

HP DecisionCenter

for the AIX, Windows®, and Itanium® operating systems

Software Version: 2.00

Best Practices for Data Warehouse Administration Guide

Document Release Date: December 2007

Software Release Date: December 2007

Last Updated: 12 October 2007



Legal Notices

Warranty

The only warranties for HP products and services are set forth in the express warranty statements accompanying such products and services. Nothing herein should be construed as constituting an additional warranty. HP shall not be liable for technical or editorial errors or omissions contained herein.

The information contained herein is subject to change without notice.

Restricted Rights Legend

Confidential computer software. Valid license from HP required for possession, use or copying. Consistent with FAR 12.211 and 12.212, Commercial Computer Software, Computer Software Documentation, and Technical Data for Commercial Items are licensed to the U.S. Government under vendor's standard commercial license.

Copyright Notices

© Copyright 2007–2008 Hewlett-Packard Development Company, L.P.

Trademark Notices

Adobe® is a trademark of Adobe Systems Incorporated.

Microsoft® and Windows® are U.S. registered trademarks of Microsoft Corporation.

UNIX® is a registered trademark of The Open Group.

Support

You can visit the HP Software support web site at:

www.hp.com/go/hpsoftwaresupport

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

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

- Search for knowledge documents of interest
- Submit and track support cases and enhancement requests
- Download software patches
- Manage support contracts
- Look up HP support contacts
- Review information about available services
- Enter into discussions with other software customers
- Research and register for software training

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

To find more information about access levels, go to:

www.hp.com/managementsoftware/access_level

To register for an HP Passport ID, go to:

www.managementsoftware.hp.com/passport-registration.html

Contents

Best Practices for Data Warehouse Administrators	6
ETL data flow.....	7
Connect-It mappings.....	8
Data synchronization.....	8
Aggregation.....	9
Hierarchy.....	9
ETL validation tools.....	10
Validating the ETL process	11
Validation status	11
Backup file names.....	12
How do I configure the etl.properties file?.....	12
How do I create the validation logging database?.....	14
How do I monitor the Connect-It log?	15
How do I use another scenario name?	16
How do I use another Connect-It service name?.....	17
Data warehouse utilities.....	18
Verification utilities.....	18
Diagnostic utilities.....	20
Disaster recovery process utilities.....	21
JDBC connections.....	24
Best practices: Data warehouse administration.....	25
Add an index to ServiceCenter data	25
Administer databases	28
Avoid NULL keys.....	29
Configure scenario schedulers.....	30
Manage the schema.....	30
Monitor Business Objects reporting server.....	31
Monitor the ETL process	31
Review system tables	32
RDS_CIT_LOG table	33
RDS_CITSCHPTR_LOG table.....	34
RDS_CITSYNC_LOG table.....	35
RDS_DBINFO table	36
RDS_ETLSYNC_LOG table.....	37
RDS_LOGININFO table.....	38
RDS_SCDCOLUMN table.....	40
RDS_SEQUENCE table	41
RDS_TIMEDIM_D table	41
RDS_UNIQUECOLUMN table.....	42
Data warehouse checklist	44
General environment	45
Database issues.....	46
Data errors.....	47

Troubleshooting hints	48
Data warehouse architecture.....	49
Design	50
Naming conventions.....	51
Dimension table system keys	52
Fact table ETL processes	52
Main dimension table changes.....	53
Referenced dimension table changes	53
Universe filter condition changes	53
Example: Data warehouse customization	54
Example: Index creation	55

Best Practices for Data Warehouse Administrators

The DecisionCenter data warehouse is the foundation of the analytic and optimization modules. Administrators can maximize the data warehouse performance with routine tasks that monitor the Extract, Load, and Transform (ETL) processes.

By adhering to the recommended best practices, administrators can ensure the DecisionCenter data warehouse reliability in production environments. The data warehouse utilities can help to:

- Provide data verification.
- Diagnose problems.
- Recover records.

See the *HP DecisionCenter Installation Guide* for detailed information about setting up your data warehouse, data sources, data targets, and mapping information. See the **HP Software Support** web site (<http://www.hp.com/go/hpsoftwaresupport/>) if you need to customize an existing data warehouse database. Most of the support areas require that you register as an HP Passport user and sign in. Many also require an active support contract. To register for an HP Passport ID, go to the **HP Passport Registration** web site (<http://www.managementsoftware.hp.com/passport-registration.html>).

Related topics

[ETL data flow](#)

[ETL validation tools](#)

[JDBC connections](#)

[Best practices: Data warehouse administration](#)

[Data warehouse checklist](#)

[Data warehouse architecture](#)

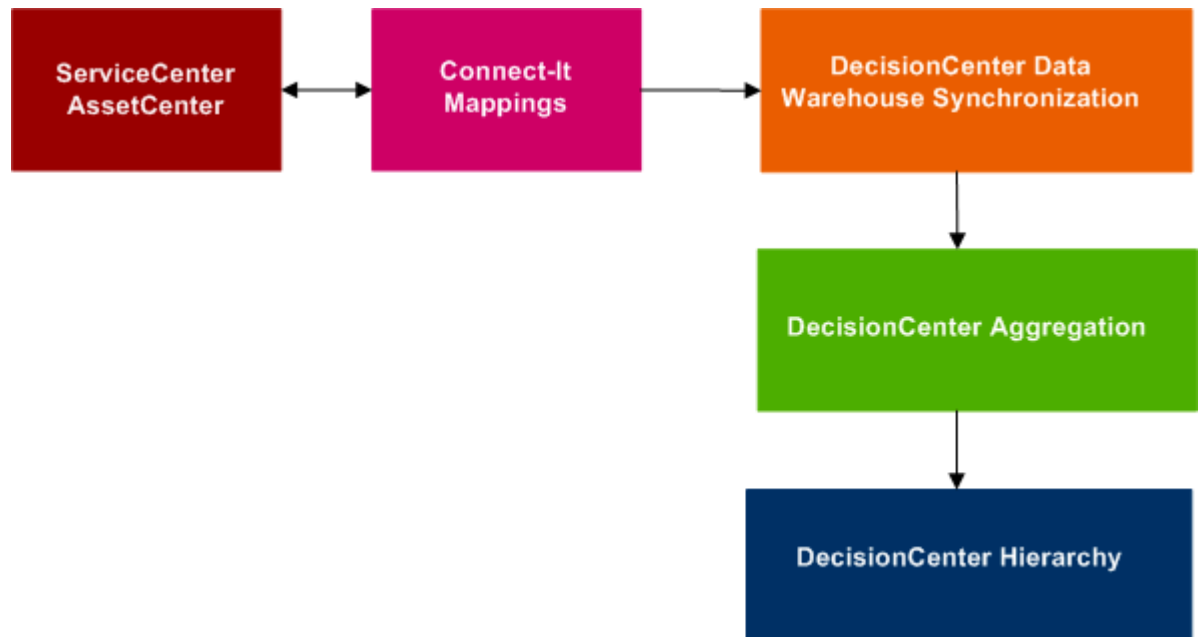
[Fact table ETL processes](#)

[Example: Data warehouse customization](#)

[Example: Index creation](#)

ETL data flow

The DecisionCenter data warehouse contains multiple data marts for IT performance analytics, Online Analytical Processing (OLAP), and operational detail information for online ad hoc query and operational reporting. The Extract, Transform, and Load (ETL) process for the DecisionCenter data warehouse handles the transfer of data from ServiceCenter and AssetCenter to provide insight and actionable IT performance analytics.



Related topics

Best Practices for Data Warehouse Administrators

ETL validation tools

JDBC connections

Best practices: Data warehouse administration

Data warehouse checklist

Data warehouse architecture

Fact table ETL processes

Example: Data warehouse customization

Example: Index creation

Connect-It mappings

Data synchronization

Aggregation

Hierarchy

Connect-It mappings

The Connect-It tool uses mapping scenarios to extract data from the source ServiceCenter or AssetCenter tables to data warehouse dimension and operational tables. Direct mappings populate operational tables while dimension mappings populate dimension tables.

Mapping type	Description
Direct	<p>The mappings for operational tables are straight mappings. Most of the transformations convert data types; for example, converting dates and array data into long text. The data warehouse delivers the transformation functions to the Connect-It component and distributes them as the rds.bas file in the Connect-It configuration data warehouse directory.</p> <p>Note: All the date values are stored as local time in the data warehouse. For ServiceCenter, the dates are converted from GMT in the ServiceCenter source to the local time of the DecisionCenter data warehouse server.</p>
Dimension	<p>The main differences between dimension table mappings and direct mappings are surrogate key generation and slowly changing dimension (SCD) processing. The Connect-It data warehouse configuration components deliver all the transformations.</p> <p>Note: Dimension surrogate keys are populated based on the DecisionCenter data warehouse RDS_SEQUENCE system table. The sequence table values are refreshed by each cycle of the DecisionCenter data warehouse synchronization.</p>

Related topics

[ETL data flow](#)

[Data synchronization](#)

[Aggregation](#)

[Hierarchy](#)

Data synchronization

DecisionCenter allows you to synchronize user records with the data warehouse. You can also synchronize your AssetCenter and ServiceCenter deletion event records in the data warehouse.

You can use the HP Connect-It Scheduler Editor to schedule automatic data synchronization.

The following table lists the files that DecisionCenter uses or creates during the synchronization process.

File name	Description
rds_ac.scn rds_sc.scn	Contains the data scenario information.
rds_ac.ini rds_sc.ini	Contains the tables that are in the sync and the time of the next scheduled data sync.
rds_ac.log rds_sc.log	Contains the data sync activity records. The system continuously appends information to this file while the service is active. You can check the log file to monitor the status of your data synchronization.

You can find the SCN and INI files in the **cit** directory of the data warehouse installation. The LOG files are in the **logs** directory of the data warehouse installation. The data warehouse installation default path is:

```
\\...\HP\DecisionCenter 2.00\ITPA-SM for ServiceCenter and Service Manager
\\...\HP\DecisionCenter 2.00\ITPA-AM for AssetCenter
```

Related topics

[ETL data flow](#)

[Connect-It mappings](#)

[Aggregation](#)

[Hierarchy](#)

Aggregation

Based on definitions in the data warehouse XML schema file, the process recalculates all DecisionCenter aggregates. To achieve the best performance, set the number of aggregations and the fields defined for each aggregation to the minimum required. All summarized measurements defined in the Business Objects universes use aggregate awareness functionality to take advantage of these pre-aggregated calculations. Set the aggregation synchronization interval based on the refresh scheduling of reports and how the frequency of the Connect-It mapping schedules.

Related topics

[ETL data flow](#)

[Connect-It mappings](#)

[Data synchronization](#)

[Hierarchy](#)

Hierarchy

The ETL process generates all DecisionCenter hierarchies based on the hierarchy section defined in the data warehouse XML schema file. For optimum performance, set the number of hierarchies and the levels of the hierarchies to the minimum required.

The DecisionCenter 1.00 hot fix patch includes synchronization transaction work control and logging for the synchronization of fact tables, aggregations and hierarchy tables.

The DecisionCenter 2.00 synchronization process has improvements in the following areas:

- Auto-indexing for dimension tables, fact tables and association tables
- Deletion process
- Logging
- Exception handling

Related topics

[ETL data flow](#)

[Connect-It mappings](#)

[Data synchronization](#)

[Aggregation](#)

ETL validation tools

DecisionCenter has a set of validation tools to verify that your Extract, Transform, and Load (ETL) processes ran correctly and that all records transferred successfully to the DecisionCenter data warehouse. These validation tools run in conjunction with the HP Connect-It scenarios that transfer data from external applications to the DecisionCenter data warehouse. When errors occur, the validation tools send e-mail notifications. There are three ETL validation tools.

- The Pre-Mapping tool creates a backup of Connect-It synchronization scheduler pointers, recovers all data warehouse sequence numbers if necessary, and sends e-mail notification if the ETL process failed.
- The Post-Mapping tool checks the last Connect-It scenario synchronization session and mapping logs. If session errors or data mapping errors occur, the Post-Mapping tool sends e-mail notification, and sets a scenario failure flag.
- The Post-Sync tool verifies the Connect-It synchronization condition of the mapping and ETL processes. If errors occur, the Post-Sync tool sends e-mail notification, and stops the Connect-It service.

For more information about Connect-It ETL processes, see the *DecisionCenter Installation Guide*.

Related topics

[Best practices for Data Warehouse Administrators](#)

[ETL data flow](#)

[JDBC connections](#)

[Best practices: Data warehouse administration](#)

[Data warehouse checklist](#)

[Data warehouse architecture](#)

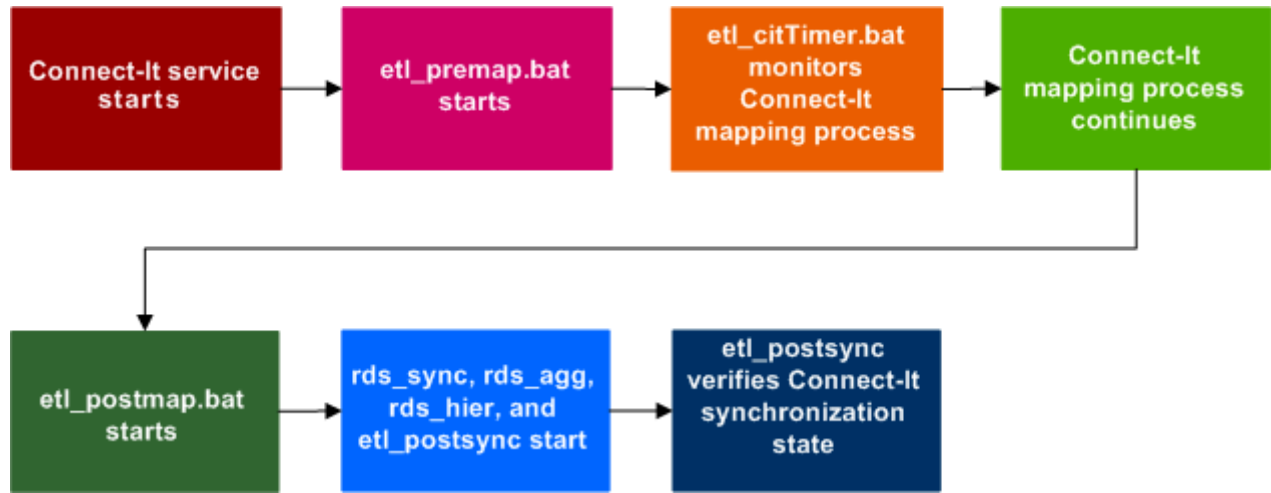
[Fact table ETL processes](#)

[Example: Data warehouse customization](#)

[Example: Index creation](#)

Validating the ETL process

When the ETL Validator is on during the Extract, Transform, and Load (ETL) process, the Connect-It service performs an update synchronization.



The Pre-Mapping tool triggers a program that monitors the Connect-It mapping process. The Post-Mapping tool then checks synchronization sessions and mapping logs. The Post-Sync tool verifies the Connect-It synchronization state. If session errors occur, administrators receive e-mail notification.

It is important that you verify the completeness of the ETL process manually before you begin the validation. To adhere to best practices, start the validators after you run the initial and update data warehouse synchronizations.

The validation tools require a logging database to monitor the log files and a configured etl.properties file. See [Related topics](#) to create the database and configure the file.

Validation status

When validation is successful, the administrator receives the following e-mail:

```
Please check CIT logs  
CIT Session ID = 50432
```

The **CIT Session ID** statement indicates that the process succeeded.

When validation fails, DecisionCenter sends an e-mail to the administrator. For example:

```
Please check CIT logs  
RDS STATE_FLAG = CIT_START  
(ETL Post Sync Validator)
```

The **STATE_FLAG** statement indicates that the Connect-It ETL process failed. You can search for the cause of the failure by viewing the appropriate:

- Connect-It logs.
- Validator logs.
- Validation failure troubleshooting guide.

Note: If you set the **stopindicator** parameter in the etl.properties file to 1, you must restart the Connect-It service after you correct the problem.

Two additional tables provide troubleshooting information.

Table	Validation information
RDS_LOGININFO	Specifies the status of the validation
RDS_CITSYNC_LOG	Verifies table name and record count information for source and target connectors mentioned in the e-mail.

Backup file names

The validation process backs up and renames the scenario INI and service LOG files to prevent overwriting the files when the scenario runs. By renaming the files, you can revert to a previous INI file or check a specific LOG for errors.

Original file name	Renamed file name
scenarioName.ini	scenarioName.ini_timestamp
serviceName.log	serviceName.log_timestamp

Related topics

[ETL validation tools](#)

[How do I create the validating logging database?](#)

[How do I configure the etl.properties file?](#)

[How do I monitor the Connect-It log?](#)

[How do I use another scenario name?](#)

[How do I use another Connect-It service name?](#)

[Data warehouse utilities](#)

How do I configure the etl.properties file?

Make sure that you set up the validation logging database before you configure the etl.properties file.

Note: Remove the commented out sections and replace the appropriate lines with the database type of your Connect-It database as well as connection information.

- 1 Open the etl.properties file with a text editor:

```
\\... \HP \DecisionCenter 2.00 \ITPA-SM \conf \
```

- 2 Update the default parameters with the correct information for your installation.

Parameter	Description
cit.os	Operating system
cit.rdbms.type	RDBMS type; either SQL Server or Oracle
cit.user	User name for the Connect-It logging database
cit.password	Password for the Connect-It logging database
etl.emailnotify	Flag that turns e-mail notification on
etl.srconnector	Name of the datasource connector in the rds_sc scenario
etl.tarconnector	Name of the target connector that points to the data warehouse in the rds_sc scenario
etl.emailserver	Name of the e-mail server
etl.adminemail	Address of the user who receives the e-mail notifications
etl.stopindicator	Indicator that identifies when to stop the Connect-It server if errors are detected. If validation fails, one of the following actions occurs: <ul style="list-style-type: none"> • When the value is 1, the Connect-It service is not stopped and runs again at its scheduled time. To maintain data integrity during the pre-mapping restore pointer process, prior to running the next update synchronization, the rds_sc.ini.bak file is restored to rds_sc.ini. • When the value is 2, the Connect-It service that runs the rds_sc update synchronization scenario is stopped, enabling the Administrator to analyze the problems.
cit.parselog	Connect-It log that contains diagnostic information to help identify where critical errors occur. <ul style="list-style-type: none"> • When set to true, the validators continue to run despite critical errors in the Connect-It log. This is the recommended option because some critical errors are not fatal and not related to data integrity. • If set to false, when the validators find a critical Type=1 error message, the validation process immediately fails and an e-mail notification is sent if the etl.emailnotify parameter is true.
jdbc.driver	JDBC driver for RDBMS

Parameter	Description
jdbc.url	JDBC url

Related topics

[ETL validation tools](#)

[Validating the ETL process](#)

[How do I create the validating logging database?](#)

[How do I monitor the Connect-It log?](#)

[How do I use another scenario name?](#)

[How do I use another Connect-It service name?](#)

[Data warehouse utilities](#)

How do I create the validation logging database?

When you set up a database that monitors the validation logs, you must create new connections and tables, then configure the database. If you use SQL Server, the Connect-It database must be case insensitive.

- 1 In SQL Server or Oracle, create database `citdb` and user `cit_dba`.
- 2 Click **Start > Programs > HP OpenView > Connect-It 3.8 en > Connect-It Scenario Builder**.
- 3 To create the connection, click the **Manage Connections** icon next to the Connection field.
 - a In the Manage connections dialog box, click **New**.
 - b Type the Name (for example, `citdb`) and Description.
 - c From the **Engine** drop-down list, select the database engine.
 - d In the selected engine, complete the connection fields.
 - e Click **Create**, then click **Close**.
- 4 From the **Connect-It Monitors database administration** dialog box, select the connection you created in Step 3 to create tables in the database.
 - a Click **Create Database**.
 - b Type the password for **Admin**.
Note: This is the Admin user for the Connect-It database. Make sure that you record the password.
 - c Click **OK** to see the list of tables that are being created. A message at the bottom of the page indicates that the database was successfully created.
 - d Check the database to verify that the tables exist.
 - e Click **Close**.
- 5 From the Connect-It main window, click **Monitors > Configure Monitors** to configure the database.
 - a Select the **session Backups** tab and **Database** tab.
 - b Set the log settings to **No Filter**, **Reject**, and **No Filter**.

- c Clear the **Limit data backed up for the scenario** option.
 - d Check **Enable monitor**.
 - e Select the connection you created in Step 3.
 - f Login as **Admin** with the password you created in Step 4.
 - g Click **Test the connection** to test the connection information.
 - h Click **Save as the default configuration**.
 - i Click **OK**.
- 6 Save the scenario.

Related topics

[ETL validation tools](#)

[Validating the ETL process](#)

[How do I configure the etl.properties file?](#)

[How do I monitor the Connect-It log?](#)

[How do I use another scenario name?](#)

[How do I use another Connect-It service name?](#)

[Data warehouse utilities](#)

How do I monitor the Connect-It log?

You can change the Post-Mapping and Post-Sync schedules in the etl_citTimer.bat file that monitors the Connect-It log.

- 1 Open the rds.scp file with a text editor:

```
\\... \HP \OpenView \Connect-It 3.80 en \config \rds \const \
```

- 2 Change ETL Validator from 0 to 1.
- 3 Save and close the file.

- 4 Open the etl_citTimer.bat file with a text editor:

```
\\... \HP \DecisionCenter 2.00 \ITPA-SM \common \bin \
```

- 5 Set JAVA_HOME to the name of the directory identified during installation. For example:

```
set JAVA_HOME=C:/j2sdk1.4.2_12
```

Set RDS_HOME to the name of the data warehouse installation directory. For example:

```
set RDS_HOME=C:/Program Files/HP/DecisionCenter 2.00/ITPA-SM/
```

- 6 Change **args0** through **args4** for com.hp.decisioncenter.dwutil.CitTimer.

Argument	Description	Default value
args0: fileName	The name and location of the Connect-It log file name where you monitor the changes	\\...\HP OpenView\Connect-It 3.80 en\bin\conitgui.log
args1: initSeconds	The length of time, in seconds, that the process is dormant before checking for changes	10
args2: sleepSeconds	Frequency to check for file changes	60
args3: stabilizeSeconds	Number of seconds the file time has no changes	180
args4: processToRun	After the file time stabilizes, the process that runs	\\...\HP\DecisionCenter 2.00\ITPA-SM\common\bin\etl_postmap.bat

- 7 Save and close the file.

Related topics

[ETL validation tools](#)

[Validating the ETL process](#)

[How do I create the validating logging database?](#)

[How do I configure the etl.properties file?](#)

[How do I use another scenario name?](#)

[How do I use another Connect-It service name?](#)

[Data warehouse utilities](#)

How do I use another scenario name?

The default Connect-It scenario is `rds_sc.scn`. You must change the validation tools to reflect your scenario name if you do not use the default.

- 1 Open the validation tool with a text editor.
- 2 Replace `rds_sc` with the name of your scenario. `RDS_HOME` is the name of your data warehouse installation directory.

Tool	File name	Replace <code>rds_sc</code> to set args0
Pre-Mapping	<RDS_HOME>\common\bin\etl_premap.bat	com.hp.decisioncenter.dwutil.ETLPreRun <code>rds_sc</code>

Post-Mapping	<RDS_HOME>\common\bin\etl_postmap.bat	com.hp.decisioncenter.dwutil.ETLRun rds_sc
Post-Sync	<RDS_HOME>\common\bin\etl_postsync.bat	com.hp.decisioncenter.dwutil.ETLPostSync rds_sc

- 3 Save and close.

Related topics

[ETL validation tools](#)

[Validating the ETL process](#)

[How do I create the validating logging database?](#)

[How do I configure the etl.properties file?](#)

[How do I monitor the Connect-It log?](#)

[How do I use another Connect-It service name?](#)

[Data warehouse utilities](#)

How do I use another Connect-It service name?

The default Connect-It service is rds_sc. You must change the validation tools to reflect your service name if you do not use the default.

- 1 Open the validation tool with a text editor.
- 2 Replace rds_sc with the name of your service. RDS_HOME is the name of your data warehouse installation directory.

Tool	File name	Replace rds_sc to set args1
Pre-Mapping	<RDS_HOME>\common\bin\etl_premap.bat	com.hp.decisioncenter.dwutil.ETLRun rds_sc rds_sc
Post-Mapping	<RDS_HOME>\common\bin\etl_postmap.bat	com.hp.decisioncenter.dwutil.ETLPostSync rds_sc rds_sc
Post-Sync	<RDS_HOME>\common\bin\etl_postsync.bat	com.hp.decisioncenter.dwutil.ETLPostSync rds_sc rds_sc

- 3 Save and close.

Related topics

[ETL validation tools](#)

[Validating the ETL process](#)

[How do I create the validating logging database?](#)

[How do I configure the etl.properties file?](#)

[How do I monitor the Connect-It log?](#)

[How do I use another scenario name?](#)

[Data warehouse utilities](#)

Data warehouse utilities

DecisionCenter utilities help administrators monitor the condition of the data warehouse. The utility tasks focus on the Extract, Transform, and Load (ETL) process to ensure data reliability.

You can run the utilities to monitor and view:

- Data verification
- Diagnosis
- Disaster recovery processes

Related topics

[ETL validation tools](#)

[Verification utilities](#)

[Diagnostic utilities](#)

[Disaster recovery process utilities](#)

Verification utilities

DecisionCenter data warehouse verification utilities validate the data warehouse tables and the source tables.

- 1 Run the verification `RDSVerificationPrep.sql` setup scripts to store the verification results. The utility creates verification check point tables:
 - `RDS_CHECKPOINT_DimSeq`
 - `RDS_CHECKPOINT_FactDIM`
 - `RDS_CHECKPOINT_RDSSC`
 - `RDS_CHECKPOINT_RDSSC_List`
- 2 Run the verification `RDSVerificationProcs.sql` scripts to create the verification procedures.
- 3 Set ServiceCenter offline to stop transaction activities on ServiceCenter so that you can verify the data or create a data snapshot of data sources with the necessary database link.
- 4 Complete the ETL synchronization process.
- 5 Run the `RDSVerificationExec.sql` templates scripts to compare ServiceCenter tables with data warehouse tables, dimension tables with fact tables, or data warehouse sequence number validation.

The following table contains the verification procedures.

Utility	Description
DC_COMPARISON_SOURCE	<p>Purpose: Compare the counts of source tables with data warehouse tables</p> <p>Insert the comparison results into RDS_CHECKPOINT_RDSSC</p> <p>Usage: Run DC_Comparison_Source(rdstable_type, rdsName, rdsKey, sourceName, rdsSourceKey);</p> <p>Example: execute DC_Comparison_Source('DIMENSION','INCIDENT', 'Z_RDSINCIDENT_DID','INCIDENTSM1','INCIDENT_ID');</p>
DC_COMPARISON_LIST_SCDC	<p>Purpose: List the keys for mismatched source tables with data warehouse tables</p> <p>Insert the comparison results into RDS_CHECKPOINT_RDSSC_LIS</p> <p>Usage: Run DC_Comparison_Source(rdstable_type, rdsName, rdsKey, sourceName, rdsSourceKey);</p> <p>Example: execute DecisionCenter DC_Comparison_List_SCDC ('DIMENSION','INCIDENT', 'Z_RDSINCIDENT_DID','INCIDENTSM1','INCIDENT_ID');</p>
DC_COMPARISON_LIST_DCSC	<p>Purpose: List the keys for mismatched data warehouse tables with source tables</p> <p>Insert the comparison results into RDS_CHECKPOINT_RDSSC_LIST</p> <p>Usage: Run DC_Comparison_List_DCSC (rdstable_type, rdsName, rdsKey, sourceName, rdsSourceKey);</p> <p>Example: execute DecisionCenter DC_Comparison_List_SCDC ('DIMENSION','INCIDENT', 'Z_RDSINCIDENT_DID','INCIDENTSM1','INCIDENT_ID');</p>
DC_COMPARISON_FACT_DIMENSION	<p>Purpose: Compare the counts of dimension and fact tables in the data warehouse</p> <p>Store the compared count results in RDS_CHECKPOINT_FACTDIM</p> <p>Store the compared ID key values in _CHECKPOINT_RDSSC_LIST</p> <p>Usage: Run DC_Comparison_Fact_Dimension(factTableName);</p> <p>Example: execute DC_Comparison_Fact_Dimension('incident_f');</p>

Utility	Description
CallFactDimProc	<p>Purpose: call procedure DC_Comparison_Fact_Dimension for all fact tables</p> <p>Usage: Run CallFactDimProc();</p> <p>Example: execute CallFactDimProc();</p>
RDSSEQCHECKPROC	<p>Purpose: Validate to ensure all the sequence numbers are current in the data warehouse</p> <p>Save compared results RDS_CHECKPOINT_DIMSEQ</p> <p>Usage: Run RDSSeqCheckProc();</p> <p>Example: execute RDSSeqCheckProc();</p>

Related topics

[Data warehouse utilities](#)

[Diagnostic utilities](#)

[Disaster recovery process utilities](#)

Diagnostic utilities

The verification utilities store the results in check point tables that compare the number of records between the source data and data warehouse and list mismatched records.

The checkpoint tables contain information that can help you to diagnose inconsistencies in the data.

Table	Description
RDS_CHECKPOINT_DimSeq	Stores dimensionDID, dimensionSeqCount, and diffNum (dimensionSeqCount- dimensionDID) data The corrected result should be Diffnum=0.
RDS_CHECKPOINT_FactDIM	Stores FactCount, DimensionCount and diffNum (FactCount-DimensionCount) data The corrected result should be Diffnum=0.
RDS_CHECKPOINT_RDSSC	Stores all the compared counts The counts in this table are critical in defining the recovery process.
RDS_CHECKPOINT_RDSSC_List	Lists all the mismatched records between ServiceCenter and the data warehouse

RDS_CHECKPOINT_RDSSC query

The following query identifies the number of records in ServiceCenter that the data warehouse does not recognize.

```
select * from RDS_CHECKPOINT_RDSSC order by scan_date desc
```

The RDS_CHECKPOINT_RDSSC_LIST table stores and lists all the mismatched records between ServiceCenter and the data warehouse.

```
select * from RDS_CHECKPOINT_RDSSC_LIST
```

The record counts in RDS_CHECKPOINT_RDSSC_LIST and RDS_CHECKPOINT_RDSSC should be the same (RDS_DEL_COUNT+SC_ADD_COUNT).

The RDS_CHECKPOINT_RDSSC table stores the number of compared records.

Table column	Description
DIFFNUM = SC_COUNT - RDS_COUNT	Identifies the result of the difference between the number of ServiceCenter records and the number of records in the data warehouse.
sc_add_count	Indicates the number of missing records that are in ServiceCenter but not in the data warehouse.
rds_del_count	Indicates the number of missed deletion records from ServiceCenter that are not in the data warehouse.

In a static environment, a successful synchronization process is

```
sc_count = rds_count AND rds_del_count = 0 AND sc_add_count = 0
```

Related topics

[Data warehouse utilities](#)

[Verification utilities](#)

[Disaster recovery process utilities](#)

Disaster recovery process utilities

DecisionCenter data warehouse disaster recovery utilities identify missing records and provide information to reactivate records.

After you identify missing records, you can recover them when you synchronize the data warehouse records.

- 1 If RDS_CHECKPOINT_RDSSC RDS_ADD_COUNT does not equal zero, check the time periods in the SYSMODTIME column.
- 6 Based on the identified missing periods, update ServiceCenter source table records where the **sysmodtime** values are equal to the missing periods.
- 7 Set **sysmodtime** to the current timestamp.
- 8 Run the data warehouse synchronization process.

The following associates.sql script example demonstrates how to correct associate tables with array fields that are activated in ServiceCenter.

```

/***** Reactivate the deleted record *****/
for directMapping table
  dimension table
  fact table
update SCRELATION c
set c.Z_RDSACTIVESTATUSIND = 'Y', c.Z_RDSDELETEDDATE = null
where c.DEPEND || c.DEPEND_FILENAME || c.SOURCE || c.SOURCE_FILENAME in
(select sckeyname from RDS_CHECKPOINT_RDSSC_LIST
where table_name='SCRELATION' and COMPINDICATOR='SC-DC' )
;
commit;

```

The following table contains the disaster recovery process procedures.

Utility	Description
DC_FACT_FIX	<p>Purpose: Update the data warehouse dimension table for missing records in the fact table</p> <p>After executing DC_FACT_FIX for all the missing dimension records, run rds_sync.bat to recover fact table missing records</p> <p>Usage: exec DC_FACT_FIX(tableName IN VARCHAR2);</p> <p>Example: exec DC_FACT_FIX('INCIDENT');</p>

Utility	Description
DC_DIM_FIX	<p>Purpose: Insure there is only one active dimension record for each ServiceCenter record</p> <p>Usage: Run DC_DIM_FIX(tableName IN VARCHAR2 , rdsSourceKey IN VARCHAR2);</p> <p>Example: exec DecisionCenter DC_DIM_FIX('OPERATOR', 'NAME');</p> <p>If RDS_CHECKPOINT_RDSSC RDS_COUNT is greater than SC_COUNT but there are no records in RDS_CHECKPOINT_RDSSC_LIST for the same table and COMPINDICATOR equal to 'SC-DC', run the procedure in Oracle SQLPlus as:</p> <pre>EXEC DC_DIM_FIX('XXX_table', 'Unique_key');</pre>
DC_Delete_Clean	<p>Purpose: Update data warehouse missing deleted record</p> <p>Usage: Run DC_Delete_Clean(rdstable_type, rdsName, rdsSourceKey);</p> <p>Example: execute DC_Delete_Clean('DIMENSION', 'INCIDENT', 'INCIDENT_ID');</p>
DC_Reactivate	<p>Purpose: Reactivate data warehouse records from inactive</p> <p>Usage: Run DC_Reactivate (rdstable_type, rdsName, rdsSourceKey);</p> <p>Example: exec DC_Reactivate ('DIMENSION', 'INCIDENT', 'INCIDENT_ID');</p> <p>If RDS_CHECKPOINT_RDSSC diffnum does not equal zero and there are records with the same table in RDS_CHECKPOINT_RDSSC_LIST, for example, CM3R_D table, run the statement in Oracle SQLPlus</p> <pre>exec DC_Reactivate('DIMENSION', 'CM3R', 'NUMBERPRGN');</pre>

Related topics

[Data warehouse utilities](#)

[Verification utilities](#)

[Diagnostic utilities](#)

JDBC connections

You can configure JDBC thick clients to work with Oracle Real Application Clusters (RAC).

The following examples show how to set up the DecisionCenter data warehouse server as an Oracle client tnsname.ora Transparent Application Failover (TAF).

```
DC =
(DESCRIPTION =
  (ADDRESS_LIST =
    (ADDRESS = (PROTOCOL = TCP)(HOST = <first instance virtual ip> )
    (PORT = <first instance's listener port>))
    (ADDRESS = (PROTOCOL = TCP)(HOST = <second instance virtual ip> )
    (PORT = <second instance's listener port>))
    (... entry for each instance...)
  )
  (LOAD_BALANCE = off)
  (FAILOVER = on)
)
(CONNECT_DATA =
(SERVER = DEDICATED)
(SERVICE_NAME = <DB service name>)
)
)
```

For example:

```
DC =
(DESCRIPTION =
(Load_Balance = OFF)
(FAILOVER = ON)
(ADDRESS_LIST =
  (ADDRESS = (PROTOCOL = TCP)(HOST = pjlworm-vip)(PORT = 39005))
  (ADDRESS = (PROTOCOL = TCP)(HOST = pier-vip)(PORT = 39005))
)
(CONNECT_DATA =
(SERVER = DEDICATED)
(SERVICE_NAME = PPG3)
)
)
```

The following example shows how to set up the DecisionCenter data warehouse server JDBC thick driver Oracle Call Interface (OCI).

- 1 From Windows, set PATH to ORACLE_HOME\bin.
- 2 Using a text editor, open rds.properties in the data warehouse installation \conf directory.
- 3 Add the JDBC driver code:

```
jdbc.driver=oracle.jdbc.driver.OracleDriver
jdbc.url=jdbc:oracle:oci:@DC
```

Related topics

[Best Practices for Data Warehouse Administrators](#)

[ETL data flow](#)

[ETL validation tools](#)

[Best practices: Data warehouse administration](#)

[Data warehouse checklist](#)

Data warehouse architecture
Fact table ETL processes
Example: Data warehouse customization
Example: Index creation

Best practices: Data warehouse administration

The data warehouse is the foundation of HP DecisionCenter. You can improve data warehouse performance and increase efficiency by adhering to some administrative best practices.

See the *HP DecisionCenter Installation Guide* for detailed information about setting up your data warehouse, data sources, data targets, and mapping information.

Related topics

Best Practices for Data Warehouse Administrators

ETL data flow
ETL validation tools
JDBC connections
Data warehouse checklist
Data warehouse architecture
Fact table ETL processes
Example: Data warehouse customization
Example: Index creation

Add an index to ServiceCenter data
Administer databases
Avoid NULL keys
Configure scenario schedulers
Manage the schema
Monitor Business Objects reporting server
Monitor the ETL process
Review system tables

Add an index to ServiceCenter data

By adding an index to your HP ServiceCenter data that is stored in a Relational Database Management System (RDBMS), you can improve the data warehouse ETL process. The index expedites data retrieval and data manipulation. Modify the scripts based on the table names from the ServiceCenter source before you run the scripts in the ServiceCenter database.

For **Oracle**, run the following script:

```
DROP INDEX ACTKNOWNERRM1_CI1;  
CREATE INDEX ACTKNOWNERRM1_CI1 ON ACTKNOWNERRM1 (SYSMODTIME);  
  
DROP INDEX ACTIVITYM1_CI1;  
CREATE INDEX ACTIVITYM1_CI1 ON ACTIVITYM1 (SYSMODTIME);  
  
DROP INDEX ACTPROBM1_CI1;  
CREATE INDEX ACTPROBM1_CI1 ON ACTPROBM1 (SYSMODTIME);  
  
DROP INDEX CLOCKS1_CI1;  
CREATE INDEX CLOCKS1_CI1 ON CLOCKS1 (SYSMODTIME);
```

```

DROP INDEX CM3RM1_CI1;
CREATE INDEX CM3RM1_CI1 ON CM3RM1 (SYSMODTIME);

DROP INDEX CM3TM1_CI1;
CREATE INDEX CM3TM1_CI1 ON CM3TM1 (SYSMODTIME);

DROP INDEX INCIDENTSM1_CI1;
CREATE INDEX INCIDENTSM1_CI1 ON INCIDENTSM1 (SYSMODTIME);

DROP INDEX KNOWNERRORM1_CI1;
CREATE INDEX KNOWNERRORM1_CI1 ON KNOWNERRORM1 (SYSMODTIME);

DROP INDEX OCMQM1_CI1;
CREATE INDEX OCMQM1_CI1 ON OCMQM1 (SYSMODTIME);

DROP INDEX OCMOM1_CI1;
CREATE INDEX OCMOM1_CI1 ON OCMOM1 (SYSMODTIME);

DROP INDEX OCMLM1_CI1;
CREATE INDEX OCMLM1_CI1 ON OCMLM1 (SYSMODTIME);

DROP INDEX OUTAGEM1_CI1;
CREATE INDEX OUTAGEM1_CI1 ON OUTAGEM1 (SYSMODTIME);

DROP INDEX PROBSUMMARYM1_CI1;
CREATE INDEX PROBSUMMARYM1_CI1 ON PROBSUMMARYM1 (SYSMODTIME);

DROP INDEX ROOTCAUSEM1_CI1;
CREATE INDEX ROOTCAUSEM1_CI1 ON ROOTCAUSEM1 (SYSMODTIME);

DROP INDEX SCRELATIONM1_CI1;
CREATE INDEX SCRELATIONM1_CI1 ON SCRELATIONM1 (SYSMODTIME);

```

For **SQL Server**, run the following script:

```

IF EXISTS (SELECT name FROM sysindexes
           WHERE name = 'ACTKNOWNERRM1_CI1')
    DROP INDEX ACTKNOWNERRM1.ACTKNOWNERRM1_CI1

GO

CREATE INDEX ACTKNOWNERRM1_CI1 ON ACTKNOWNERRM1 (SYSMODTIME)
GO

IF EXISTS (SELECT name FROM sysindexes
           WHERE name = 'ACTIVITYM1_CI1')
    DROP INDEX ACTIVITYM1.ACTIVITYM1_CI1

GO

CREATE INDEX ACTIVITYM1_CI1 ON ACTIVITYM1 (SYSMODTIME)
GO

IF EXISTS (SELECT name FROM sysindexes
           WHERE name = 'ACTPROBM1_CI1')
    DROP INDEX ACTPROBM1.ACTPROBM1_CI1

GO

CREATE INDEX ACTPROBM1_CI1 ON ACTPROBM1 (SYSMODTIME)
GO

IF EXISTS (SELECT name FROM sysindexes
           WHERE name = 'CLOCKSM1_CI1')
    DROP INDEX CLOCKSM1.CLOCKSM1_CI1

GO

CREATE INDEX CLOCKSM1_CI1 ON CLOCKSM1 (SYSMODTIME)
GO

IF EXISTS (SELECT name FROM sysindexes
           WHERE name = 'CM3RM1_CI1')

```

```

        DROP INDEX CM3RM1.CM3RM1_CI1
GO
CREATE INDEX CM3RM1_CI1 ON CM3RM1 (SYSMODTIME)
GO
IF EXISTS (SELECT name FROM sysindexes
           WHERE name = 'CM3TM1_CI1')
        DROP INDEX CM3TM1.CM3TM1_CI1
GO
CREATE INDEX CM3TM1_CI1 ON CM3TM1 (SYSMODTIME)
GO
IF EXISTS (SELECT name FROM sysindexes
           WHERE name = 'INCIDENTSM1_CI1')
        DROP INDEX INCIDENTSM1.INCIDENTSM1_CI1
GO
CREATE INDEX INCIDENTSM1_CI1 ON INCIDENTSM1 (SYSMODTIME)
GO
IF EXISTS (SELECT name FROM sysindexes
           WHERE name = 'KNOWNERRORM1_CI1')
        DROP INDEX KNOWNERRORM1.KNOWNERRORM1_CI1
GO
CREATE INDEX KNOWNERRORM1_CI1 ON KNOWNERRORM1 (SYSMODTIME)
GO
IF EXISTS (SELECT name FROM sysindexes
           WHERE name = 'OCMQM1_CI1')
        DROP INDEX OCMQM1.OCMQM1_CI1
GO
CREATE INDEX OCMQM1_CI1 ON OCMQM1 (SYSMODTIME)
GO
IF EXISTS (SELECT name FROM sysindexes
           WHERE name = 'OCMOM1_CI1')
        DROP INDEX OCMOM1.OCMOM1_CI1
GO
CREATE INDEX OCMOM1_CI1 ON OCMOM1 (SYSMODTIME)
GO
IF EXISTS (SELECT name FROM sysindexes
           WHERE name = 'OCMLM1_CI1')
        DROP INDEX OCMLM1.OCMLM1_CI1
GO
CREATE INDEX OCMLM1_CI1 ON OCMLM1 (SYSMODTIME)
GO
IF EXISTS (SELECT name FROM sysindexes
           WHERE name = 'OUTAGEM1_CI1')
        DROP INDEX OUTAGEM1.OUTAGEM1_CI1
GO
CREATE INDEX OUTAGEM1_CI1 ON OUTAGEM1 (SYSMODTIME)
GO
IF EXISTS (SELECT name FROM sysindexes
           WHERE name = 'PROBSUMMARYM1_CI1')
        DROP INDEX PROBSUMMARYM1.PROBSUMMARYM1_CI1
GO
CREATE INDEX PROBSUMMARYM1_CI1 ON PROBSUMMARYM1 (SYSMODTIME)
GO

```

```

IF EXISTS (SELECT name FROM sysindexes
           WHERE name = 'ROOTCAUSEM1_CI1')
           DROP INDEX ROOTCAUSEM1.ROOTCAUSEM1_CI1

GO

CREATE INDEX ROOTCAUSEM1_CI1 ON ROOTCAUSEM1 (SYSMODTIME)
GO

IF EXISTS (SELECT name FROM sysindexes
           WHERE name = 'SCRELATIONM1_CI1')
           DROP INDEX SCRELATIONM1.SCRELATIONM1_CI1

GO

CREATE INDEX SCRELATIONM1_CI1 ON SCRELATIONM1 (SYSMODTIME)
GO

```

Related topics

Best practices: Data warehouse administration

Administer databases

Avoid NULL keys

Configure scenario schedulers

Manage the schema

Monitor Business Objects reporting server

Monitor the ETL process

Review system tables

Administer databases

HP DecisionCenter interacts with several components:

- Data warehouse database
- Business Objects repositories
- HP Connect-It database

The following tasks can help you maintain the integrity of your data and improve analytical process performance.

- Back up all databases on a regular basis.
- Monitor the Extract, Transform, and Load (ETL) process and reporting activity.
- For the data warehouse database, create additional indexes for:
 - Frequently-used ad hoc attributes
 - **sysmodtime** for all ServiceCenter data source tables
- Monitor database tablespaces and storage to ensure adequate disk space. Increase the tablespaces as needed.

Optionally, archive old data and create partition conditions to access the data. To ease a shortage of disk space, you can archive large tables such as `probsum_d` (`probsum_f`), `clock_d` (`clock_f`), and `outage_d` (`outage_f`) to different partitions and build a data connection to old data ad hoc query activities.

The partition condition for a dimension table is:

```
Z_RDSTRANSLASTIND = 'N'
```

or

```
Z_RDSACTIVESTATUSIND = 'N'
```

The partition condition for a fact table is:

```
Z_RDSOBSOLETEIND = 'Y'
```

Related topics

Best practices: Data warehouse administration

Add an index to ServiceCenter data

Avoid NULL keys

Configure scenario schedulers

Manage the schema

Monitor Business Objects reporting server

Monitor the ETL process

Review system tables

Avoid NULL keys

The presence of NULL in a key field can negatively affect the efficiency and accuracy of your ETL process. Most keys are required fields. When you have multi-field unique keys, it is not uncommon for one of the set to have a NULL value.

You can avoid NULL keys in multiple levels that include ServiceCenter and Connect-It.

- Use form validations in your source HP ServiceCenter and HP AssetCenter applications to enforce the entry of at least a default value for all key fields. See your application documentation for details about enforcing data integrity.
- In HP Connect-It, include key fields in a function that replaces NULL values with a default value. Most keys in the DecisionCenter data warehouse are character fields. The `CSTR()` function replaces NULL with a zero-length string. For example:

```
CSTR([location])
```

- In your source application, add a new key field to the table that the server automatically generates. Make sure that it cannot be edited by the end user. Designate this field as unique and make sure it appears first in the list of keys for the table. See your application documentation for details about editing keys.

Note: This action should only be performed by a knowledgeable application administrator.

Related topics

Best practices: Data warehouse administration

Add an index to ServiceCenter data

Administer databases

Configure scenario schedulers

Manage the schema

Monitor Business Objects reporting server

Monitor the ETL process

Review system tables

Configure scenario schedulers

The default data warehouse synchronization runs daily. Only one scheduler is defined for all the data synchronization updates. To optimize the synchronization, add different interval settings based on your report refresh requirements.

You can set operation data in DecisionCenter Analytics to refresh multiple times daily. Schedule update synchronization intervals no more than every two hours to avoid refreshing the data too frequently. Refreshing too often may lead to an unstable data warehouse environment.

- 1 From the HP Connect-It service console, modify the RDS_ALL scheduler to set multiple update synchronization intervals, as necessary.
- 2 See the *HP Connect-It User's Guide* for more information about setting times by Itemized list or periodical.
- 3 Verify that the scheduled mappings for preMapping, produced documents, and postMapping are in the right order.
- 4 Monitor the data warehouse Connect-It service schedule pointer updates in the rds_sc.ini file of the `cit` directory in the data warehouse installation:
\\...\\HP\\DecisionCenter 2.00\\ITPA-SM. If you need to synchronize any tables fully, remove the line of the source table. The entire table fully synchronizes the next time that the update process runs.

Related topics

Best practices: Data warehouse administration

Add an index to ServiceCenter data

Administer databases

Avoid NULL keys

Manage the schema

Monitor Business Objects reporting server

Monitor the ETL process

Review system tables

Manage the schema

The data warehouse database contains the schema based on the HP ServiceCenter or HP AssetCenter data source. The schema can include new tables, new fields, and changes to existing fields in your data. You must add the changes to the schema if you want the data warehouse database to include the changes in your source data.

The `rds_etl.xml` file is in the `conf` directory of the data warehouse installation:
\\...\\HP\\DecisionCenter 2.00\\ITPA-SM.

- Track reporting and analytical requirements.
- Track data source schema changes.
- Plan data warehouse schema modifications or extension implementation.
- Make schema changes such as adding dimension columns, changing dimension columns, and changing fact key columns.

Related topics

Best practices: Data warehouse administration

Add an index to ServiceCenter data

Administer databases

Avoid NULL keys
Configure scenario schedulers
Monitor Business Objects reporting server
Monitor the ETL process
Review system tables

Monitor Business Objects reporting server

The Business Objects reporting server works with DecisionCenter to support reporting, ad hoc queries, and analysis of your IT performance data. See the Business Objects documentation for detailed information about Business Objects administration processes.

- Monitor the Business Objects reporting server memory and CPU usage to ensure adequate memory and disk space. Add vertical or horizontal server processes as needed.
- Monitor the peak concurrent sessions to ensure that you have enough concurrent licenses.
- Monitor the scheduled job completion to ensure that your jobs ran successfully.
- Set proper security profile and access rights in the Central Management Console (CMC) for DecisionCenter reporting groups and users. The CMC is responsible for the Business Objects Enterprise system that includes user management, server management, and server groups.

Related topics

Best practices: Data warehouse administration

Add an index to ServiceCenter data

Administer databases

Avoid NULL keys

Configure scenario schedulers

Manage the schema

Monitor the ETL process

Review system tables

Monitor the ETL process

The Extract, Transform, and Load (ETL) process is how data gets loaded into the data warehouse. You can ensure data integrity by regularly verifying the information in the system tables and system logs. The log files are in the **logs** directory of the data warehouse installation: \\... \ITPA-SM.

Verification task	Action
Query the data warehouse system table information.	Run the query for the RDS_ETLSYNC_LOG table to ensure that at least one record is created for each update synchronization. For example: <pre>select * from RDS_ETLSYNC_LOG where TABLE_NAME = 'RDS_LAST_SYNC' order by SYNC_DATETIME desc</pre>

Verification task	Action
Monitor ETL log files.	<p>The ETL produces log files that you can access to troubleshoot errors. The log files are in the data warehouse installation <code>logs</code> directory:</p> <pre> etl_citTimer.log etl_postmapping.log rds_etl_agg.log rds_etl_hier.log rds_etl_runscript.log rds_etl_sync.log *.err rds_sc.log rds_validator.log </pre> <ul style="list-style-type: none"> • Review the log files daily. • Review *.err files that are not empty. • After you correct the errors that appear in the *.err files, delete or zip the files. • Check the timestamp of the files to determine if new errors have occurred. • To view additional data warehouse synchronization process logging information, you can use the debug function in the <code>rds.properties</code> file located in <code>\\...\ITPA-SM\conf\</code>. <pre>bidebug='true'</pre>

Related topics

[Best practices: Data warehouse administration](#)

[Add an index to ServiceCenter data](#)

[Administer databases](#)

[Avoid NULL keys](#)

[Configure scenario schedulers](#)

[Manage the schema](#)

[Monitor Business Objects reporting server](#)

[Review system tables](#)

Review system tables

The system tables in the data warehouse provide information about schema metadata, the data warehouse environment, Extract, Transform, and Load (ETL) processing, and time dimension reporting. By reviewing the content of the tables daily, you can:

- Track unique columns and fields.
- Identify Slowly Changing Dimension (SCD) columns for dimension tables.
- View information about the data warehouse deployment environment.
- Verify the ETL process status.
- Find the available sequence value for unique system keys.
- Check the ETL synchronization volume.
- Examine time dimension information.

Table purpose	Table name
Track schema metadata	RDS_SCDCOLUMN RDS_SEQUENCE RDS_UNIQUECOLUMN
Store environment data	RDS_DBINFO
Track ETL processing	RDS_CIT_LOG RDS_CITSCHPTR_LOG RDS_CITSYNC_LOG RDS_ETLSYNC_LOG RDS_LOGINFO
Provide time measurements	RDS_TIMEDIM_D

One of the most useful tables for data warehouse administrators is the RDS_ETLSYNC_LOG table, which tracks the internal ETL processing status and scans the synchronization for dimension tables.

See **Related topics** for more detailed information about the individual system tables.

Related topics

[Best practices: Data warehouse administration](#)

[Add an index to ServiceCenter data](#)

[Administer databases](#)

[Avoid NULL keys](#)

[Configure scenario schedulers](#)

[Manage the schema](#)

[Monitor Business Objects reporting server](#)

[Monitor the ETL process](#)

[RDS_CIT_LOG table](#)

[RDS_CITSCHPTR_LOG table](#)

[RDS_CITSYNC_LOG table](#)

[RDS_DBINFO table](#)

[RDS_ETLSYNC_LOG table](#)

[RDS_LOGINFO table](#)

[RDS_SCDCOLUMN table](#)

[RDS_SEQUENCE table](#)

[RDS_TIMEDIM_D table](#)

[RDS_UNIQUECOLUMN table](#)

RDS_CIT_LOG table

The RDS_CIT_LOG table tracks the internal data warehouse Extract, Transform, and Load (ETL) processing status.

Name	Null?	Type	Description
ID	Not NULL	NUMBER(10)	Unique system key

Name	Null?	Type	Description
STATE_FLAG		VARCHAR2(60)	ETL process status
CIT_TIME		DATE	ETL process time

The STATE_FLAG values show the ETL process status.

Value	Description
CIT_START	Data warehouse HP Connect-It scenario initial or periodic synchronization starts
CIT_FAILED	Connect-It scenario initial or periodic synchronization fails
RDS_SYNC_FACT_START	Data warehouse fact tables population starts
RDS_SYNC_ASSOCIATE_START	Data warehouse associate tables population starts
RDS_SYNC_DELTRAN_START	Data warehouse deletion synchronization starts
RDS_SYNC_SYSTEM_START	Data warehouse system tables population starts
RDS_AGG_START	Data warehouse aggregation tables population starts
RDS_HIER_START	Data warehouse hierarchy tables population starts
COMPLETED	ETL process is completed

Related topics

[RDS_CITSCHPTR_LOG table](#)
[RDS_CITSYNC_LOG table](#)
[RDS_DBINFO table](#)
[RDS_ETLSYNC_LOG table](#)
[RDS_LOGINFO table](#)
[RDS_SCDCOLUMN table](#)
[RDS_SEQUENCE table](#)
[RDS_TIMEDIM_D table](#)
[RDS_UNIQUECOLUMN table](#)

RDS_CITSCHPTR_LOG table

The RDS_CITSCHPTR_LOG table tracks the Connect-It schedule pointer records, which enable you to recover only those records that were modified or created since the last scenario startup.

Name	Null?	Type	Description
RDS_CITSCHEDLOG_ID	Not NULL	NUMBER	Unique ID
TABLE_NAME	Not NULL	VARCHAR(80)	Source table name
LASTSYNC_DATE		DATE	Last synchronization time
SCENARIO_NAME	Not NULL	VARCHAR(30)	Connect-It scenario name
Z_RDSCREATEDDATE		DATE	Date record was created
Z_RDSMODIFIEDDATE		DATE	Date record was modified

Related topics

[RDS_CIT_LOG table](#)
[RDS_CITSYNC_LOG table](#)
[RDS_DBINFO table](#)
[RDS_ETLSYNC_LOG table](#)
[RDS_LOGININFO table](#)
[RDS_SCDCOLUMN table](#)
[RDS_SEQUENCE table](#)
[RDS_TIMEDIM_D table](#)
[RDS_UNIQUECOLUMN table](#)

RDS_CITSYNC_LOG table

The RDS_CITSCHPTR_LOG table tracks records mapped in source and target connectors.

Name	Null?	Type	Description
RDS_CITSYNCCLOG_ID	Not NULL	NUMBER	Unique ID
TABLE_NAME	Not NULL	VARCHAR(80)	Source / target table name
CONNECTOR_NAME	Not NULL	VARCHAR(80)	Connect-It connector name
SYNC_RECORD_COUNT	Not NULL	NUMBER	Number of synchronized records

Name	Null?	Type	Description
DATE_LOGGED		DATE	Date logged in CITLOG table in Connect-It monitoring database
ETYPE	Not NULL	NUMBER	Type of log in CITLOG table in Connect-It monitoring database
LLOGID	Not NULL	NUMBER	Unique ID in CITLOG table
LSESSIONID	Not NULL	NUMBER	Unique Connect-It session ID
SCENARIO_NAME	Not NULL	VARCHAR(30)	Connect-It scenario name
Z_RDSCREATEDDATE		DATE	Date record was created

Related topics

[RDS_CIT_LOG table](#)
[RDS_CITSCHPTR_LOG table](#)
[RDS_DBINFO table](#)
[RDS_ETLSYNC_LOG table](#)
[RDS_LOGININFO table](#)
[RDS_SCDCOLUMN table](#)
[RDS_SEQUENCE table](#)
[RDS_TIMEDIM_D table](#)
[RDS_UNIQUECOLUMN table](#)

RDS_DBINFO table

The RDS_DBINFO table stores information about the data warehouse deployment environment.

Name	Null?	Type	Description
Z_RDS_DBINFO_ID	Not NULL	NUMBER(10)	Unique system key
OSTYPE		VARCHAR2(20)	Operating systems type
DATABASETYPE		VARCHAR2(20)	Database type
DATABASENAME		VARCHAR2(20)	Database name

Name	Null?	Type	Description
DATABASEURL		VARCHAR2(100)	JDBC URL to connect
USERID		VARCHAR2(20)	User name to access the database
USERPASSWORD		VARCHAR2(20)	User password
PROG_PATH		VARCHAR2(100)	Data warehouse installation path
PROG_NAME		VARCHAR2(100)	Data warehouse program name

Related topics

[RDS_CIT_LOG table](#)
[RDS_CITSCHPTR_LOG table](#)
[RDS_CITSYNC_LOG table](#)
[RDS_ETLSYNC_LOG table](#)
[RDS_LOGININFO table](#)
[RDS_SCDCOLUMN table](#)
[RDS_SEQUENCE table](#)
[RDS_TIMEDIM_D table](#)
[RDS_UNIQUECOLUMN table](#)

RDS_ETLSYNC_LOG table

The RDS_ETLSYNC_LOG table provides information about the Extract, Transform, and Load (ETL) synchronization volume.

The following example contains RDS_ETLSYNC_LOG information.

Name	Null?	Type	Description
RDS_ETLSYNC_LOG_ID	Not NULL	NUMBER	Unique system key
TABLE_NAME	Not NULL	VARCHAR2(80)	Populated table name
TABLE_ACTION	Not NULL	VARCHAR2(80)	ETL process updating action codes: Add Update Update with SCD Delete Log
SYNC_RECORD_COUNT	Not	NUMBER	Total record number of the updating

Name	Null?	Type	Description
	NULL		described by the action code
SYNC_DATETIME		DATE	Sync date time
Z_RDSCREATEDDATE		DATE	Track record creation date time

Related topics

[RDS_CIT_LOG table](#)
[RDS_CITSCHPTR_LOG table](#)
[RDS_CITSYNC_LOG table](#)
[RDS_DBINFO table](#)
[RDS_LOGININFO table](#)
[RDS_SCDCOLUMN table](#)
[RDS_SEQUENCE table](#)
[RDS_TIMEDIM_D table](#)
[RDS_UNIQUECOLUMN table](#)

RDS_LOGININFO table

The RDS_LOGININFO table tracks the status of the Extract, Transform, and Load process.

Name	Null?	Type	Description
RDS_LOGININFO_ID	Not NULL	NUMBER	Unique ID
CHKPOINT_NAME	Not NULL	VARCHAR(30)	Name of validating tool
STATUS	Not NULL	VARCHAR(80)	Process status
ERRORCODE		NUMBER	Error code from CITLOG table in Connect-It monitoring database
ERRORDESC		VARCHAR(255)	Error description
CONTEXT		VARCHAR(255)	Message context from CITLOG table in Connect-It monitoring database
LSESSIONID		NUMBER	Unique Connect-It session ID

Name	Null?	Type	Description
SCENARIO_NAME	Not NULL	VARCHAR(30)	Connect-It scenario name
Z_RDSCREATEDDATE		DATE	Date record was created

Related topics

[RDS_CIT_LOG table](#)
[RDS_CITSCHPTR_LOG table](#)
[RDS_CITSYNC_LOG table](#)
[RDS_DBINFO table](#)
[RDS_ETLSYNC_LOG table](#)
[RDS_SCDCOLUMN table](#)
[RDS_SEQUENCE table](#)
[RDS_TIMEDIM_D table](#)
[RDS_UNIQUECOLUMN table](#)

RDS_SCDCOLUMN table

The RDS_SCDCOLUMN table tracks the Slowly Changing Dimension (SCD) columns for dimension tables. The data warehouse HP Connect-It scenario uses this information to update the information with a simple overwrite procedure or the SCD Type Two update in the Extract, Transform, and Load (ETL) process. The SCD Type Two updating process creates a new record and marks the existing record as *old* in dimension tables.

The following example contains RDS_SCDCOLUMN table columns in an Oracle database.

Name	Null?	Type	Description
LRDS_SCDCOLUMN_ID	Not NULL	NUMBER(10)	Unique system key
TABLE_NAME	Not NULL	VARCHAR2(80)	The name of the table as it appears in the data warehouse database, without any suffixes
SRC_SCDCOLUMN	Not NULL	VARCHAR2(80)	The name of the column as it appears in the original data source. If this contains a period, it must have single quotation marks.
RDS_SCDCOLUMN	Not NULL	VARCHAR2(80)	The name of the column as it appears in the data warehouse
TYPE_SCDCOLUMN	Not NULL	VARCHAR2(30)	The data type of the column in the original source data
TABLE_SCDNUMBER	Not NULL	NUMBER(10)	The sequence number of the SCD column if more than one

Related topics

[RDS_CIT_LOG table](#)

[RDS_CITSCHPTR_LOG table](#)

[RDS_CITSYNC_LOG table](#)

[RDS_DBINFO table](#)

[RDS_ETLSYNC_LOG table](#)

[RDS_LOGINFO table](#)

[RDS_SEQUENCE table](#)

[RDS_TIMEDIM_D table](#)

[RDS_UNIQUECOLUMN table](#)

RDS_SEQUENCE table

The RDS_SEQUENCE table stores the next available sequence value for the unique system keys for all data warehouse tables.

Name	Null?	Type	Description
LRDS_SEQUENCE_ID	Not NULL	NUMBER(10)	Unique system key
TABLE_NAME	Not NULL	VARCHAR2(80)	Data warehouse table name
TABLE_SEQUENCE	Not NULL	NUMBER(0)	Sequence value

Related topics

[RDS_CIT_LOG table](#)
[RDS_CITSCHPTR_LOG table](#)
[RDS_CITSYNC_LOG table](#)
[RDS_DBINFO table](#)
[RDS_ETLSYNC_LOG table](#)
[RDS_LOGININFO table](#)
[RDS_SCDCOLUMN table](#)
[RDS_TIMEDIM_D table](#)
[RDS_UNIQUECOLUMN table](#)

RDS_TIMEDIM_D table

The data warehouse time dimension table provides the mechanism for DecisionCenter reporting based on days, months, quarters and years.

All data warehouse aggregation tables end with _AGG. Table name conventions pertain to weeks, quarters and months, and years. RDS_TIMEDIM_D stores the calendar days from 1995 to 2010. The data warehouse installer sets the start-year and end-year values, and saves them in the rds.properties file of the data warehouse installation `conf` directory: \\...\HP\DecisionCenter 2.00\ITPA-SM.

```
rds.startDate=1995  
rds.endDate=2010
```

The following example contains time dimension table columns information.

Name	Null?	Type
RDS_TIMEDIM_DID	Not NULL	NUMBER
FULLDATE		VARCHAR2(20)
WEEK		VARCHAR2(10)

Name	Null?	Type
MONTH		VARCHAR2(20)
QUARTER		VARACHAR2(4)
YEAR		VARCHARS2(10)
FISCALPERIOD		VARCHAR2(4)
LASTDAYINMONTHFLAG		CHAR(1)
FULLDATE_D		DATE
MONTH_NUMERIC		NUMBER
QUALIFIED_QUARTER		VARCHAR(15)
QUALIFIED_MONTH		VARCHAR(15)
QUALIFIED_WEEK		VARCHAR(15)
FISCALYEAR		VARCHAR(10)

Related topics

[RDS_CIT_LOG table](#)
[RDS_CITSCHPTR_LOG table](#)
[RDS_CITSYNC_LOG table](#)
[RDS_DBINFO table](#)
[RDS_ETLSYNC_LOG table](#)
[RDS_LOGININFO table](#)
[RDS_SCDCOLUMN table](#)
[RDS_SEQUENCE table](#)
[RDS_UNIQUECOLUMN table](#)

RDS_UNIQUECOLUMN table

The RDS_UNIQUECOLUMN table tracks the unique columns for the reporting tables. It also includes unique fields for the ServiceCenter and AssetCenter source files. The data warehouse schema creation program uses this information to create unique indexes for data warehouse reporting tables. The HP Connect-It scenarios use the unique field values to add new or update existing records in the Extract, Transform, and Load process.

The following example contains RDS_UNIQUECOLUMN table columns in an Oracle database.

Name	Null?	Type	Description
LRDS_UNIQUECOLUMN_ID	Not NULL	NUMBER(10)	Unique system key
TABLE_NAME	Not NULL	VARCHAR2(80)	The name of the table as it appears in the data warehouse database, without any suffixes
TABLE_TYPE		VARCHAR2(80)	Data warehouse table type as one of following values: DIMENSION DIRECT_MAPPING
DS_TABLE_NAME	Not NULL	VARCHAR2(30)	The name of the table or file as it appears in the original ServiceCenter or AssetCenter data source
DS_TABLE_TYPE	Not NULL	VARCHAR2(30)	File type has the following values: DATA_SOURCE (as a source file) PARENT_DIM Note: Use PARENT_DIM information for cascading delete. When the value of DS_TABLE_NAME is deleted from the data source, the corresponding record in the TABLE_NAME table is marked as inactive.
SRC_UNIQUECOLUMN	Not NULL	VARCHAR2(80)	The name of the column as it appears in the original data source. If this contains a period, it must have single quotation marks.
RDS_UNIQUECOLUMN	Not NULL	VARCHAR2(80)	The name of the column as it appears in the data warehouse
TYPE_UNIQUECOLUMN		VARCHAR2(30)	The data type of the column in the original data source
TABLE_UNIQUENUMBER	Not NULL	VARCHAR2(10)	Index of the field in the unique key. Note: If the table uses a single field key, it is always 1. If the key contains multiple fields, this lists the field order in the key.

In the following example, the CLOCK table uses four fields to build its key.

DS_TABLE_NAME	SRC_UNIQUECOLUMN	RDS_UNIQUECOLUMN	TABLE_UNIQUENUMBER
clocks	TYPE	TYPEPRGN	1
clocks	'KEY.CHAR'	KEY_CHAR	2
clocks	'KEY.NUMERIC'	KEY_NUMERIC_KEY	3
clocks	NAME	NAME	4

Related topics

- [RDS_CIT_LOG table](#)
- [RDS_CITSCHPTR_LOG table](#)
- [RDS_CITSYNC_LOG table](#)
- [RDS_DBINFO table](#)
- [RDS_ETLSYNC_LOG table](#)
- [RDS_LOGININFO table](#)
- [RDS_SCDCOLUMN table](#)
- [RDS_SEQUENCE table](#)
- [RDS_TIMEDIM_D table](#)

Data warehouse checklist

By performing routine daily tasks, you can maximize the data warehouse performance. Make sure that you regularly monitor the data warehouse ETL server, perform backup and maintenance on all databases, and monitor the Business Objects server.

Monitor the data warehouse ETL server	Maintain all databases	Monitor the Business Objects server
<ul style="list-style-type: none"> • Check the data warehouse Connect-It service schedule pointers updates. • Query the data warehouse system tables information. • Check DecisionCenter log files and clean the rds\logs directory as needed. • Check the Connect-It log files. • Check the data warehouse data synchronization process. 	<ul style="list-style-type: none"> • Regularly back up the data warehouse database, Business Objects repositories, and Connect-It Admin database. • Check database tablespaces and storages. • Optionally, clean up old data. 	<ul style="list-style-type: none"> • Check memory and CPU usage. • Check peak concurrent sessions. • Check scheduled job completion. • Check personal storage free disk space.

Print this checklist to help you maintain the DecisionCenter data warehouse. Gathering this information may require the combined expertise of a ServiceCenter, Service Manager or AssetCenter Administrator, Database Administrator, and

Professional Services personnel to arrive at the decisions and complete the required actions. As you complete each question, mark it as finished. The answers to these questions will help you set up and monitor the data warehouse.

General environment

Print this checklist to help you maintain the DecisionCenter data warehouse. Gather the information about the RDBMS, ServiceCenter, Service Manager, or AssetCenter servers in your environment. As you answer each question, mark it completed.

Check when completed	Question
	How many servers do you use for DecisionCenter components in the production environment?
	What version of the components do you have in the production environment? <ul style="list-style-type: none"> • Data source database: ServiceCenter, Service Manager, AssetCenter • Data source RDBMS database: Oracle, SQL Server • Data warehouse RDBMS database • HP Connect-It • Operating system • Application server
	Are the servers in the same domain and time zone?
	Is the data source database server the same as the data warehouse database server?
	What backup procedures do you use for the databases? Do you use DecisionCenter backup server files such as the Connect-It scheduler pointer file (\\...\rds_sc\cit\rds_sc.ini) and log files (\\...\rds_sc\logs)? What is the backup frequency of databases and files?
	What is the average update volume for extracted data source tables?
	How many production and test environments do you have?
	How many monitoring processes are enabled for Connect-It scenarios?

Database issues

Print this checklist to help you maintain the DecisionCenter data warehouse. You may require the expertise of a Database Administrator to gather the information about the RDBMS. These questions are for an Oracle environment. As you answer each question, mark it completed.

Check when completed	Question
	If Oracle database failover or switch overs exist, do you use Real Application Clusters (RAC) or Data Guide?
	Are data warehouse and data source databases involved in the failover or switch over?
	If you use Data Guide, does the data warehouse schema guide it?
	How are DecisionCenter database ODBC and JDBC connections reconfigured for the failover or switch over?
	For RAC failover ODBC implementation: <ul style="list-style-type: none">• Is the Oracle Call Interface (OCI) connector with tnsname lookup configured correctly for Oracle Fast Connection Failover (FCF)?• Do you have the RAC multiple instance with a single database or multiple databases?
	What type of failover do you reference, instance or database?
	If an instance failover occurs, was the primary instance completely unavailable?
	Did the DecisionCenter Extract, Transform, and Load (ETL) processes stop for the switch over or failover?
	Have any failover situations previously occurred? If so, <ul style="list-style-type: none">• What happened with the DecisionCenter ETL processes?• What Oracle error messages were in the DecisionCenter log files?

Data errors

Print this checklist to help you maintain the DecisionCenter data warehouse. You may require the expertise of a ServiceCenter, Service Manager, or AssetCenter Administrator to gather the information about the data. As you answer each question, mark it completed.

Check when completed	Question
	What information can you provide about the missing data in the data warehouse?
	When did the ETL process that caused the data corruption first fail? What were the causes?
	Other than the mass delete event situation, do you have system problems with the Connect-It ServiceCenter or AssetCenter connector sessions?
	How frequently do you experience lost database connections?
	What are the circumstances when scenario services previously stopped or had interruptions?
	What, if any, ETL issues did you have when you changed to Daylight Saving time?
	What ETL auditing reports do you generate?

Troubleshooting hints

Print this checklist to help you maintain the DecisionCenter data warehouse. This section provides some basic information for anticipating and correcting issues related to the data warehouse. Before you attempt to troubleshoot a problem, ensure that you have DecisionCenter system files readily available. As you read each issue, mark it completed.

Check when completed	Issue
	<p>Do you have access to the DecisionCenter system files? When you call Support, you must have these files.</p> <ul style="list-style-type: none">• rds_sc.scn or rds_ac.scn• rds.bas• conitgui.log• rds_sc.ini or rds_ac.ini and all back up files• All *.log and *.err files in the \\...\rds_sc\logs or \\...\rds_ac\logs directory, provided the Connect-It scenario trace settings are on• rds_etl.xml or rdsac_etl.xml• rds.properties
	<p>Do you have necessary access rights for all supporting servers, including the database server?</p>
	<p>Can you change the Connect-It scenario to add auditing programs to parse mapping process logs that determine a successful run or failed case for ETL process administration notifications?</p>
	<p>Can you replace data warehouse java executables to add necessary trace and log statements?</p>
	<p>Can you create a new Connect-It administration database for scenario auditing and monitoring processes? The administration database can create a new schema and tables under that schema. Samples of Connect-It tables are available in citsession and citlog.</p>
	<p>A known error with a workaround can correct memory leak issues in ServiceCenter 6.1.3. To repair the memory leak, upgrade to ServiceCenter 6.1.4 or to Connect-It 3.6.1 6186.</p>

Related topics

[Best Practices for Data Warehouse Administrators](#)

[ETL data flow](#)

[ETL validation tools](#)

[JDBC connections](#)

[Best practices: Data warehouse administration](#)

Data warehouse architecture
Fact table ETL processes
Example: Data warehouse customization
Example: Index creation

Data warehouse architecture

The data warehouse is a repository of integrated information that is available for queries and analysis and includes Extract, Transform, and Load (ETL) processes and metadata.

Some advantages of the data warehouse include:

- Reduced reporting pressure on the database
- Ability for administrators to restructure data to speed up data analysis and reporting capabilities
- Ease of use in generating new reports
- Ability for administrators to create clean data that does not require wholesale changes to the ServiceCenter or AssetCenter transactional system or business processes
- Ability to use hierarchy analysis
- Ability to use aggregate tables for better performance
- Availability of historical analysis by providing a data source that supports a longer span of time

Related topics

Best Practices for Data Warehouse Administrators

ETL data flow

ETL validation tools

JDBC connections

Best practices: Data warehouse administration

Data warehouse checklist

Fact table ETL processes

Example: Data warehouse customization

Example: Index creation

Design

Naming conventions

Dimension table system keys

Design

The data warehouse uses dimensional modeling to present data in a standard framework that is intuitive and allows for high performance access. It is extendable and provides easy management of aggregates.

The data warehouse defines tables based on the following categories.

Table category	Description
Dimension	<p>HP Connect-It maps most dimension table fields directly from ServiceCenter and AssetCenter tables. Dimension tables add additional data warehouse Extract, Transform, and Load (ETL) attributes.</p> <p>Tables that include historical or time-associated data can be dimension tables. You can use dimension tables as search criteria for multiple ServiceCenter and AssetCenter reporting modules.</p> <p>Each dimension table has the Z_RDS(tablename)ID system key. The surrogate keys allow the data warehouse to:</p> <ul style="list-style-type: none">• Assign a new key version for Slowly Changing Dimensions.• Encode uncertain, not known, not recorded, and null record types. <p>To adhere to best practices, base every join between dimension and fact tables on surrogate keys.</p>
Fact	<p>Fact tables are the data warehouse tables created for measurements for the associated dimension tables that the reporting modules use.</p>
Associate	<p>You can use associate tables to bridge dimension tables in the data warehouse. They can normalize many-to-many relationships and prevent data inaccuracies caused by ambiguous joins.</p>
Direct mapping	<p>You can use ServiceCenter and AssetCenter transaction tables for detail reporting and as search criteria. These tables directly map from ServiceCenter and AssetCenter into the data warehouse.</p>
Aggregate	<p>You can summarize your data more quickly by grouping the measures on common sort fields.</p>
Hierarchy	<p>In ServiceCenter, hierarchy tables contain fields to create parent-child relationships between records.</p> <p>AssetCenter builds hierarchy tables from the fullpath field in structured data.</p> <p>The maximum hierarchy level is 10.</p>

Table category	Description
System	System tables track the data warehouse schema metadata, user security, ETL process status, and time dimension. All system table names, except BIRECORDDELETE, begin with RDS_.

Related topics

[Data warehouse architecture](#)

[Naming conventions](#)

[Dimension table system keys](#)

Naming conventions

Data warehouse tables use consistent naming conventions.

- Dimension table names end with _D and have a maximum length of 18 characters. The primary key for a dimension table is Z_RDSXXX_DID.
- Dimension table names use the first 15 characters from the original source table names without the S.
- Fact table names end with _F.
- Measurement names for fact tables do not use the count character.
- Aggregate table names end with _AGG.
- Associate or bridge table names use the first eight characters from the two associated table names with an underscore (_) as the separator.
- Direct mapping table names use the same name as the source name.
- The underscore (_) character replaces the dot (.) character when converting ServiceCenter and AssetCenter field names to data warehouse field names. For example, parent.name becomes parent_name.
- The maximum Direct Mapping Indexes length is 18 characters.
- Data warehouse system table names begin with RDS_.

Related topics

[Data warehouse architecture](#)

[Design](#)

[Dimension table system keys](#)

Dimension table system keys

Dimension table attributes provide information to help track changes to data warehouse table reporting.

Action	Result
New records are added to data warehouse dimension tables (XXX_D)	Z_RDSCREATEDDATE is set to the current time
Records are updated to data warehouse dimension tables (XXX_D)	Z_RDSLASTMODDATE is set to the current time
Records are deleted from ServiceCenter	Dimension table attributes are updated: Z_RDSACTIVESTATUSIND = 'N' Z_RDSDELETEDDATE is set to the deleted synchronization time Z_RDSLASTMODDATE is set to the current time
The records are updated as slowly changing dimension (SCD)	<ul style="list-style-type: none">The original dimension table record attributes are updated: Z_RDSTRANSLASTIND='N' Z_RDSTRANSENDDATE is set to the current timeThe new record is added as ADD caseZ_RDSTRANSLASTIND = 'N' or Z_RDSACTIVESTATUSIND = 'N' means the dimension record is no longer active

Related topics

[Data warehouse architecture](#)

[Design](#)

[Naming conventions](#)

Fact table ETL processes

Data warehouse fact tables update frequently because they accumulate snapshots with measurements, date fields, and status changes. The dimensional foreign keys and measurements change when the information changes.

The creation date time of a fact record is the same as the main dimension table, which stores the context of recorded line items, such as incidents or invoices. The creation date time field in the fact table is the last modified column to track when the fact record was last updated.

Fact table ETL processes parallel the data warehouse dimensional modeling techniques. Fact tables show changes to the main dimension table, referenced dimension tables, and universe filter conditions.

Main dimension table changes

The ETL process causes changes to the main dimension table.

- A new line item, such as incident or portfolio, is created, adding a new fact record.
- The existing line item is updated. A matching fact record is updated only when the dimensional foreign keys or measurements are changed.
- The existing line item slowly changing dimension (SCD) is updated. The matching fact record main foreign key is updated from the old key to the new SCD3 key.
- The existing line item is marked as deleted. The matching fact record is marked *obsolete*.

Referenced dimension table changes

The ETL process causes changes to the referenced dimension tables.

- A new record is created in the referenced dimension tables without changing the fact tables.
- An existing referenced dimensional record is updated without changing the fact tables.
- An existing referenced dimensional SCD record is updated. The related dimensional foreign key must be updated with the new surrogate key value in all the related fact records.
- An existing referenced dimensional record is deleted without changing the fact tables.

Universe filter condition changes

The ETL process causes changes when universe filter conditions for query or reporting based on the fact table and dimensional tables have updates.

- The fact table is not *obsolete*.
- Dimensional foreign keys are greater than 5 (≥ 5).
- Referenced dimensions must include obsolete records.

Related topics

[Best Practices for Data Warehouse Administrators](#)

[ETL data flow](#)

[JDBC connections](#)

[Best practices: Data warehouse administration](#)

[Data warehouse checklist](#)

[Data warehouse architecture](#)

[Example: Data warehouse customization](#)

[Example: Index creation](#)

Example: Data warehouse customization

You can customize the data warehouse to meet your specific requirements. In the following example, the administrator adds a new dimension column to the dept_d table, modifies a dimension column in the device_d table, and changes the fact key column.

```

/*****
/****  Add new dimension column for dept_d table *****/
/*****
--Step 1, Clean up DEPT_D and DEPT_F, if you want to keep all history, you
can skip this step
--delete from DEPT_D where Z_RDSDEPT_DID > 5;
--delete from DEPT_F where Z_RDSDEPT_DID > 5;
--commit;

--Step 2, Alter table to add a new column, please modify the name and data
type, modification required!!!
ALTER TABLE DEPT_D
  ADD RBFfield varchar2(60);

--Step 3, Modify manually the connect.it scenario to add the mapping for
the new field from SC to RDS, modification required!!!

--Step 4,  Modify rds_sc.ini to rollback scheduler pointer for deptSrc
back to April 01, 2007, modification required!!!

/*****
/*****Modify dimension column for device_d table *****/
/*****
--Step 1, Clean up DEVICE_D and DEVICE_F, if you want to keep all history,
you can skip this step
--delete from DEVICE_D where Z_RDSDEVICE_DID > 5;
--delete from DEVICE_F where Z_RDSDEVICE_DID > 5;
--commit;

--Step 2, Alter table to add a new column, please modify the name and data
type, modification required!!!
ALTER TABLE DEVICE_D
  MODIFY columnName column_type;

--Step 3, Modify manually the connect.it scenario mapping for the modified
field from SC to RDS, modification required!!!

--Step 4,  Modify rds_sc.ini to rollback scehduler pointer for deviceSrc
back to April 01, 2007, modification required!!!

/*****
/*****Change Fact key column, such as rootcaus_f *****/
/*****
--Step 1, Clean up ROOTCAUS_F
delete from ROOTCAUS_F where Z_RDSDEVICE_DID > 5;
commit;

--Step 2, Modify rds.xml for the key field definition, modification
required!!!

/* change the line from
<factKey name="START_TIMEID" fieldName="PROBLEM_START_TIME"
tableName="RDS_TIMEDIM_D" matchFieldName="FULLDATE"/>
to
<factKey name="START_TIMEID" fieldName="OPEN_TIME"
tableName="RDS_TIMEDIM_D" matchFieldName="FULLDATE"/>
*/

--Step 3, after the next wake up, ROOTCAUS_F should be repopulated

```

