

Data Integration and ETL with Oracle Warehouse Builder: Part 2

Volume I • Student Guide

D70627GC20

Edition 2.0

December 2010

D71328

ORACLE®

Author

Richard Green

Key Contributor

David Allan

**Technical Contributors
and Reviewers**

Michelle Bird

Herbert Bradbury

Dan Gallagher

Gerry Jurrens

Joseph Klein

John Leigh

Padmaja Potineni

Puja Singh

Vishwanath Sreeraman

Gary Tripp

Editors

Daniel Milne

Amitha Narayan

Graphic Designer

Rajiv Chandrabhanu

Publishers

Sumesh Koshy

Veena Narasimhan

Jobi Varghese

Copyright © 2010, Oracle and/or its affiliates. All rights reserved.

Disclaimer

This document contains proprietary information and is protected by copyright and other intellectual property laws. You may copy and print this document solely for your own use in an Oracle training course. The document may not be modified or altered in any way. Except where your use constitutes "fair use" under copyright law, you may not use, share, download, upload, copy, print, display, perform, reproduce, publish, license, post, transmit, or distribute this document in whole or in part without the express authorization of Oracle.

The information contained in this document is subject to change without notice. If you find any problems in the document, please report them in writing to: Oracle University, 500 Oracle Parkway, Redwood Shores, California 94065 USA. This document is not warranted to be error-free.

Restricted Rights Notice

If this documentation is delivered to the United States Government or anyone using the documentation on behalf of the United States Government, the following notice is applicable:

U.S. GOVERNMENT RIGHTS

The U.S. Government's rights to use, modify, reproduce, release, perform, display, or disclose these training materials are restricted by the terms of the applicable Oracle license agreement and/or the applicable U.S. Government contract.

Trademark Notice

Oracle and Java are registered trademarks of Oracle and/or its affiliates. Other names may be trademarks of their respective owners.

Contents

Preface

I Introduction

- Objectives 1-2
- Course Objectives 1-3
- Agenda: Day 1 1-4
- Agenda: Day 2 1-5
- Appendices 1-6
- Questions About You 1-7
- Summary 1-8

1 Administrative Tasks in Warehouse Builder

- Objectives 1-2
- Lesson Agenda 1-3
- Enterprise ETL License Extends Core In-Database ETL 1-4
- Scenario Requiring Multiple Configurations 1-6
- Loose Coupling of Metadata with Physical Implementations 1-7
- Each Configuration Records the Physical Settings of Its Target Environment 1-8
- Multiple Named Configurations: Why and How 1-9
- Deploying a Selected Configuration 1-10
- Configurations, Control Centers, and Locations 1-11
- Creating Additional Configurations and Control Centers 1-12
- Setting the Active Configuration 1-13
- Configuring Objects Across All Configurations 1-14
- Cloning a Database to Create *Identical* Dev-QA-Prod Systems 1-15
- Practice 1-1 Overview: Using Multiple Named Configurations 1-17
- Lesson Agenda 1-18
- Configuration Templates 1-19
- Creating a New Configuration Template 1-21
- Associating a Configuration Template with a Configuration 1-22
- Overriding Default Configuration Settings 1-23
- Practice 1-2 Overview: Using Configuration Templates 1-24
- Lesson Agenda 1-25
- Control Center Service Failover on RAC 1-26
- RAC Deployment 1-27

Supported and Unsupported RAC Features 1-28
Steps for Setting Up OWB in a RAC Environment 1-29
Lesson Agenda 1-30
Scheduling Concepts 1-31
Creating a Schedule Module 1-32
Creating a Schedule 1-33
Create Schedule Wizard: Name and Description 1-34
Create Schedule Wizard: Start and End Time 1-35
Create Schedule Wizard: Frequency and Repeat Interval 1-36
Monthly Schedule 1-37
Editing a Schedule 1-38
Run One Week into the Next Quarter 1-39
Associating an Executable Object with a Schedule 1-40
Assigning a Schedule to an Executable Object 1-41
Quiz 1-42
Summary 1-43
Practice 1-3 Overview: Creating a Schedule 1-44

2 Managing Metadata

Objectives 2-2
Lesson Agenda 2-3
Managing Metadata by Using Lineage and Impact Analysis Diagrams 2-4
Integrated Lineage and Impact Analysis 2-5
Invoking Lineage and Impact Analysis 2-6
Diagram-Based Analysis 2-7
Lineage and Impact Analysis Within the Same Diagram 2-8
Lineage Analysis 2-9
Lesson Agenda 2-10
Using the Change Propagation Dialog 2-11
Lesson Agenda 2-12
User-Defined Properties (UDPs) 2-13
User-Defined Icons 2-15
Creating Icon Sets 2-16
Icon Set Usage 2-17
User-Defined Objects 2-18
UDOs Participate in Lineage and Impact Analysis 2-20
Quiz 2-21
Practice 2-1 Overview: Managing Metadata 2-22
Lesson Agenda 2-23
What Is Pluggable Mapping 2-24
Using Pluggable Mappings 2-25

- Example of Pluggable Mapping 2-26
- Advantages of Pluggable Mappings 2-27
- Creating a Pluggable Mapping 2-28
- Provide the Pluggable Mapping Name 2-29
- Define the Input and Output Signature Groups 2-30
- Specify the Input Signature 2-31
- Specify the Output Signature 2-32
- The Pluggable Mapping Editor Opens 2-33
- Add the Source Tables 2-34
- Add a Joiner Operator to the Pluggable Mapping 2-35
- Define the Groups for the Joiner Operator 2-36
- Map the Inputs to the Joiner Operator 2-37
- Map the Joiner Outputs to the Output Signature 2-38
- Practice 2-2 Overview: Creating a Pluggable Mapping 2-39
- Lesson Agenda 2-40
- Advanced Activity Types in Process Flows 2-41
- Using Variables and Parameters in Process Flows 2-42
- Passing Data Between Activities by Assigning into Local Variables 2-43
- Child Process Using Input and Output Parameters 2-44
- New OWB 11g Release 2 Activity Types in Process Flows 2-45
- Creating Activity Templates 2-46
- Assign Activity 2-48
- For Loop Activity 2-49
- Route and While Loop Activities 2-50
- While Loop: Example 2-51
- Notification Activity 2-52
- Set Status Activity 2-53
- Data Auditor 2-54
- Native Relational Object Support 2-55
- Heterogeneous Predefined SQL Transformations 2-57
- Summary 2-59

3 Accessing Non-Oracle Sources

- Objectives 3-2
- Lesson Agenda 3-3
- Extensible Framework of OWB 11g Release 2 3-4
- What Is a Code Template? 3-5
- An Open Extensibility Framework 3-6
- Benefits of Extensible Code Templates 3-7
- What Is a Code Template Mapping? 3-8
- Typical Usage of Code Templates 3-9

Lesson Agenda 3-10
Location of Seeded Code Templates 3-11
Quiz 3-12
Lesson Agenda 3-13
Imported CT Opened in OWB CT Editor 3-14
Practice 3-1 Overview: Examine Built-In and Imported Code Templates 3-15
Lesson Agenda 3-16
Creating New Code Templates 3-17
Lesson Agenda 3-18
What Is an Integration Platform 3-19
Define New Integration Platforms in OWB 3-20
Setting Up Platform Properties 3-21
Creating Data Type Mappings 3-22
Adding Connectivity 3-23
Using Code Templates in Mappings 3-24
Hybrid Design: Combine Template Extract with PL/SQL 3-25
Quiz 3-26
Summary 3-27

4 Designing Mappings with the Oracle Data Integration Enterprise Edition License

Objectives 4-2
Lesson Agenda 4-3
Traditional Versus CT Mappings 4-4
Integrate CTs into an OWB Mapping Framework 4-5
Additional Tasks for Using CTs in Mappings 4-6
Lesson Agenda 4-7
Execution Units in a CT Mapping 4-8
Execution View Versus Logical View 4-9
Default Configuration and Its Execution View 4-10
Alternate Configuration and Its Execution View 4-11
Mechanics of Template-Based Mappings 4-12
New Properties in the Location Dialog Box Support Code Template Mappings 4-13
Location Information for CT-Based Mappings 4-14
Staging Table Prefixes 4-15
Create a List of Favorite Operators 4-16
Quiz 4-17
Lesson Agenda 4-18
Convert a Classic Mapping to a CT Mapping That Utilizes Data Pump 4-19
Copy and Paste a Classic Mapping to a Template Mapping Node 4-20
Execution View of a Code Template Version of the Mapping 4-21

Assign the USE_DATA_PUMP Code Template 4-22
Lesson Agenda 4-23
Reminder of CT-Based Mapping Architecture 4-24
OWB Topology with CCA on Oracle Target Server 4-25
Privileges for J2EE Users and Agents 4-26
1. Starting the Default Agent 4-28
2. Setting Up Credentials for Default Agent 4-29
3. Mapping Modules and Agent Locations 4-30
Agents Need Information About Locations 4-31
Installing the OWB Agent on Another System 4-32
OWB Infrastructure with Agent on Oracle Target 4-33
OWB Infrastructure with Agent on Remote Source 4-34
Monitoring Deployments with HRAB Reports 4-35
Quiz 4-36
Summary 4-37
Practice 4-1 Overview: Using Code Templates in Mappings 4-38

5 Right-Time Data Warehousing with OWB

Objectives 5-2
Lesson Agenda 5-3
What Is Meant by Real-Time Data Warehousing 5-4
What Refresh Frequency Does OWB Support? 5-5
OWB Supports Multiple Loading Approaches 5-6
Trickle Feed: Start/Stop 5-7
Mapping Can Consume or Produce Data to And from Queues 5-8
OWB Supports Two Types of Advanced Queues 5-9
Streams Queues Can Stage Logical Change Records (LCR) or Non-LCR
Messages 5-10
Consume from a Queue 5-11
Produce Information into a Queue 5-12
Queue-to-Queue Propagation 5-13
Lesson Agenda 5-14
Building a Trickle Feed Mapping 5-15
Building a Trickle Feed Mapping: Queue Operator Wizard 5-16
Select the Message Queue as Driver 5-17
Select Batch or Real-Time Source 5-18
Select Message Type 5-19
Add Object Type Expander for Payload 5-20
Add the Mapping Target Table 5-21
Execution Type Automatically Set to Trickle 5-22
Define the Streams Administrator User 5-23

Configure the Mapping Real-Time Parameters 5-24
Practice 5-1 Overview: Building a Trickle Feed Mapping 5-25
Lesson Agenda 5-26
Change Data Capture Framework 5-27
Simplified CDC Mapping Steps 5-28
Choose How to Perform CDC 5-29
Select Tables Upon Which to Perform CDC 5-30
Start the CDC Capture 5-31
Define Subscribers 5-32
Define Mapping to Consume Changes 5-33
Define Execution Units and Deploy 5-34
Using Web Services to Administer CDC: Select Run-Time Services 5-35
Using Web Services to Administer CDC: Edit the Web Service 5-36
Quiz 5-37
Summary 5-38
Practice 5-2 Overview: Steps for Using CDC Code Templates for CDC 5-39

6 Defining Relational Models

Objectives 6-2
Lesson Agenda 6-3
Lesson Scenario 6-4
Classifying the Data 6-5
Dimension Tables 6-6
Determining Granularity 6-7
Dimensional Design Using OWB 6-8
Defining Dimensions 6-9
Creating a Dimension by Using the Wizard 6-10
Create Dimension Wizard: Name and Description 6-11
Create Dimension Wizard: Storage Type 6-12
Create Dimension Wizard: Dimension Attributes 6-13
Create Dimension Wizard: Levels 6-14
Create Dimension Wizard: Level Attributes 6-15
Create Dimension Wizard: Data Policy 6-16
Create Dimension Wizard: Pre Create Settings 6-17
Create Dimension Wizard: Dimension Creation Progress 6-18
Continuing with the Dimension Editor 6-19
Configuring a Dimension 6-20
Quiz 6-21
Practice 6-1: Create a Dimension by Using the Wizard 6-22
Lesson Agenda 6-23
Creating or Editing a Dimension by Using the Dimension Editor 6-24

Invoking the Dimension Editor 6-25
Specifying Product Name and Roles 6-26
Selecting the Relational Storage Type 6-27
Defining Attributes 6-28
Defining Levels and Assigning Level Attributes 6-29
Defining a Hierarchy of Levels 6-30
Specifying Slowly Changing Dimension Type 6-31
Relational Implementation: Star 6-32
Relational Implementation: Snowflake 6-33
Binding 6-34
Binding Dimension Attributes to the Implementation Table 6-35
Unbinding 6-36
Dimension Details: Data Viewer Tab 6-37
Quiz 6-38
Practice 6-2: Create a Dimension by Using the Editor 6-39
Lesson Agenda 6-40
Creating a Time Dimension by Using the Create Time Dimension Wizard 6-41
Create Time Dimension Wizard: Name and Description 6-42
Specify Data Storage Type 6-43
Specify the Range of Data to Be Stored 6-44
Select Calendar Hierarchy and Hierarchy Levels 6-45
Pre Create Settings Are Displayed 6-46
Progress Panel Is Displayed 6-47
Create Time Dimension Wizard: Summary 6-48
Practice 6-3: Create a `TIMES` Dimension by Using the Create Time Dimension Wizard 6-49
Lesson Agenda 6-50
Defining a Cube by Using the Wizard 6-51
Create Cube Wizard: Name and Description 6-52
Select the Storage Type 6-53
Specify the Dimensions of the Cube 6-54
Define the Measures 6-55
Create Cube Wizard: Summary 6-56
Editing a Cube 6-57
Assigning Roles 6-58
Configuring Cubes 6-59
Lesson Agenda 6-60
Designing ETL Mappings 6-61
Mapping to Load the `PRODUCTS` Dimension 6-62
Mapping to Load the `SALES` Cube 6-63

Generating the Mappings 6-65
Summary 6-66
Practice 6-4: Create a Cubeby Using the Editor 6-67

7 More Relational Dimensional Modeling

Objectives 7-2
Lesson Agenda 7-3
Initial Versus Incremental Loads 7-4
Two Categories of DW Updating: Data and Metadata 7-5
Preserving History While Updating Fact Data 7-6
Two Ways to Refresh the Fact Table 7-7
Capturing Changed Data for Refresh 7-8
Wholesale Data Replacement 7-9
Comparison of Database Instances 7-10
Time and Date Stamping 7-11
Database Triggers 7-12
Using a Database Log 7-13
Verdict 7-14
Applying the Changes to Data 7-15
Overwriting a Record 7-16
Adding a New Record 7-17
Adding a Current Field 7-18
Maintaining History 7-19
History Preserved 7-20
Dimensions and Cubes Automatically Handle Update Via MERGE 7-21
Three Refresh Scenarios for Refreshing Target Tables 7-22
Target Uses a Sequence-Generated Key 7-23
Change “Match by Constraint” to “No Constraints” 7-24
Setting Attribute Properties for Synthetic Keys 7-25
Setting Loading Properties 7-26
Update Operation Conditions 7-28
Choosing the DML Load Type 7-29
Specifying an Update Target Condition 7-31
CDC Template Mappings: Another Method for Updating Changed Data 7-32
Quiz 7-33
Lesson Agenda 7-34
The Challenge of Managing Orphans 7-35
How OWB Manages Orphans 7-36
Lesson Agenda 7-37
ROLAP Implementation of Dimensional Objects 7-38
Support for Cube-Organized Materialized Views 7-39

Configuring the Cube 7-40
Storage Type: ROLAP with CUBE MVs 7-41
Using Compressed Cube Technology with Sparse Dimensions 7-42
View the Code After Deploying Dimensions and Cube 7-43
Examining the Cube in Analytic Workspace Manager 7-44
Use SQL Developer to Test Queries with and Without Query Rewrite 7-45
Execution Plan Without Query Rewrite 7-46
Execution Plan with Query Rewrite 7-47
Lesson Agenda 7-48
What Is a Slowly Changing Dimension? 7-49
Types of Slowly Changing Dimensions 7-50
Type 1 SCD: Does Not Store History 7-51
Type 2 SCD: Preserves Complete History 7-52
Type 3 SCD: Stores Only the Previous Value 7-53
Creating a Type 2 SCD by Using the Dimension Editor 7-54
Applying the Two Change-Tracking Attributes to the Lowest Level 7-55
Creating a Type 2 SCD 7-56
Binding the Attribute to Its Implementation Table 7-58
Synchronize Mapping That Loads Type 2 SCD 7-59
Creating a Type 2 SCD by Using the Wizard 7-60
Dimension Operator in a Mapping 7-61
Creating a Type 3 SCD 7-62
Attributes in a Type 3 SCD 7-63
Quiz 7-64
Summary 7-65
Practice 7-1: Creating a Type 2 Slowly Changing Dimension 7-66

8 Modeling Multidimensional OLAP Dimensions and Cubes

Objectives 8-2
Lesson Agenda 8-3
What Is OLAP? 8-4
Examining an OLAP Question 8-5
Lesson Agenda 8-7
Multidimensional Data Types 8-8
Implementing a Dimensional Data Model with Multidimensional Data Types 8-9
Dimensional Model 8-10
Measures 8-11
Measure Types 8-12
Example of Measures in a Report 8-13
Dimensions 8-14
Example of Dimensions in a Report 8-15

Quiz 8-16
Hierarchies 8-17
Hierarchy: Example 8-18
Levels 8-19
Types of Hierarchy 8-20
Attributes 8-21
Attributes: Examples 8-22
Dimensional Model Summarized 8-23
Lesson Agenda 8-24
Analytic Workspace 8-25
Analytic Workspace: Creation and Maintenance Privileges 8-27
OLAP DML 8-28
Lesson Agenda 8-29
Tools to Build an Analytic Workspace 8-30
Quiz 8-31
Lesson Agenda 8-32
Dimensional Modeling 8-33
Enabling OLAP Solutions 8-34
Storage Management 8-35
Dimensional Modeling Using OWB 8-36
Hierarchy Support: Value-Based 8-37
Create a Value-Based Hierarchy 8-38
Lesson Agenda 8-39
Calculated Measures 8-40
OWB Calculated Measures 8-41
Generating Calculated Measures 8-42
Sparsity 8-43
Compress Cube 8-44
Partition Cube 8-45
Cost-Based Aggregation 8-46
Level-Based Aggregation 8-47
Lesson Agenda 8-48
Differences Between OLAP and Relational Loading 8-49
No Relational Tables to Bind 8-50
Partially Predefined OLAP Module 8-51
Examine the Predefined Dimensions and Mappings 8-52
Define the Sales Cube 8-53
Create an OLAP Target User 8-54
Specify Whether to Create a Location 8-55
Location Created 8-56
Associate the Module with the Target Location 8-57

Register the Target Location 8-58
Deploying OLAP Objects with Control Center Manager 8-59
Executing OLAP Mappings 8-60
View Cube Data in the Data Viewer 8-61
Quiz 8-62
Summary 8-63
Practice 8-1: Overview 8-64

Appendix A: Practices and Solutions

Appendix B: Configuring Warehouse Builder in RAC Environments

Objectives B-2
Scope of RAC Discussed in This Appendix B-3
Clusters B-4
Oracle Real Application Clusters (RAC) B-5
Benefits of RAC B-6
OWB Certification on RAC B-7
Typical Connect Failover Scenarios B-8
Control Center Service Failover on RAC B-9
Supported and Unsupported RAC Features B-10
Lesson Progress Report B-11
Single Shared Disk Versus Local Disk Per Node B-12
Extra Tasks Required of Non-Shared Local Disk Environments B-13
Lesson Progress Report B-14
Installing OWB on Real Application Clusters B-15
OWB Installation Decisions for RAC B-16
Installing OWB with Oracle Universal Installer B-17
Specifying Oracle Home for All Nodes B-18
Selecting Cluster or Local Installation B-19
Selecting Cluster Installation B-20
Installing to the Same Path on All Nodes B-21
Installation Includes Copying to Remote Nodes B-22
Executing the root.sh Configuration Script in Each Cluster Node B-23
Lesson Progress Report B-24
Installing the Repository to One Node and Registering It to Other Nodes B-25
Installing the OWB Repository with the Repository Assistant B-26
Connection Information for the Node B-27
Finishing Installation of the Repository to a Node B-28
Rerunning the Repository Assistant to Register the Repository on All Other
Nodes B-29
Finishing OWB Repository Registration B-30

Only One Database Parameter Specifically Recommended for RAC B-31
`rtrepos.properties` Must Be Replicated to All Nodes B-32
Moving a Copy of `rtrepos.properties` to Each Node B-33
OWB RAC Locations Use Net Service Names B-34
Sample `TNSNAMES.ORA` File B-35
RAC Deployment B-36
RAC Deployment: Alternate Node B-37
Lesson Progress Report B-38
Logging In to OWB Browser B-39
Select the Service Node Report B-40
Service Node Report Shows the Status of Nodes B-41
Disabling a Node B-42
Enabling a Node B-43
Lesson Progress Report B-44
Useful Diagnostics for OWB RAC Problems B-45
Using SQL*Plus Scripts to Test Availability of Control Center Service B-46
Result of Not Replicating `rtrepos.properties` to a Node B-47
Using OWB With or Without a Control Center Service B-48
Further Study of RAC B-49
Summary B-50

Appendix C: Service-Oriented Architectures

Objectives C-2
Topics C-3
What Are Services? C-4
Service-Oriented Architecture C-5
Topics C-6
Full Support for SOA C-7
Publishing a Web Service C-8
Topics C-10
Top-Down Web Service Design Approach C-11
Topics C-12
Executing Web Services C-13
Topics C-14
Methods for Creating a Mapping That Uses Web Services C-15
Topics C-16
Consuming a Web Service C-17
Public Web Services Can Be Used by OWB Process Flows C-18
Topics C-19
Accessing Web Services Securely C-20

J2EE Roles for Java Runtime Security C-21
Demonstration Viewlet C-22
Summary C-23

Appendix D: Importing COBOL Copybooks

Objectives D-2
Flexible File Import Options in OWB 11g Release 2 D-3
Enhanced Support for COBOL Copybook Import D-4
Importing a Source COBOL Copybook D-5
Create a Location Pointing to the Copybook D-6
Start the COBOL Copybook Import D-7
Begin the Process for Selecting Copybooks to Import D-8
Select the Copybook D-9
Click Session Options D-10
Examine Advanced Import Session Options D-11
Invoke the Import D-12
Examine the Newly Created File D-13
View the Imported Fields D-14
Summary D-15

