. If you chose to use a formula in the data classification to populate dimension slices, you must provide a value consistent with the formula naming convention. For example, if you chose the YYYY formula, select January 1, 2013 for the FROM value and December 31, 2013 for the TO value.
7. Click **Save**.

CHAPTER 19

Smart Partitioning Segmentation Policies

This chapter includes the following topics:

Smart Partitioning Segmentation Policies Overview

The first time you want to create segments from a segmentation group, create a segmentation policy and schedule it to run. After you implement the segmentation policy, you can run management operations on the segments.

A segmentation policy defines the data classification that you want to apply to a segmentation group. When you run the segmentation policy, the ILM Engine creates a segment for each dimension slice you configured in the data classification.

After you run a segmentation policy, choose a method of periodic segment creation for the policy. The method of periodic segment creation determines how the ILM Engine creates segments for the segmentation group in the future, as the default segment becomes too large.

After you create segments, you can run management operations on an individual segment or a segment set. You can make segments read-only, move segments to another storage classification, or compress segments.

RELATED TOPICS:

- "Creating a Segmentation Policy" on page 230
- " Check Indexes for Segmentation Job" on page 26

Segmentation Policy Process

A segmentation policy defines the data classification that you apply to a segmentation group. When you create a segmentation policy, you configure steps and properties to determine how the policy runs. Create and run a segmentation policy to create segments.

Before you create a segmentation policy, you must create a data classification. A data classification is a policy composed of segmentation criteria such as dimensions and dimension slices. Data classifications apply consistent segmentation criteria to a segmentation group. You can reuse data classifications in multiple segmentation policies.

When you create a segmentation policy, you complete the following steps:

1. Select the data classification that you want to associate with a segmentation group. A segmentation group is a group of tables based on business function, such as order management or accounts receivable. You create segmentation groups in the Enterprise Data Manager. If you want to use interim table processing to pre-process the business and dimension rules associated with a segmentation group, you must configure interim table processing in the Enterprise Data Manager.

2. Select the segmentation group that you want to create segments for. A segmentation group is a group of tables based on business function, such as order management or accounts receivable. You create segmentation groups in the Enterprise Data Manager. You can run the same segmentation policy on multiple segmentation groups.

3. After you select the segmentation group, review and configure the tablespace and data file properties for each segment. The tablespace and date file properties determine where and how Data Archive stores the segments. When you run the segmentation policy, the ILM Engine creates a tablespace for each segment.

4. If necessary, you can save a draft of a segmentation policy and exit the wizard at any point in the creation process. If you save a draft of a segmentation policy after you add a segmentation group to the policy, the policy will appear on the **Segmentation Policy** tab with a status of **Generated**. As long as the segmentation policy is in the **Generated** status, you can edit the policy or add a table to the segmentation group in the Enterprise Data Manager.

5. Configure the segmentation policy steps, then schedule the policy to run. The ILM Engine runs a series of steps and sub-steps to create segments for the first time. You can skip any step or pause the job run after any step. You can also configure Data Archive to notify you after each step in the process is complete. When you configure the policy steps you have the option to configure advanced segmentation parameters. After you run the policy, it appears as **Implemented** on the **Segmentation Policy** tab. Once a policy is implemented, you cannot edit it. You can run multiple segmentation policies in parallel.

6. If necessary, you can run the unpartition tables standalone job to reverse the segmentation process and return the tables to their original state. For more information about standalone jobs, see the chapter *Scheduling Jobs*.

Related Topics:

-
-

Segmentation Policy Properties

When you create a segmentation policy, you configure tablespace and data file properties. Before you schedule the policy to run, you configure how the ILM Engine runs each step in the policy. You can also

configure advanced segmentation parameters, such as options for index creation and the option to drop the original tables after you create segments.

Tablespace and Data File Properties

The ILM Engine creates a tablespace that contains at least one data file for each segment in the segmentation policy. Before you run the segmentation policy, configure the tablespace and data file properties for each segment. For more information about tablespace and data file properties, see the database software documentation.

The following table describes tablespace and data file properties:

Property	Description
Segment Group	Name of the segmentation group you want to create segments for. This field is read-only.
Table Space Name	Name of the tablespace where the segmentation group resides. This field is read-only.
Storage Class	Name of the storage classification where the segmentation group resides.
Logging	Generates redo records. Select yes or no. Default is yes.
Force Logging	Forces the generation of redo records for all operations. Select yes or no. Default is yes.
Extent Management	Type of extent management. Default is local.
Space Management	Type of space management. Default is auto.
Auto Allocate Data Files	Automatically allocates the data files.
Segment Name	Name of the segment you want to create a tablespace for. This field is read-only.
Compression Type	Type of compression applied to the global tablespace. Choose either **No Compression** or **For OLTP** . For information about Oracle OLTP and other types of compression, refer to database documentation.
Allocation Type	Type of allocation. Select auto allocate or uniform. Default is auto allocate.
Description	Description of the data file. This field is optional.
File Size	Maximum size of the data file. Default is 100 MB.
Auto Extensible	Increases the size of the data file if it requires more space in the database.
Max File Size	Maximum size to which the data file can extend if you select auto extensible.
Increment By	Size to increase the data file by if you select auto extensible.

Segmentation Policy Steps

Before you run a segmentation policy, configure the policy steps. You can skip any step except the Segment Data step. You can also pause the policy run after each step.

If you choose to skip a step in the policy, Data Archive might also select related steps in the policy to skip. For example, if you skip the Create Audit Snapshot - Before step, Data Archive also selects **Skip** for the Create Audit Snapshot - After and Compare Audit Snapshots - Summary steps. If Data Archive selects **Skip** for a job step, you can clear the check box to run the step if required.

The following table describes the segmentation policy steps:

Step	Description
Create audit snapshot - before	Creates a snapshot of the state of the objects, such as indexes and triggers, that belong to the segmentation group before the ILM Engine moves any data.
Create optimization indexes	Creates optimization indexes to improve smart partitioning performance. Indexes are used during the segment data step to generate the selection statements that create segments. This step is applicable only if you have defined optimization indexes in the Enterprise Data Manager.
Preprocess business rules	Uses interim tables to preprocess the business rules you applied to the segmentation group. This step is visible only if you selected to use interim table processing in the Enterprise Data Manager.
Get table row count per segment	Estimates the size of the segments that will be created when you run a segmentation policy. Default is skip.
Allocate datafiles to mount points	Estimates the size of tables in the segmentation group based on row count and determines how many datafiles to allocate to a particular tablespace. Default is skip.
Generate explain plan	Creates a plan for how the smart partitioning process will run the SQL statements that create segments. This plan can help determine how efficiently the statements will run. Default is skip.
Create tablespaces	Creates the tablespaces for the new segments.
Segment data	Required. Moves the data into the new tablespaces.
Drop optimization indexes	Drops any optimization indexes you created in the Enterprise Data Manager.
Create audit snapshot - after	Creates a snapshot of the state of the objects, such as indexes and triggers, that belong to the segmentation group after the ILM Engine moves the data into the segments.
Compare audit snapshots - summary	Compares and summarizes the before and after audit snapshots.
Compile invalid objects	Compiles any not valid objects, such as stored procedures and views, that remain after the ILM Engine creates segments.
Enable access policy	Enables the access policies you applied to the segmentation group.

Step	Description
Collect segmentation group statistics	Collects database statistics, such as row count and size, for the segmentation group.
Clean up after segmentation	Cleans up temporary objects, such as tables and indexes, that are no longer valid after segment creation.

Advanced Segmentation Parameters

When you configure the policy steps you can also configure advanced segmentation parameters. Advanced parameters contain options such as index creation and temporary tables.

The following table describes advanced segmentation parameters:

Field	Description
Add Part Column to Non Unique Index	Adds the partition column to a non-unique index.
Index Creation Option	Select global, local, or global for unique indexes. The smart partitioning process recreates indexes according to the segmentation policy you configure. If the source table index was previously partitioned, the segmentation process might drop the original partitions. Default is local.
Bitmap Index Option	Select local bitmap or global regular. Default is local bitmap.
Parallel Degree	The degree of parallelism the ILM Engine uses to create and alter tables and indexes. Default is 4.
Parallel Processes	Number of parallel processes the ILM Engine uses in the compile invalid objects step. Default is 4.
DB Policy Type	Select dynamic, context sensitive, or shared context sensitive.
Estimate Percentage	The sampling percentage. Valid range is 0-100. Default is 10.
Apps Stats Option	Select no histogram or all global indexed columns auto histogram.
Drop Managed Renamed Tables	Drops the original tables after you create segments.
Count History Segments	Flag that indicates if you want to merge archived history data with production data when you run the segmentation policy. If set to Yes, the ILM Engine includes the history data during the Create Audit Snapshots steps and the Compare Audit Snapshot step of the segmentation process.
Drop Existing Interim Tables	If you have enabled interim table processing to pre-process business rules, the partitioning process can drop any existing interim tables for the segmentation group and recreate the interim tables. Select Yes to drop existing interim tables and recreate them. If you select No and interim tables exist for the segmentation group, the partitioning process will fail.

Creating a Segmentation Policy

Create and run a segmentation policy to create segments for a segmentation group. When you create the segmentation policy, you associate a segmentation group with a data classification and optionally configure advanced properties. Then, schedule the project to run. You can view both task and database session details of each job step while the job runs.

1. Click **Workbench** > **Manage Segmentation.**

2. Click the **Segmentation Policies** tab.

3. On the left side of the **Manage Segmentation** screen, select the source connection where the application data you want to create segments for resides.

   The segmentation groups you created on the source connection appear in the upper pane.

4. Click **Actions** > **New Segmentation Policy**.

   The **New Segmentation Policy** tab opens.

5. Select the data classification that you want to associate with the segmentation group.

   Select a dimension underneath the data classification to view the dimension slices for that dimension.

6. Click **Next**.

   The **Select Segmentation Groups** screen appears.

7. On the right side of the screen, click **Actions** > **Add Segmentation Group**.

   The **Segmentation Group LOV** window appears.

8. Select the segmentation group that you want to create segments for and click **Add**.

   The ILM Engine generates metadata for the segmentation group you selected. The Data Archive UI displays the dimension slices associated with the data classification that you chose and the properties for each dimension slice.

9. To review or edit the tablespace and data file allocation parameters that apply to each segment, click the **Tablespace** link for each dimension slice.

   The **Table Space and Data File Association** window appears.

10. Review or edit the tablespace and data file allocation parameters. Then click **Save**.

11. Click **Next**.

    The **Define Parameters and Steps** page appears.

12. Configure the segmentation policy job steps. You can choose to skip any step, pause the job after any step, or receive a notification after any step.

    The **Run Before** and **Run After** functionality is not supported.

13. To configure advanced options such as index creation options, click the **Define Parameters** option.

14. To save the policy as a draft to run later, click **Save Draft**.

15. To run or schedule the policy, click **Publish and Schedule**.

    The **Schedule Job** page appears.

16. To run the job immediately, select **Immediately**. To schedule the job to run in the future, select **On** and choose a date and time.

17. Optionally, you can enter an email address to receive a notification when the policy run completes, terminates, or returns an error.

18. Click **Schedule**.

19. Optionally, you can view the task details and SQL statements for each job step that is running, completed, or in an error state. Click **Jobs** > **Monitor Jobs**.

    The **Monitor Jobs** page appears.

20. To expand the job details, click the arrow next to the Job ID of the segmentation policy.

21. To view task details and database details for a job step, click the name of a job step.

22. Click **View Job Step Details**.

    The **Monitor** window appears.

23. Click a task name.

24. In the lower pane of the window, click **Task Details** or **Database Session Details**. Then select the command or SQL ID that you want to view the details of.

Related Topics:

- "Smart Partitioning Segmentation Policies Overview" on page 225
- "Segmentation Policy Process" on page 226
- " Check Indexes for Segmentation Job" on page 26

Business Rule Analysis Reports

If you chose to enable business rule pre-processing for the segmentation group in the Enterprise Data Manager, you can generate a business rule analysis report before you run the segmentation policy. The business rule analysis report gives a graphical representation of how many records will be active versus inactive after business rules are applied.

When you create a business rule analysis report, you first specify the period of time that you want the report to cover. You can choose from a yearly, quarterly, or monthly report.

You can generate three types of reports. The transaction summary report is a graphical representation of all segmentation group records that will be processed using all business rules applied to the driving table. The report details how many records will be active and how many records will be inactive for the segmentation group. Active records will remain the default segment while inactive records will move to the appropriate segment in accordance with business rules.

The transaction summary by period report details how many records will be active versus inactive by the report period you specified. For example, if the segmentation group contains data from 2005 to 2013, the report displays a bar graph for each year that details how many records will be active and how many records will be inactive for the given year after all business rules are applied.

If you want to view how many records will be active versus inactive for a particular business rule instead of all business rules, create a business rule details by period report. The business rule details by period report displays a graphical representation of active versus inactive records by the time period and specific business rule that you select.

If you copy a segmentation group, enable business rules pre-processing, and try to generate a business rule analysis report for the copied segmentation group, you will receive an error. You can view business rule analysis reports for one segmentation group at a time, for example either the original group or the copied group.

Creating a Business Rule Analysis Report

You can create a business rule analysis report to preview how many records will be active versus inactive after business rules for the segmentation group are applied.

1. Click **Workbench** > **Manage Segmentation.**
2. Click the **Segmentation Policies** tab.
3. On the left side of the **Manage Segmentation** screen, select the source connection where the segmentation group you want to create a business rule analysis report for resides.

   The segmentation groups you created on the source connection appear in the upper pane.
4. Select the segmentation group you want to create a report for.
5. Click **Actions** > **View Business Rule Analysis Report**.

   The **Business Rule Analysis Report** window opens.
6. In the **Period** drop-down menu, select the period of time that you want the report to cover.
7. In the **Driving Table** drop-down menu, select the driving table that you want to use to generate the report.
8. In the **Report Type** drop-down menu, select the type of report you want to generate. Choose from transaction summary report, transaction summary by period report, or business rule details by period report.
9. If you chose to generate a business rule details by period report, select the business rules to be included in the report from the **Business Rules** drop-down menu. You can select all business rules or specific business rules.
10. Click **Generate Report** to generate the report in the Data Archive user interface, or **Download Report** to download the report in Microsoft Excel format.

Segment Creation for Implemented Segmentation Policies

You can create segments for a segmentation policy after you run it. Before you create a segment for an implemented segmentation policy, choose a method of periodic segment creation to apply to the segmentation group.

The first time you run a segmentation policy, the ILM Engine creates a segment for each dimension slice in the data classification, plus a default segment for new transactions and transactions that do not meet the business rule criteria of other segments.

Over a period of time, the default segment becomes larger and database performance degrades. To avoid performance degradation you must periodically create a new default segment, split the current default segment, or create a new non-default segment to move data to. You might also create a new non-default segment if you need to segment data that was previously missed, or to add a new organization to an organization dimension in the segmentation policy.

Before you can create a segment for an implemented segmentation group, you must create a dimension slice to specify the value of the segment. If you applied a data classification that includes a formula to the segmentation group, the ILM Engine creates and populates the dimension slice when you create the segment. If you did not apply a data classification that uses a formula, you must manually create the dimension slice.

Before you create the new segment, decide which method of periodic segment creation to apply to the segmentation group. The method of periodic segment creation determines how the ILM Engine will create segments in the future as the default segment becomes too large or if you need to manually create a new segment. When you create the new segment you can also choose to compress the segment or make it read-only.

You can change the method of periodic segment creation. Before you can change the method of periodic segment creation, you must delete any segments in the group that are generated and not implemented. When you delete a generated segment, you must delete the most-recently generated segment first.

Create a New Default Segment

You can create a new default segment for a segmentation policy after the end of a specified time period. Create a new default segment when the data classification for the segmentation group contains only one dimension.

If the transactions in the default segment are primarily from the last time period, such as a year or quarter, and the default does not have a large number of active transactions, you can create a new default segment. When you create a new default segment, the ILM Engine creates the new default segment and renames the old default segment based on the time period of the transactions that it contains, for example 2013_Q2.

The ILM Engine moves active transactions and transactions that do not meet the business criteria of the renamed 2013_Q2 segment from the renamed segment to the new default segment. This method requires repeated row-level data movement of active transactions. Do not use this method if the default segment has many open transactions.

Note: You can generate one new default segment at a time. Before you can generate another default segment, you must either delete or implement any existing generated default segments.

Before you run the job to create the new segment, you can configure the job steps. The following table lists each of the job steps to create a new default segment:

Step	Description
New segment internal task	Required. Internal task that prepares the job to run.
Preprocess business rules	Uses interim tables to preprocess the business rules you applied to the segmentation group. This step is visible only if you selected to use interim table processing in the Enterprise Data Manager.
Create new segments	Required. Creates the table spaces, grants required schema access, and creates the segments.
Incremental move policy segments to default	Estimates the size of tables in the segmentation group based on row count and determines how many datafiles to allocate to a particular tablespace. Default is skip.
Collect segmentation group statistics	Collects database statistics, such as row count and size, for the segmentation group.
Drop interim tables	If you have enabled interim table processing to pre-process business rules, this step drops the interim tables.
Clean up after segmentation	Cleans up temporary objects, such as tables and indexes, that are no longer valid after segment creation.

Split the Default Segment

If the default segment contains more than one time period of data, you can split the default segment.

If you have a large amount of data in the default segment that spans more than one time period, use the split default segment method. The split default segment method creates multiple new segments from the default segment and moves the transactions from the default segment into the appropriate segment for each time period.

For example, the default segment contains data from three quarters, including the current quarter. The split default method creates three new segments, one for each of the two quarters that have ended, plus a new default segment. The ILM Engine moves closed transactions to the new segment for the quarter in which they belong. Then the ILM Engine moves new and active transactions to the new default segment.

The split default method of segment creation works with both single and multidimensional data classifications. This method is ideal when the application environment has complex business rules with long transaction life cycles.

To split the default segment, the application database must be offline.

Note: If the policy contains multiple segments in a Generated state and you want to delete one of the segments, you must delete the most recently generated segment first.

Before you run the job to split the default segment, you can configure the job steps. The following table lists each of the job steps to split the default segment:

Step	Description
New segment internal task	Required. Internal task that prepares the job to run.
Create audit snapshot - before	Creates a snapshot of the state of the objects, such as indexes and triggers, that belong to the segmentation group before the ILM Engine moves any data.
Create optimization indexes	Creates optimization indexes to improve smart partitioning performance. Indexes are used during the segment data step to generate the selection statements that create segments. This step is applicable only if you have defined optimization indexes in the Enterprise Data Manager.
Disable access policy	Disables the access policy on the default segment.
Create new segments	Required. Creates the table spaces, grants required schema access, and creates the segments.
Split single segment to multi segments	Required. Moves data to the new segment.
Drop optimization indexes	Drops any optimization indexes you created in the Enterprise Data Manager.
Create audit snapshot - after	Creates a snapshot of the state of the objects, such as indexes and triggers, that belong to the segmentation group after the ILM Engine moves the data into the segments.
Compare audit snapshots - summary	Compares and summarizes the before and after audit snapshots.
Compile invalid objects	Compiles any not valid objects, such as stored procedures and views, that remain after the ILM Engine creates segments.
Enable access policy	Enables the access policies you applied to the segmentation group.

Step	Description
Collect segmentation group statistics	Collects database statistics, such as row count and size, for the segmentation group.
Clean up after segmentation	Cleans up temporary objects, such as tables and indexes, that are no longer valid after segment creation.

Create a Non-Default Segment

You can create a new non-default segment. When you create a new non-default segment, the ILM Engine moves data incrementally from the default segment to the new non-default segment.

Create a new non-default segment when the data classification associated with the segmentation group is multi-dimensional and you want to keep the default segment small by moving data incrementally.

You can create a new non-default segment at any time. When you move data into a new non-default segment, the process does not affect the default segment and does not require that the database is offline.

You might also create a new non-default segment to move data out of the default segment that was previously missed.

Note: You can generate one new non-default segment at a time. Before you can generate another non-default segment, you must either delete or implement any existing generated non-default segments.

Before you run the job to create the new segment, you can configure the job steps. The following table lists each of the job steps to create a new non-default segment:

Step	Description
New segment internal task	Required. Internal task that prepares the job to run.
Create new segments	Required. Creates the table spaces, grants required schema access, and creates the segments.
Incremental move policy segments to default	Moves data from the default segment to the new non-default segment.
Collect segmentation group statistics	Collects database statistics, such as row count and size, for the segmentation group.
Clean up after segmentation	Cleans up temporary objects, such as tables and indexes, that are no longer valid after segment creation.

Creating a Segment for an Implemented Segmentation Group

You can periodically create a segment for an implemented segmentation policy to manage the growth of the default segment.

Before you create a segment for an implemented segmentation policy, create a dimension slice that specifies the value of the segment. If the data classification you applied to the segmentation group uses a formula, the ILM Engine will create and populate the dimension slice.

1. Click **Workbench** > **Manage Segmentation**.

   The **Manage Segmentation** window appears.

2. Select the segmentation policy that you want to create a segment for.

   The status of the segmentation policy must be **Implemented**.

3. If you want to change the method of periodic segment creation, click **Actions** > **Edit Periodic Segmentation Creation**.

4. Select a method of periodic segment creation and click **Ok**.

5. In the bottom pane of the **Manage Segmentation** window, click **Actions** > **Create**.

   The ILM Engine generates metadata for the new segment and the **Create New Segment Job Parameters** window appears.

6. Click the **Tablespace** link to edit or review the segment tablespace or data file allocation parameters.

7. If you want to compress the segment or make it read-only, select the box under the **Compressed** or **Read-Only** columns.

8. Click **Next**.

   The **Parameters and Steps** page appears.

9. Configure the job steps. You can pause the job after any step or receive a notification after any step.

10. Click **Publish**.

    The **Schedule Job** page appears.

11. To make create the segment immediately, select **Immediately**. To schedule the job for the future, select **On** and choose a date and time.

    The **Monitor Jobs** window appears. The create new segment job is paused at the top of the jobs list.

12. To run the job, select the create new segment job and click **Resume Job**.

Deleting a Generated Segment

If necessary, you can delete a segment that has been generated and not yet implemented.

1. Click **Workbench** > **Manage Segmentation**.

   The **Manage Segmentation** window appears.

2. In the top pane, select the segmentation policy that contains the segment that you want to delete.

3. In the bottom pane, select the segment that you want to delete.

   The status of the segment must be **Generated**.

4. Click **Actions** > **Delete**.

Editing the Method of Periodic Segment Creation

You can edit the method of periodic segment creation for a segmentation group. Before you can edit the method, you must delete any segments that are generated but not yet implemented.

1. Click **Workbench** > **Manage Segmentation**.

2. In the top pane, select the segmentation group that you want to edit.

3. Click **Actions** > **Edit Periodic Segment Creation**.

   The **Edit Periodic Segment Creation** window appears.

4. Select the method that you want to apply to the segmentation group.

5. Click **Ok**.

Segment Management

You can perform management operations on segments after you run a segmentation policy to create them.

After you run a segmentation policy, you can compress a segment, make it read-only, or move it to another storage classification. You can also create segment sets to perform an action on multiple segments at one time.

Segment Merging

You can merge multiple segments into one segment. To merge segments, create a segment set and then run the merge partitions into single partition standalone job.

You might want to merge segments to combine multiple history segments into one larger segment. For example, you can merge four different quarter segments into one segment that spans the entire year.

Before you can merge segments, create a segment set that contains the segments that you want to merge. For example, you might want to merge the segments "2015_01," "2015_02," "2015_03," and "2015_04" into one history segment that spans data for the year 2015. First, create a segment set that contains the four quarter segments. The name of the segment set must be unique and descriptive of the segment set.

After you create the segment set, run the merge partitions into single partition standalone job. When you run the merge partitions job, you select parameters such as the source connection name and the name of the segment set that you created. You can provide a tablespace for the merged segment, but if you do not, the job creates the segment in the existing tablespace of the most recent segment. If you provide a different tablespace, you must create the tablespace before you run the job.

After the job completes successfully, the older segments are merged into the most recent segment. In the **Manage Segmentation** window, the older segments have the status "Merged." The most recent segment is renamed to be a combination of the older segments. For instance, the most recent segment might be named "2015_01_2015_02_2015_03_2015_04." This segment has the status "Implemented."

If you change your mind about merging a segment set, you can replace the merged segment with the original segments. To revert the merge process, run the replace merged partitions with original partitions standalone job. For more on the replace merged partitions job, see the *Scheduling Jobs* chapter.

You can continue to create other segment sets and then run the merge job on those segment sets, provided that the segments exist in the same segmentation group. If you want to run the merge job multiple times for a segmentation group, mark the "FinalMergeJob" parameter as "N" until you run the final merge job for the segmentation group. On the final merge job for the segmentation group, set the "FinalMergeJob" parameter to Y.

You can repeat the merge process for each segmentation group in a data classification. You must complete all of the merge jobs for each segmentation group in a data classification before you run the "clean up after merged partitions" job. The clean up job drops the empty partitions and tablespaces, and merges the segment metadata. After you run the clean up job, the older individual segments are removed from the **Manage Segmentation** window and only the merged segment remains.

As a best practice, conduct a full backup before you merge segments.

Merging Segments

To merge segments, create a segment set and then run the merge partitions into single partition standalone job. When you finish merging the partitions, run the clean up after merge partitions standalone job to make the changes final.

1. Create a segment set that contains the segments that you want to merge together. For more information, see the topic "Segment Sets" in this chapter.

2. Schedule the merge partitions to single partition standalone job. For more information about the standalone job, see the *Scheduling Jobs* chapter.

   The merge partitions into single partition job merges the data from the older partitions into the most recent partition.

3. When the standalone job completes successfully, click **Workbench** > **Manage Segmentation**.

4. In the lower pane of the **Manage Segmentation** window, verify that the segment set has merged correctly. The older segments in the set will have the status "Merged." The job renames the most recent segment to a combination of the merged segment names. This segment has the status "Implemented."

5. If for any reason you need to reverse the merge, run the replace merged partitions with original partitions standalone job before you continue.

6. Repeat steps 1-4 for each segment set that you want to merge.

7. When you finish merging segments, run the clean up after merge partitions standalone job. For more information, see the *Scheduling Jobs* chapter.

   The run clean up after merge partitions job drops the empty partitions and merges the partition metadata. After you run the job, only the merged partition appears in the **Manage Segmentation** window.

Segment Compression

You can compress the tablespaces and indexes for a specific segment. Compress a segment if you want to reduce the amount of disk space that the segment requires. When you compress a segment, you can optionally configure advanced compression properties.

The ILM Engine uses native database technology to compress a segment. If you want to compress multiple segments at once, create a segment set and run the compress segments standalone job.

If the segmentation group that contains the segments you want to compress has a bitmap index and any of the segments in the group are read-only, you must either turn the read-only segments to read-write mode or drop the bitmap index before you compress any segments. This is because Oracle requires that you mark bitmap indexes as unusable before you compress segments for the first time. When compression is complete, the segmentation process cannot rebuild the index for segments that are read-only.

Segment Compression Properties

Before you compress a segment, review the segment compression properties. You can optionally configure advanced compression properties.

The following table describes segment compression properties:

Property	Description
Segmentation Group Name	The name of the segmentation group that the segment belongs to. This field is read-only.
Segment Name	The name of the segment that you want to compress. You enter the segment name when you create the dimension slice. This field is read-only.
Tablespace Name	The name of the tablespace where the segment resides. This field is read-only.
Storage Class Name	The name of the storage classification that the segment resides in. This field is read-only.
Remove Old Tablespaces	Removes old tablespaces. This option is selected by default.
Logging Options	Choose logging or no logging. Default is no logging.
Parallel Degree	Default is 4.

Compressing a Segment

When you configure a segment for compression, you can schedule it to run or you can run it immediately.

1. Click **Workbench** > **Manage Segmentation**.

   The **Manage Segmentation** window appears.

2. Select the segmentation policy that includes the segment you want to compress.

   Each segment in the policy appears in the lower pane of the window.

3. Select the segment that you want to compress.

4. Click **Actions** > **Compress.**

   The **Compress** window appears.

5. Optionally, click **Show Advanced Parameters** to remove old tablespaces after compression, select the logging option, or change the parallel degree.

6. Click **Schedule**.

   The **Schedule Job** window appears.

7. To compress the segment immediately, select **Immediately**. To schedule the compression job for the future, select **On** and choose a date and time.

8. Optionally, enter an email address to receive a notification when the compression job completes, terminates, or returns an error.

9. Click **Schedule**.

Read-only Segments

You can make the tablespaces for a specific segment read-only. Make segment tablespaces read-only to prevent transactional data in a segment from being modified.

Read-only Segment Properties

Review the properties to make a segment read-only. The properties are read-only.

The following table describes read-only segment properties:

Properties	Description
Segmentation Group Name	The name of the segmentation group that the segment belongs to.
Segment Name	The name of the segment that you want to make read-only. You enter the segment name when you create the dimension slice.
Tablespace Name	The name of the tablespace where the segment resides.
Storage Class Name	The name of the storage classification that the segment resides in.

Making a Segment Read-only

When you configure a segment to be read-only, you can schedule it to run or you can run it immediately.

1. Click **Workbench** > **Manage Segmentation**.

   The **Manage Segmentation** window appears.

2. Select the segmentation policy that includes the segment you want to compress.

   Each segment in the policy appears in the lower pane of the window.

3. Select the segment that you want to make read-only.

4. Click **Actions** > **Read Only.**

   The **Read Only** window appears.

5. Click **Schedule**.

   The **Schedule Job** page appears.

6. To make the segment read-only immediately, select **Immediately**. To schedule the job for the future, select **On** and choose a date and time.

7. Optionally, enter an email address to receive a notification when the job completes, terminates, or returns an error.

8. Click **Schedule**.

Storage Classification Movement

When you run a segmentation policy, the ILM Engine assigns segments to the default storage classification. You can move a segment to a different storage classification.

You might want to move a segment to another storage classification if the segment will not be accessed frequently and can reside on a slower disk. A Data Archive administrator can create storage classifications.

Segment Storage Classifications Properties

Before you move a segment to a different storage classification, review the storage classification properties and choose a storage classification to move the segment to.

The following table describes storage classification properties:

Property	Description
Segmentation Group Name	The name of the segmentation group that the segment belongs to. This field is read-only.
Segment Name	The name of the segment that you want to move. You enter the segment name when you create the dimension slice. This field is read-only.
Tablespace Name	The name of the tablespace where the segment resides. This field is read-only.
Compress Status	Indicates whether the segment is compressed. This field is read-only.
Current Storage Class Name	Name of the storage classification where the segment currently resides. This field is read-only.
Read Only	Indicates whether the segment is read-only. This field is read-only.
Storage Class Name	Storage classification that you want to move the segment to. Choose from the drop-down menu of available storage classifications.

Moving Segments to a Different Storage Classification

To move a segment to a different storage classification, select the storage classification that you want to move the segment to. You can schedule the job to run or you can run it immediately.

1. Click **Workbench** > **Manage Segmentation**.

   The **Manage Segmentation** window appears.

2. Select the segmentation policy that includes the segment you want to compress.

   Each segment in the policy appears in the lower pane of the window.

3. Select the segment that you want to move to a different storage classification.

4. Click **Actions** > **Move Storage Classification.**

   The **Move Storage Classification** window appears.

5. Select the storage classification that you want to move the segment to.

6. Click **Schedule**.

   The **Schedule Job** page appears.

7. To move the segment immediately, select **Immediately**. To schedule the job for the future, select **On** and choose a date and time.

8. Optionally, enter an email address to receive a notification when the compression job completes, terminates, or returns an error.

9. Click **Schedule**.

Segment Sets

A segment set is a set of segments that you group together to perform a single action on. If you want to run a standalone job on multiple segments at once, create a segment set.

If you want to compress multiple segments at once, make multiple segments read-only, or move multiple segments to a different storage classification, you must create a segment set. After you create a segment set, you can run a standalone job on the entire segment set. When you run a job on a segment set, the ILM Engine runs the job on only the segments in the set.

You can add a segment to multiple segment sets. If you want to run another job on a segment set, you can edit a segment set to include new segments or remove segments.

Creating a Segment Set

To run a standalone job on multiple segments at once, create a segment set.

1. Click **Workbench** > **Manage Segmentation**.

2. Click the **Segmentation Policy** tab.

3. Select the segmentation policy that includes the segments you want to create a set of.

4. In the bottom pane, select the **Segment Sets** tab.

5. Click **Actions** > **Create/Edit Segment Sets**.

   The **Create/Edit Segment Sets** window appears.

6. Click **Actions** > **Add Set**.

   All segments in the segmentation policy appear in **Associated Segments** pane.

7. In the **Name** field, enter a name for the segment set and click the checkmark.

8. In the bottom pane, click **No** next to the name of each segment that you want to include in the set. After you click **No**, select the checkbox that appears.

9. Click **Save**.

   The segment set appears under the **Segment Sets** tab on the **Manage Segmentation** screen.

CHAPTER 20

Smart Partitioning Access Policies

This chapter includes the following topics:

Smart Partitioning Access Policies Overview

An access policy establishes rules that limit user or program access to specified segments. Access policies increase application performance because they limit the number of segments that a user or program can query.

Access policies reduce the amount of data that the database retrieves, which improves both query performance and application response time. When a user or program queries the database, the database performs operations on only the segments specified in the user or program access policy.

When you create an access policy, you apply it to one dimension of a segmentation group. For example, you want to limit an application user's access on a general ledger segmentation group to one year of data. You create the user access policy and apply it to the time dimension of the segmentation group. Create an access policy for each dimension associated with the segmentation group.

Before you create access policies based on program or user, create a default access policy. The default access policy applies to anything that can query the database. Access policies that you create for specific programs or users override the default access policy.

Access policies are dynamic and can be changed depending on business needs. Users can optionally manage their own access policies.

Access Policy Example

Most application users and programs in your organization rarely need access to more than one year of application data. You configure the default access policy to limit access on all segmentation groups to one year of data. A program that accesses the general ledger segmentation group to run month-end processes, however, needs to access only quarterly data. You create an access policy specifically for this program. You also create an application user access policy for a business analyst who needs access to three years of data on all segmentation groups.

Default Access Policy

The default access policy is an access policy that applies to anything that can query the database. Access policies you create for programs or users override the default access policy.

Before you create a default access policy, decide how much data a typical program or user in your organization needs to access. The default policy should be broad enough to allow most users in your organization to access the data they need without frequently adjusting their access policies.

You can configure the default access policy to apply to all segmentation groups. If you want to change the default access policy for a specific segmentation group, you can override the default policy with a policy that applies to a specific segmentation group.

For example, you can configure the default policy to allow two years of data access for all segmentation groups. You can then create a default policy for the accounts receivable segmentation group that allows access to one year of data. When you create the second policy, Data Archive gives it a higher sequence number so that it overrides the first. Each time you create an access policy, Data Archive assigns a higher sequence number. A Data Archive administrator can change the sequence number of any policy except the default policy. The default policy sequence must be number one.

When a user or program queries the database, Data Archive starts with the default access policy and successively checks each access policy with a higher sequence number. Data Archive applies the policy with the highest sequence number unless you configure an access policy to override other policies or exempt a user or program from all access policies.

Oracle E-Business Suite User Access Policy

You can create user access policies for Oracle E-Business Suite users. Oracle E-Business Suite user access policies are based on user profiles within Oracle E-Business Suite.

Before you create an Oracle E-Business Suite access policy in Data Archive, create a user profile with the Application Developer in Oracle E-Business Suite. The internal name for the Oracle E-Business Suite user profile must match the name you want to give the access policy. When you create the access policy in Data Archive, set the user profile value to match the value that you want Oracle E-Business Suite to return to Data Archive. For example, set the user profile value to two if you want to allow access to two years of data.

Example

Within Data Archive you set a default data access policy of two years. You then configure the inventory manager user profile in Oracle E-Business Suite to allow access to three years of data. When you configure the access policy in Data Archive, Data Archive connects to Oracle E-Business Suite to return the values in the Oracle E-Business Suite user profile.

PeopleSoft User Access Policy

You can create user access policies for PeopleSoft users. PeopleSoft user access policies limit user access to segmentation groups you created for PeopleSoft applications.

Before you create a PeopleSoft access policy in Data Archive, create a user profile within PeopleSoft. The internal name for the user profile must match the name you give the access policy in Data Archive.

Data Archive identifies the user who accesses the database based on the session value set by PeopleSoft. If an access policy exists within Data Archive for the Peoplesoft user profile name, the ILM Engine only queries the segments that the user has access to.

Database User Access Policy

You can create an access policy for a database user. The database user name must match the name you assign the policy to.

You might create an access policy for a database user when the database user is used to connect to an interface, such as an application interface. For example, a specific reporting group might use a database user to connect to the application database.

OS User Access Policy

You can create access policies for operating system users. Create an operating system user access policy if you need to restrict access to segments but cannot specify a program, application, or database user.

For example, if your environment has business objects that connect to a general application schema, but you cannot identify the name of the program or database user, you can limit access with an operating system user access policy. The database stores the operating system user in the session variables table. Data Archive compares the operating system user in the session variable table to the operating system user assigned to the access policy within Data Archive. If the values match, Data Archive applies the policy you configured for the operating system user.

Program Access Policy

You can create an access policy for a program. Program access policies help optimize program performance.

Create an access policy for a program based on the function of the program, such as running reports. For example, you might create a program access policy for a transaction processor that needs to access current data. You could limit access for the transaction processor to one quarter of data. You might also create an access policy for a receivables aging reporting tool that needs to view historical data. You configure the policy to allow the receivables aging reporting tool access to three years of historical data.

Access Policy Properties

Enter access policy properties to create and configure an access policy.

The following table describes access policy properties:

Field	Description
Policy Type	The type of access policy you want to create. Select from the drop-down menu of available policy types.
Assigned To	The user or program that you want to assign the policy to.
Dimension	The dimension that you want to apply the access policy to. Select from the drop-down menu of available dimensions.
Segmentation Group	The segmentation group that you want to apply the access policy to. Available segmentation groups are listed in the drop-down menu. Select the **All** button to apply the policy to every segmentation group.
Scope Type	The scope type of the policy. Select date offset, list, or exempt. - Date offset. Choose date offset when you apply an access policy to a time dimension to restrict access by date. - List. Choose list when you apply an access policy to a dimension other than time. - Exempt. Choose exempt if you want to bypass all access policies, including the default policy, to see all data for a segmentation group.
Offset Value	The amount of data you want the policy holder to be able to access if you chose a scope type of date offset. For example, if you want the user to be able to access two years of data, enter two in the **Value** field and select **Year** from the **Unit** drop-down menu. You can select a unit of year, quarter, or month.
Override	Option to override all access policies. If you want the access policy to override access policies with higher sequence numbers, select Override.
Enable	Enables the access policy.

Creating an Access Policy

Create an access policy to limit access to specific segments. When you create an access policy, you select the segmentation group or groups and the dimension that the policy applies to.

1. Click on **Workbench** > **Manage Segmentation**.
2. Click on the **Access Policy** tab.
3. Click the **Actions** button and select **New.**

   The **New Access Policy** page appears.
4. Select the policy type from the drop-down menu.
5. Enter the user or program that you want to assign the policy to.
6. Select the dimension to associate with the access policy.
7. Select the segmentation group or groups to apply the policy to.

8. Select the scope type.

9. If you selected date offset for the scope type, enter the offset value.

10. Select the **Override** checkbox if you want the access policy to override all access policies.

11. Select the **Enable** checkbox to enable the policy.

CHAPTER 21

Language Settings

This chapter includes the following topics:

Language Settings Overview

You can change language settings in the Information Lifecycle Management web user interface at the browser level.

Changing Browser Language Settings

The browser language settings allow you to change the browser's default language to your preferred language.

To change the language settings at the browser level:

1. Open the Internet Explorer web browser.
2. Click Tools > Internet Options.
3. Click the General tab.
4. Click Languages.
5. In the Language Preference window, click Add.
6. Select the language.

    For example: French(France) [fr-FR] to change the language to French.
7. Click OK.

APPENDIX A

Data Vault Datatype Conversion

This appendix includes the following topics:

Data Vault Datatype Conversion Overview

When you archive to a file, the Data Archive project includes the Data Vault Loader job. One of the tasks that the Data Vault Loader job performs is the datatype conversion between the source database and Data Vault. The Data Vault Loader job can convert datatypes from Oracle, Microsoft SQL Server, IBM DB2 UDB, and Teradata databases.

If you have custom datatypes on your source database, you must manually map each custom datatype to a Data Vault datatype before you run the Data Vault Loader job. For information on how to map custom datatypes, see the *Data Archive Administrator Guide*.

Data Vault

The Data Vault Loader job converts native datatypes based on the database.

Oracle Datatypes

The following table describes how the Data Vault Loader job converts Oracle native datatypes to Data Vault datatypes:

Oracle Datatype	Data Vault Datatype
CHAR	- If length is less than or equal to 32768, converts to CHAR(length). - If length is greater than 32768, converts to CLOB.
VARCHAR	- If length is less than or equal to 32768, converts to VARCHAR(length). - If length is greater than 32768, converts to CLOB.

Oracle Datatype	Data Vault Datatype
DATE	TIMESTAMP
TIMESTAMP	TIMESTAMP
TIMESTAMP WITH TIME ZONE	VARCHAR(100)
TIMESTAMP(9) WITH TIME ZONE	VARCHAR
INTERVAL YEAR TO MONTH	VARCHAR
INTERVAL YEAR(4) TO MONTH	VARCHAR
INTERVAL DAY TO SECOND	VARCHAR
INTERVAL DAY(9) TO SECOND(9)	VARCHAR
BFILE	BLOB
BLOB	BLOB
CLOB	CLOB
RAW (2000)	VARCHAR (4000)
LONG RAW	BLOB
LONG	BLOB
NUMBER	DECIMAL
FLOAT	- If length is less than or equal to 63, converts to REAL. - If length is greater than 63 but less than or equal to 126, converts to FLOAT(53). - If length is greater than 126, converts to FLOAT.
DECIMAL	DECIMAL
DEC	DECIMAL
INT	INTEGER
INTEGER	INTEGER
SMALLINT	SMALLINT
REAL	- If length is less than or equal to 63, converts to REAL. - If length is greater than 63 but less than or equal to 126, converts to FLOAT(53). - If length is greater than 126, converts to FLOAT.
DOUBLE PRECISION	DOUBLE PRECISION

Oracle Datatype	Data Vault Datatype
ROWID	VARCHAR(18)
UROWID	VARCHAR

Manual Conversion of Oracle Datatypes

The following table describes the Oracle datatypes that require manual conversion:

Source Datatype	Data Vault Datatype
NUMBER	If the length is more than 120, change the datatype to DOUBLE in the metadata XML file.

IBM DB2 Datatypes

The following table describes how the Data Vault Loader job converts IBM DB2 native datatypes to Data Vault datatypes:

IBM DB2 Datatype	Data Vault Datatype
CHARACTER	- If length is less than or equal to 32768, converts to CHAR(length). - If length is greater than 32768, converts to CLOB.
VARCHAR	- If length is less than or equal to 32768, converts to VARCHAR(length). - If length is greater than 32768, converts to CLOB.
LONG VARCHAR	CLOB
DATE	TIMESTAMP
TIME	VARCHAR
TIMESTAMP	TIMESTAMP
BLOB	BLOB
SMALLINT	SMALLINT
INTEGER	INTEGER
BIGINT	DECIMAL(19)
REAL	- If length is less than or equal to 63, converts to REAL. - If length is greater than 63 but less than or equal to 126, converts to FLOAT(53). - If length is greater than 126, converts to FLOAT.
FLOAT	- If length is less than or equal to 63, converts to REAL. - If length is greater than 63 but less than or equal to 126, converts to FLOAT(53). - If length is greater than 126, converts to FLOAT.

IBM DB2 Datatype	Data Vault Datatype
DOUBLE PRECISION	DOUBLE PRECISION
DECIMAL	DECIMAL

Manual Conversion of IBM DB2 Datatypes

The following table describes the IBM DB2 datatypes that require manual conversion:

Source Datatype	Data Vault Datatype
GRAPHIC	Change to VARCHAR in the metadata XML file.
VARGRAPHIC	Change to VARCHAR in the metadata XML file.
LONG VARGRAPHIC	Change to CLOB in the metadata XML file.
DBCLOB	Change to CLOB in the metadata XML file.

Microsoft SQL Server Datatypes

The following table describes how the Data Vault Loader job converts Microsoft SQL Server native datatypes to Data Vault datatypes:

SQL Server Datatype	Data Vault Service Datatype
CHAR	- If length is less than or equal to 32768, converts to CHAR(length). - If length is greater than 32768, converts to CLOB.
CHAR(8000)	CHAR(8000)
TEXT	CLOB
VARCHAR	VARCHAR
VARCHAR(8000)	VARCHAR(8000)
VARCHAR(MAX)	CLOB
DATE	TIMESTAMP
DATETIME2(7)	VARCHAR
DATETIME	TIMESTAMP
DATETIMEOFFSET	VARCHAR(100)
SMALLDATETIME	TIMESTAMP
TIME	VARCHAR(100)

SQL Server Datatype	Data Vault Service Datatype
BIGINT	DECIMAL(19)
BIT	VARCHAR(10)
DECIMAL	DECIMAL
INT	INTEGER
MONEY	DECIMAL(19,4)
NUMERIC	DECIMAL(19)
SMALLINT	SMALLINT
SMALLMONEY	INTEGER
TINYINT	INTEGER
FLOAT	- If length is less than or equal to 63, converts to REAL. - If length is greater than 63 but less than or equal to 126, converts to FLOAT(53). - If length is greater than 126, converts to FLOAT.
REAL	- If length is less than or equal to 63, converts to REAL. - If length is greater than 63 but less than or equal to 126, converts to FLOAT(53). - If length is greater than 126, converts to FLOAT.
BINARY	BLOB
BINARY(8000)	BLOB
IMAGE	BLOB
VARBINARY	BLOB
VARBINARY(8000)	BLOB
VARBINARY(MAX)	BLOB

Teradata Datatypes

The following table describes how the Data Vault Loader job converts Teradata native datatypes to Data Vault datatypes:

Teradata Datatype	Data Vault Datatype
BYTE	BLOB
BYTE(64000)	BLOB
VARBYTE(64000)	BLOB
BLOB	BLOB

Teradata Datatype	Data Vault Datatype
CHARACTER	- If length is less than or equal to 32768, converts to CHAR(length). - If length is greater than 32768, converts to CLOB.
CHARACTER(32000)	CLOB
LONG VARCHAR	CLOB
VARCHAR(32000)	CLOB
LONGVARGRAPHIC	CLOB
VARGRAPHIC	- If length is less than or equal to 8192, converts to VARCHAR(4x). - If length is greater than 8192, converts to CLOB.
CLOB	CLOB
DATE	TIMESTAMP
TIME	VARCHAR
TIME WITH TIMEZONE	VARCHAR
TIMESTAMP	VARCHAR
TIMESTAMP WITH TIMEZONE	VARCHAR
INTERVAL YEAR	VARCHAR
INTERVAL YEAR(4)	VARCHAR
INTERVAL YEAR(4) TO MONTH	VARCHAR
INTERVAL YEAR TO MONTH	VARCHAR(100)
INTERVAL MONTH	VARCHAR
INTERVAL MONTH(4)	VARCHAR
INTERVAL DAY	VARCHAR
INTERVAL DAY(4)	VARCHAR
INTERVAL DAY TO HOUR	VARCHAR(100)
INTERVAL DAY(4) TO HOUR	VARCHAR
INTERVAL DAY TO MINUTE	VARCHAR
INTERVAL DAY(4) TO MINUTE	VARCHAR
INTERVAL DAY TO SECOND	VARCHAR
INTERVAL DAY(4) TO SECOND	VARCHAR

Teradata Datatype	Data Vault Datatype
INTERVAL HOUR	VARCHAR
INTERVAL HOUR(4)	VARCHAR
INTERVAL HOUR TO MINUTE	VARCHAR
INTERVAL HOUR(4) TO MINUTE	VARCHAR
INTERVAL HOUR TO SECOND	VARCHAR
INTERVAL HOUR(4) TO SECOND(6)	VARCHAR
INTERVAL MINUTE	VARCHAR
INTERVAL MINUTE(4)	VARCHAR
INTERVAL MINUTE TO SECOND	VARCHAR
INTERVAL MINUTE(4) TO SECOND(6)	VARCHAR
INTERVAL SECOND	VARCHAR
INTERVAL SECOND(4,6)	VARCHAR
SMALLINT	SMALLINT
INTEGER	INTEGER
BIGINT	DECIMAL(19)
BYTEINT	INTEGER
REAL	- If length is less than or equal to 63, converts to REAL. - If length is greater than 63 but less than or equal to 126, converts to FLOAT(53). - If length is greater than 126, converts to FLOAT.
FLOAT	- If length is less than or equal to 63, converts to REAL. - If length is greater than 63 but less than or equal to 126, converts to FLOAT(53). - If length is greater than 126, converts to FLOAT.
DOUBLE PRECISION	- If length is less than or equal to 63, converts to REAL. - If length is greater than 63 but less than or equal to 126, converts to FLOAT(53). - If length is greater than 126, converts to FLOAT.
DECIMAL	DECIMAL
PERIOD(DATE)	UDCLOB
PERIOD(TIMESTAMP WITH TIMEZONE)	UDCLOB

Teradata Datatype	Data Vault Datatype
PERIOD(TIME WITH TIMEZONE)	UDCLOB
PERIOD(TIMESTAMP)	UDCLOB

Datatype Conversion Errors

The Data Vault Loader job may not always convert datatypes automatically. The Data Vault Loader job may not be able to find a generic datatype to convert when it creates objects in Data Vault. Or, the Data Vault Loader job may have conversion errors or warnings when it loads numeric data to Data Vault. In either case, you must convert the datatypes manually.

Convert the datatypes if you receive one of the following errors in the log file:

Unsupported Datatype Error

If the Data Vault Loader job processes an unsupported datatype and is not able to identify a generic datatype, you receive an error message in the log file. The log file indicates the unsupported datatype, the table and column the datatype is included in, and the metadata xml file location that requires the change.

Conversion Error

If the Data Vault Loader job loads columns that have large numeric data or long strings of multibyte characters, it generates conversion warnings in the log file. The log file indicates the reason why the job failed if there is a problem converting data.

Converting Datatypes for Unsupported Datatype Errors

Manually convert datatypes to resolve the unsupported datatype error in the Data Vault Loader log file.

1. Access the log file.

2. Search the log for the unsupported datatype error.

   Use the error description to find the table and corresponding datatype that you need to convert.

3. Access the metadata xml file that the ILM engine generated.

   Use the error description to find the file location.

4. Change the datatype to a supported Data Vault Service datatype.

   Use the error description to find the table and corresponding datatype that you need to convert.

5. Resume the Data Vault Loader job.

Converting Datatypes for Conversion Errors

Manually modify columns to resolve the conversion errors generated by the Data Vault Loader job. Data Archive ships with sample scripts that you could modify to make column changes to the Data Vault. You can use the SQL Worksheet to run the scripts on the Data Vault.

1. Use a text editor to open the sample script installed with Data Archive.

2. Modify the script for the column changes you need to make.

3. Run the script.

4. Resume the Data Vault Loader job.

If the conversion error involves a change in the size of a column, use the following sample script to change the size of columns in the Data Vault:

```
CREATE DOMAIN "dbo"."D_TESTCHANGESIZE_NAME_TMP" VARCHAR(105);
ALTER TABLE "dbo"."TESTCHANGESIZE" ADD COLUMN
"NAME_TMP" "dbo"."D_TESTCHANGESIZE_NAME_TMP";
ALTER TABLE "dbo"."TESTCHANGESIZE" DROP COLUMN NAME;
ALTER TABLE "dbo"."TESTCHANGESIZE" RENAME COLUMN "NAME_TEMP" "NAME";
DROP DOMAIN "dbo"."D_TESTCHANGESIZE_NAME";
RENAME DOMAIN "dbo"."D_TESTCHANGESIZE_NAME_TMP" "D_TESTCHANGESIZE_NAME";
COMMIT;
```

If the conversion error involves a change in the number of columns in the table, use the following sample script to add columns to the table in the Data Vault:

```
CREATE DOMAIN "dbo"."D_TESTCHANGESIZE_DOB" TIMESTAMP;
ALTER TABLE "dbo"."TESTCHANGESIZE" ADD COLUMN "DOB" "dbo"."D_TESTCHANGESIZE_DOB";
COMMIT;
```

Special Characters in Data Vault

This appendix includes the following topic:

Special Characters in Data Vault

When you retire data to the Data Vault, the presence of special characters in the table, column, or schema names of the data might cause the Data Vault Loader job or other Data Vault functions, such as retention, legal hold, or Data Visualization, to fail.

Before you create a retirement project, check for the presence of special characters in table, column, and schema names.

Supported Special Characters

Certain special characters are supported for schema, table, and column names in Data Vault.

Table and Column Names

The following special characters are supported for table and oclumn names in Data Vault: _ $ ~ ! @ ^ & - + = : < > ? | { }

Schema/Database Names

The following special characters are supported for schema and database names in Data Vault: _ $

Note: Schema names should not start with _ or $.

SAP Application Retirement Supported HR Clusters

This appendix includes the following topics:

SAP Application Retirement Supported HR Clusters Overview

SAP applications encode data in an SAP format. You can only read the data through the SAP application layer, not the database layer. The retirement job reads and transforms the SAP encoded data for certain clusters.

The retirement job logs in to the SAP system and uses the ABAP import command to read the SAP encoded cluster data. The job transforms the encrypted data into XML format during the extraction process. The XML structure depends on the cluster ID in the transparent table.

The job reads data from the PCL1, PCL2, PCL3, PCL4, and PCL5 transparent HR tables. The tables store sensitive human resource data. The job transforms the data by cluster IDs. The transformation is not supported for all cluster IDs. The job also transforms data for the all cluster IDs for the STXL text table.

PCL1 Cluster IDs

The retirement job converts encrypted data into XML format depending on the HR cluster ID.

The following table lists the PCL1 cluster IDs that the retirement job transforms into XML:

SAP Cluster ID	SAP Cluster Description
B1	PDC data
G1	Group incentive wages
L1	Ind. incentive wages
PC	Personnel calendar
TA	General data for travel expense acctg
TC	Credit card data for travel expense acctg
TE	HR travel expenses international
TX	Texts for PREL

PCL2 Cluster IDs

The retirement job converts encrypted data into XML format depending on the HR cluster ID.

The following table lists the PCL2 cluster IDs that the retirement job transforms into XML:

SAP Cluster ID	SAP Cluster Description
A2	Payroll results Argentina: Annual consolidation
AA	Payroll results (Saudi Arabia)
AB	Payroll results (Albania)
AD	Payroll results (Andorra)
AE	Payroll results (garnishment)
AG	Payroll results (Antigua)
AH	Payroll results (Afghanistan)
AM	Payroll results (Armenia)
AO	Payroll results (Angola)
AR	Payroll results (Argentina)

SAP Cluster ID	SAP Cluster Description
AW	Payroll results (Aruba)
AZ	Payroll results (Azerbaijan)
B2	PDC data (month)
BA	Payroll results (Bosnia and Herzegovina)
BB	Payroll results (Barbados)
BD	Payroll results (Bangladesh)
BF	Payroll results (Burkina Faso)
BG	Payroll results (Bulgaria)
BH	Payroll results (Bahrain)
BI	Payroll results (Burundi)
BJ	Payroll results (Benin)
BM	Payroll results (Bermuda)
BN	Payroll results (Brunei)
BO	Payroll results (Bolivia)
BR	Payroll results (Brazil)
BS	Payroll results (Bahamas)
BT	Payroll results (Bhutan)
BW	Payroll results (Botswana)
BY	Payroll results (Belarus)
BZ	Payroll results (Belize)
CA	Cluster directory (CU enhancement for archiving)
CF	Payroll results (Central African Republic)
CG	Payroll results (Democratic Republic of the Congo)
CI	Payroll results (Ivory Coast)
CK	Payroll results (Cook Islands)
CL	Payroll results (Chile)
CM	Payroll results (Cameroon)

SAP Cluster ID	SAP Cluster Description
CN	Payroll results (China)
CO	Payroll results (Colombia)
CV	Payroll results (Cape Verde)
DJ	Payroll results (Djibouti)
DM	Payroll results (Dominica)
DO	Payroll results (Dominican Republic)
DP	Garnishments (DE)
DQ	Garnishment directory
DY	Payroll results (Algeria)
EC	Payroll results (Ecuador)
EE	Payroll results (Estonia)
EG	Payroll results (Egypt)
EI	Payroll results (Ethiopia)
EM	Payroll results (United Arab Emirates)
ER	Payroll results (Eritrea)
ES	Payroll results (totals)
ET	Payroll results (totals)
FI	Payroll results (Finland)
FJ	Payroll results (Fiji)
FM	Payroll results (Federated States of Micronesia)
FO	Payroll results (Faroe Islands)
GA	Payroll results (Gabon)
GD	Payroll results (Grenada)
GE	Payroll results (Georgia)
GF	Payroll results (French Guiana)
GH	Payroll results (Ghana)
GL	Payroll results (Greenland)

SAP Cluster ID	SAP Cluster Description
GM	Payroll results (Gambia)
GN	Payroll results (Guinea)
GP	Payroll reconciliations (GB)
GQ	Payroll results (Equatorial Guinea)
GR	Payroll results (Greece)
GT	Payroll results (Guatemala)
GU	Payroll results (Guam)
GW	Payroll results (Guinea-Bissau)
GY	Payroll results (Guyana)
HK	Payroll results (HongKong)
HN	Payroll results (Honduras)
HR	Payroll results (Croatia)
HT	Payroll results (Haiti)
IC	Payroll results (Iceland)
ID	Interface toolbox - Directory for interface results
IE	Payroll results (Ireland)
IF	Payroll interface
IL	Payroll results (Israel)
IN	Payroll results India
IQ	Payroll results (Iraq)
IR	Payroll results (Iran)
IS	Payroll results (Indonesia)
JM	Payroll results (Jamaica)
JO	Payroll results (Jordan)
KE	Payroll results (Kenya)
KG	Payroll results (Kyrgyzstan)
KH	Payroll results (Cambodia)

SAP Cluster ID	SAP Cluster Description
KI	Payroll results (Kiribati)
KM	Payroll results (Comoros)
KO	Payroll results (Democratic Republic of the Congo)
KP	Payroll results (North Korea)
KR	Payroll results (Korea)
KT	Payroll results (Saint Kitts and Nevis)
KU	Payroll results (Cuba)
KW	Payroll results (Kuwait)
KZ	Payroll results (Kazakhstan)
LC	Payroll results (Saint Lucia)
LE	Payroll results (Lebanon)
LI	Payroll results (Liechtenstein)
LK	Payroll results (Sri Lanka)
LO	Payroll results (Laos)
LR	Payroll results (Liberia)
LS	Payroll results (Lesotho)
LT	Payroll results (Lithuania)
LU	Payroll results (Luxembourg)
LV	Payroll results (Latvia)
LY	Payroll results (Libya)
MA	Payroll results (Morocco)
MC	Payroll results (Monaco)
MD	Payroll results (Moldova)
MG	Payroll results (Madagascar)
MH	Payroll results (Marshall Islands)
MK	Payroll results (Macedonia, FYR of)
ML	Payroll results (Mali)

SAP Cluster ID	SAP Cluster Description
MM	Payroll results (Myanmar)
MN	Payroll results (Mongolia)
MP	Payroll results (Northern Mariana Islands)
MR	Payroll results (Mauritania)
MT	Payroll results (Malta)
MU	Payroll results (Mauritius)
MV	Payroll results (Maldives)
MW	Payroll results (Malawi)
MX	Payroll results (Mexico)
MZ	Payroll results (Mozambique)
NA	Payroll results (Namibia)
NC	Payroll results (New Caledonia)
NE	Payroll results (Niger)
NG	Payroll results (Nigeria)
NP	Payroll results (Nepal)
NR	Payroll results (Nauru)
NZ	Payroll results (New Zealand)
OD	Civil service - staggered payment (directory)
OM	Payroll results (Oman)
PA	Payroll results (Panama)
PE	Payroll results (Peru)
PF	Payroll results (French Polynesia)
PG	Payroll results (Papua New Guinea)
PH	Payroll results (Philippines)
PK	Payroll results (Pakistan)
PL	Payroll results (Poland)
PO	Payroll results (Puerto Rico)

SAP Cluster ID	SAP Cluster Description
PS	Schema
PT	Schema
PU	Payroll results (Paraguay)
PW	Payroll results (Palau)
Q0	Statements (international)
QA	Payroll results (Qatar)
RA	Payroll results (Austria)
RB	Payroll results (Belgium)
RC	Payroll results (Switzerland), all data
RD	Payroll results (Germany), all data
RE	Payroll results (Spain)
RF	Payroll results (France), all data
RG	Payroll results (Britain)
RH	Payroll results (Hungary)
RI	Payroll results (Italy)
RJ	Payroll results (Japan)
RK	Payroll results (Canada)
RL	Payroll results (Malaysia)
RM	Payroll results (Denmark)
RN	Payroll results (Netherlands)
RO	Payroll results (Romania)
RP	Payroll results (Portugal)
RQ	Payroll results (Australia)
RR	Payroll results (Singapore)
RS	Payroll results (Sweden)
RT	Payroll results (Czech Republic)
RU	Payroll results (USA)

SAP Cluster ID	SAP Cluster Description
RV	Payroll results (Norway)
RW	Payroll results (South Africa)
RX	Payroll results (international), all data
SC	Payroll results (Seychelles)
SD	Payroll results (Sudan)
SI	Payroll results (Slovenia)
SK	Payroll results (Slovakia)
SL	Payroll results (Sierra Leone)
SM	Payroll results (San Marino)
SN	Payroll results (Senegal)
SO	Payroll results (Somalia)
SR	Payroll results (Suriname)
SV	Payroll results (El Salvador)
SY	Payroll results (Syria)
SZ	Payroll results (Swaziland)
TD	Payroll results (Chad)
TG	Payroll results (Togo)
TH	Payroll results (Thailand)
TJ	Payroll results (Tajikistan)
TM	Payroll results (Tunisia)
TN	Payroll results (Taiwan)
TO	Payroll results (Tonga)
TP	Payroll results (East Timor)
TR	Payroll results (Turkey)
TT	Payroll results (Trinidad and Tobago)
TU	Payroll results (Turkmenistan)
TV	Payroll results (Tuvalu)

SAP Cluster ID	SAP Cluster Description
TZ	Payroll results (Tanzania)
UA	Payroll results (Ukraine)
UG	Payroll results (Uganda)
UN	Payroll results (United Nations)
UR	Payroll results (Russia)
UY	Payroll results (Uruguay)
UZ	Payroll results (Uzbekistan)
VA	Payroll results (Vatican City)
VC	Payroll results (Saint Vincent)
VE	Payroll results (Venezuela)
VN	Payroll results (Vietnam)
VU	Payroll results (Vanuatu)
WA	Payroll results (Rwanda)
WS	Payroll results (Samoa)
XA	Special run (Austria)
XD	Payroll results (Germany), all data
XE	Payroll results (Spain)
XM	Payroll results (Denmark)
YE	Payroll results (Yemen)
YJ	Cluster YJ (year-end adjustment) (Japan)
YS	Payroll results (Solomon Islands)
YT	Payroll results (Sao Tome)
YU	Payroll results (Serbia and Montenegro)
ZL	Time wage types/work schedule
ZM	Payroll results (Zambia)
ZS	VBL/SPF onetime payment
ZV	SPF data

SAP Cluster ID	SAP Cluster Description
ZW	Payroll results (Zimbabwe)
ZY	Payroll results (Cyprus)

PCL3 Cluster IDs

The retirement job converts encrypted data into XML format depending on the HR cluster ID.

The following table lists the PCL3 cluster IDs that the retirement job transforms into XML:

SAP Cluster ID	SAP Cluster Description
AL	Letters for actions
AN	Notepad for actions
AP	Applicant actions
TY	Applicant data texts

PCL4 Cluster IDs

The retirement job converts encrypted data into XML format depending on the HR cluster ID.

The following table lists the PCL4 cluster IDs that the retirement job transforms into XML:

SAP Cluster ID	SAP Cluster Description
LA	Long-term receipts for PREL
PR	Logging of report start (T599R)
SA	Short-term receipts for PREL

PCL5 Cluster IDs

The retirement job converts SAP encoded data into XML format depending on the HR cluster ID.

The retirement job transforms data into XML for the PCL5 AL cluster ID.

Glossary

access policy

A policy that determines the segments that an individual user can access. Access policies can restrict access to segments based on the application user, database user, or OS user.

administrator

The administrator is the Data Archive super user. The administrator includes privileges for tasks such as creating and managing users, setting up security groups, defining the system profile, configuring repositories, scheduling jobs, creating archive projects, and restoring database archives.

application

A list of brand names that identify a range of ERP applications, such as Oracle and PeopleSoft.

application module

A list of supported Application Modules for a particular Application Version.

application version

A list of versions for a particular application.

archive

Informatica refers to a backed up database as an Archive. From the process perspective, it also means archiving data from an ERP/CRM instance to an online database using business rules.

archive (Database)

A database archive is defined as moving data from an ERP/CRM instance to another database or flat file.

archive (File)

File archive is defined as moving data from an ERP/CRM instance to one or more BCP files.

archive action

Data Archive enables data to be copied from Data Source to Data Target (Archive Only), deleted from Data Source after the backup operation (Archive and Purge), or deleted from the Data Source as specified in the Archive Project (Purge Only).

archive definition

An Archive Definition defines what data is archived, where it is archived from, and where it is archived to.

archive engine

A set of software components that work together to archive data.

archive project

An archive project defines what data is to be archived, where it is archived from, and where it is archived to.

business rule

A business rule is criteria that determines if a transaction is eligible to be archived or moved from the default segment.

business table

Any Table that (with other Tables and Interim Tables) contributes to define an Entity.

column level retention

A rule that you apply to a retention policy to base the retention period for records in an entity on a column in an entity table. The expiration date for each record in the entity equals the retention period plus the date in the column. For example, records in the EMPLOYEE entity expire 10 years after the date stored in the EMP.TERMINATION_DATE column.

custom object

An Object (Application / Entity) defined by an Enterprise Data Manager Developer.

dashboard

A component that generates information about the rate of data growth in a given source database. The information provided by this component is used to plan the archive strategy.

Data Archive

A product of the Information Lifecycle Management Suite, which provides flexible archive / purge functionality that quickly, reduces the overall size of production databases. Archiving for performance, compliance, and retirement are the three main use cases for Data Archive.

Data Archive metadata

Metadata that is specific to a particular ILM product.

data classification

A policy that defines the data set in segments based on criteria that you define, such as dimensions.

data destination

The database containing the archived data.

Data Model Metadata

Metadata that is common to Data Archive.

Data Privacy

A product of the Information Lifecycle Management Suite, which provides a comprehensive automated data privacy solution for sensitive data stored in everything from mission critical production systems to internal QA and development environments to long term data repositories over a heterogeneous IT infrastructure.

data source

The database containing the data to be archived.

Data Subset

A product of the Information Lifecycle Management Software, which enables organizations to create smaller, targeted databases for project teams that maintain application integrity while taking up a fraction of the space.

data target

The database containing the archived data. This might alternatively be referred to as History Database.

Data Vault access role

User-defined role that restricts access to archived data in the Data Discovery portal. You assign access roles directly to entities and users. For further restriction, you can assign an access role to the archive or retirement project.

de-referencing

The process of specifying inter-Column Constraints for JOIN queries when more than two Tables are involved. Enterprise Data Manager dictates that two Tables are specified first and then a Column in the Second Table is De-Referenced to specify a Third Table (and so on) for building the final JOIN query.

dimension

An attribute that defines the criteria, such as time, to create segments when you perform smart partitioning.

dimension rule

A rule that defines the parameters, such as date range, to create segments associated with the dimension.

dimension slice

A subset of dimensional data based on criteria that you configure for the dimension.

entity

An entity is a hierarchy of ERP/CRM tables whose data collectively comprise a set of business transactions. Each table in the hierarchy is connected to one or more other tables via primary key/foreign key relationships. Entity table relationships are defined in a set of metadata tables.

ERP application

ERP applications are software suites used to create business transaction documents such as purchase orders and sales orders.

expiration date

The date at which a record is eligible for deletion. The expiration date equals the retention period plus a date that is determined by the retention policy rules. The Purge Expired Records job uses this date to determine which records to delete from the Data Vault. A record expires on its expiration date at 12:00:00 a.m. The time is local to the machine that hosts the Data Vault Service. A record with an indefinite retention period does not have an expiration date.

expression-based retention

A rule that you apply to a retention policy to base the retention period for records in an entity on a date value that an expression returns. The expiration date for each record in the entity equals the retention period plus the date value from the expression. For example, records in the CUSTOMER table expire five years after the last order date, which is stored as an integer.

formula

A procedure that determines the value of a new time dimension slice based on the end value of the previous dimension slice.

general retention

A rule that you apply to a retention policy to base the retention period for records in an entity on the archive job date. The expiration date for each record in the entity equals the retention period plus the archive job date. For example, records in an entity expire five years after the archive job date.

history database

Database used to store data that is marked for archive in a database archive project.

home database

The database containing metadata and other tables used to persist application data. This is contained in the Schema, usually known as "AMHOME".

interim

Temporary tables generated for an Entity for data archive.

job

A process scheduled for running at a particular point in time. Jobs in Data Archive include Archive Projects or Standalone Jobs.

metadata

Metadata is data about data. It not only contains details about the structure of database tables and objects, but also information on how data is extracted, transformed, and loaded from source to target. It can also contain information about the origin of the data.

Metadata contains Business Rules that determines whether a given transaction is archivable.

production database

Database which stores Transaction data generated by the ERP Application.

restore (Cycle)

A restore operation based on a pre-created Archive Project.

restore (Full)

A restore operation that involves complete restoration of data from Data Target to Data Source.

restore (Transaction)

A restore operation that is carried out by precisely specifying an Entity, Interim Tables and related WHERE clauses for relevant data extractions.

retention management

The process of storing records in the Data Vault and deleting records from the Data Vault. The retention management process allows you to create retention policies for records. You can create retention policies for records as part of a data archive project or a retirement archive project.

retention policy

The set of rules that determine the retention period for records in an entity. For example, an organization must retain insurance policy records for five years after the most recent message date in the policy, claims, or client message table. You apply a retention policy to an entity. You can apply a retention policy rule to a subset of records in the entity.

security group

A Security Group is used to limit a User (during an Archive Project) to archive only data that conforms to certain scope restrictions for an Entity.

segment

A set of data that you create through smart partitioning to optimize application performance. After you run a segmentation policy to create segments, you can perform operations such as limiting the data access or compressing the data.

segmentation group

A group of tables based on a business function, such as order management or human resources, for which you want to perform smart partitioning. A segmentation group defines the database and application relationships for the tables.

segmentation policy

A policy that defines the data classification that you apply to a segmentation group. When you create a policy, you also choose the method of periodic segment creation and schedule the policy to run.

segment set

A set of segments that you group together so you can perform a single action on all the segments, such as compression.

smart partitioning

A process that divides application data into segments based on rules and dimensions that you configure.

staging schema

A staging schema is created at the Data Source for validations during data backup.

standard object

An Object (Application / Entity) which is pre-defined by the ERP Application and is a candidate for Import into Enterprise Data Manager.

system-defined role

Default role that includes a predefined set of privileges. You assign roles to users. Any user with the role can perform all of the privileges that are included in the role.

system profile

Technical information about certain parameters, usually related to the Home Schema, or Archive Users comprises the System Profile.

transaction data

Transaction data contain the information within the business documents created using the master data, such as purchase orders, sales orders etc. Transactional Data can change very often and is not constant. Transaction data is created using ERP applications.

Transaction data is located in relational database table hierarchies. These hierarchies enforce top-down data dependencies, in which a parent table has one or more child tables whose data is linked together using primary key/foreign key relationships.

translation column

A column which is a candidate for a JOIN query on two or more Tables. This is applicable to Independent Archive, as Referential Data is backed up exclusively for XML based Archives and not Database archive, which contain only Transaction Data.

Note that such a scenario occurs when a Data Archive job execution extracts data from two Tables that are indirectly related to each other through Constrains on an intermediary Table.

user

Data Archive is accessible through an authentication mechanism, only to individuals who have an account created by the Administrator. The level of access is governed by the assigned system-defined roles.

user profile

Basic information about a user, such as name, email address, password, and role assignments.

INDEX

U

Update Retention Policy job
reports

Update Retention Policy job *(continued)*
running