

Data Integration and ETL with Oracle Warehouse Builder: Part 1

Volume I • Student Guide

D70625GC20

Edition 2.0

November 2010

D69902

ORACLE®

Author

Richard Green

Key Contributor

David Allan

**Technical Contributors
and Reviewers**

John Leigh

Puja Singh

Padmaja Potineni

David Wenberg

Joyce Scapicchio

Michelle Bird

Herbert Bradbury

Gerry Jurrens

Tom O'Shaughnessy

Barry Cook

Antonio Romero

Kevin Yang

David Dong

S. Matt Taylor Jr.

Graphic Editor

Satish Bettegowda

Editors

Vijayalakshmi Narasimhan

Daniel Milne

Publishers

Giri Venugopal

Michael Sebastian Almeida

Sumesh Koshy

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
- Course Agenda: Day 1 1-4
- Course Agenda: Day 2 1-5
- Course Agenda: Day 3 1-6
- Appendixes 1-7
- Questions About You 1-8
- Summary 1-9

1 Installing and Setting Up the Warehouse Builder Environment

- Objectives 1-2
- Lesson Agenda 1-3
- What Is Oracle Warehouse Builder? 1-4
- Producing Quality Information 1-5
- Managing the Full Life Cycle of Metadata and Data 1-6
- Basic Process Flow of Design and Deployment 1-7
- Oracle Warehouse Builder Licensing 1-9
- Connectivity Options 1-12
- BI Environment from an OWB Perspective 1-13
- Full Support for SOA 1-14
- Lesson Agenda 1-15
- Installing Oracle Warehouse Builder 11.2 1-16
- Simpler OWB 11.2 Installation 1-17
- OWBSYS Schema 1-18
- OWB 11.2 Server Installation 1-19
- OWB 11.2 Stand-Alone Installation 1-20
- Using OWB 11.2 with Database 10g R2 1-21
- Script for Creating OWBSYS in Database 10g R2 1-22
- Making OWB Oracle Home Known to Database 10g R2 1-23
- Unlock OWBSYS and OWBSYS_AUDIT 1-24
- Quiz 1-25
- Lesson Agenda 1-26
- Warehouse Builder Workspace 1-27
- Workspace Installation 1-28
- Workspace Owners and Users: Job Functions 1-29
- Starting the Repository Assistant 1-30
- Using the Repository Assistant to Manage Workspaces 1-31
- Dropping a Workspace 1-34
- OWB Product Components 1-35
- Lesson Agenda 1-36
- Supported Operating Systems 1-37
- Oracle Database Support: Source, Target, Repository 1-38
- Installing Optional Components 1-39

Simplified Installation of Oracle Workflow 1-41
Lesson Agenda 1-43
Classroom Configuration 1-44
Oracle Warehouse Builder Documentation and Resources 1-45
Quiz 1-46
Summary 1-47
Practice 1-1 Overview: Creating a Workspace, Workspace Owner, and a Workspace User 1-48

2 Getting Started with Warehouse Builder

Objectives 2-2
Lesson Agenda 2-3
Logging In to OWB Design Center 2-4
OWB 11.2 Uses the Oracle IDE Interface 2-5
Design Center: Start Page 2-6
Design Center: Dockable Panels 2-7
Projects Navigator, Locations Navigator, and Globals Panel 2-8
Structure Panel 2-9
Structure Panel Versus Object Editor 2-10
Property Inspector 2-11
Graphical Navigator 2-12
Advanced Find 2-14
Help Menu and the OWB Help Center 2-16
Managing Deployment with the Control Center Manager 2-18
Quiz 2-19
Lesson Agenda 2-21
Projects 2-22
Objects in a Project 2-23
Objects Within a Project 2-25
Objects Within an Oracle Module 2-29
Objects Within a Module 2-30
Organizing Metadata Using Foldering 2-32
Lesson Agenda 2-35
Locations, Connectors, and Control Centers 2-36
Globals Navigator Objects 2-38
Lesson Agenda 2-41
Creating a Project 2-42
Deleting a Project 2-43
Lesson Agenda 2-44
Setting Project Preferences 2-45
Recent Logon Drop-Down List 2-48
Enabling Recent Logon List 2-49
Scripting in Design Center 2-50
Quiz 2-51
Summary 2-53
Practice 2-1 Overview: Examining the Design Center Interface 2-54

3 Understanding the Warehouse Builder Architecture

- Objectives 3-2
- Lesson Agenda 3-3
- Warehouse Builder Development Cycle 3-4
- Architecture for Design, Deployment, Execution 3-5
- OWB Topology with CCA on Oracle Target Server 3-11
- Configurations, Control Centers, Locations 3-12
- Lesson Agenda 3-14
- Locations 3-15
- Creating Locations 3-16
- Metadata and Data Locations 3-17
- Registering Locations 3-18
- Unregistering Locations 3-19
- Alternatives for Unregistering Locations 3-20
- Connectors 3-21
- Quiz 3-23
- Control Centers 3-24
- Control Center and Locations 3-26
- Configurations and Control Centers 3-27
- Lesson Agenda 3-28
- Creating Target Schemas 3-29
- Registering DB User as an OWB User 3-31
- Using the Security Node to Register an OWB User 3-32
- Roles and Privileges of Warehouse Builder Users 3-33
- Registering an Oracle Workflow User 3-34
- Course: Predefined and New Objects 3-35
- Import the CLASS_PROJECT1.mdl File 3-36
- Quiz 3-37
- Summary 3-38
- Practice 3-1 Overview: Setting Up a Pre-Populated Project 3-39

4 Defining Source Metadata

- Objectives 4-2
- Lesson Agenda 4-3
- Data Warehouse Implementation: Typical Steps 4-4
- Metadata to Data 4-6
- Difference Between Obtaining Relational and Flat File Source Metadata 4-7
- Source Module 4-8
- Lesson Agenda 4-9
- Data Sources 4-10
- Creating a Flat File Source Module 4-11
- Create Module Wizard: Connection Information 4-12
- Metadata to Data 4-14
- Importing Source Metadata 4-15
- Selecting the Flat File for Sampling 4-16
- View the Sample File 4-17
- Flat File Sample Wizard: Welcome 4-18
- Describing the File 4-19
- Sampling Simple Delimited File 4-20
- Setting Field Properties 4-21
- Flat File Sample Wizard: Summary 4-22

Advanced: Record Organization 4-23
Advanced: Selecting the File Format 4-24
Advanced: File Layout 4-25
Advanced: Field Properties 4-26
Quiz 4-27
Lesson Agenda 4-29
Data Sources 4-30
Creating an Oracle Module 4-31
Create Module Wizard: Name and Description 4-32
Create Module Wizard: Connection Information 4-33
Editing the Oracle Database Location 4-34
Finishing Module Creation and Proceeding to Import 4-36
Import Metadata Wizard: Filter Information 4-37
Selecting the Tables for Import 4-38
Completing the Metadata Import 4-39
Viewing Data 4-40
Quiz 4-41
Summary 4-42
Practice 4-1 Overview: Defining and Importing Flat Files and a Relational Source 4-43

5 Defining ETL Mappings for Staging Data

Objectives 5-2
Lesson Agenda 5-3
Purpose of a Staging Area 5-4
Data Staging Area 5-5
Data Warehouse Implementation: Typical Steps 5-7
Metadata to Data 5-8
Defining Staging Area in OWB 5-9
Lesson Agenda 5-10
OWB Mapping 5-11
Building a Mapping 5-12
Creating a New Mapping 5-13
Mapping Editor Canvas 5-14
Tools to Manage the Mapping Editor Canvas Operators 5-15
Grouping and Ungrouping 5-16
Spotlighting in the Mapping Editor 5-18
Practice 5-1: Working in the Mapping Editor 5-19
Lesson Agenda 5-20
Creating Mappings for Flat File Sources 5-21
What Are External Tables? 5-22
Creating an External Table 5-23
Create External Table: File Selection 5-24
Create External Table: Locations 5-25
Configuring External Tables 5-26
Practice 5-2: Creating a Staging Module and an External Table 5-27
Including the External Table in a Mapping 5-28
Creating a Staging Table 5-29
Create and Bind 5-30
Adding an Attribute 5-31
Levels of Synchronizing Changes 5-32
Synchronizing: Outbound and Inbound 5-33

- Synchronize All in a Mapping 5-35
- Quiz 5-36
- Lesson Agenda 5-38
- Using a Filter Operator 5-39
- Using the Expression Builder to Define a Filter Condition 5-40
- Using a Data Generator 5-41
- Lesson Agenda 5-42
- Mapping Source Columns to the Target 5-43
- Using the Automapper in the Mapping Editor 5-44
- Lesson Agenda 5-45
- Expression Operator 5-46
- Copying Operators Across Mappings 5-47
- Expression Editing and Copying 5-48
- Quiz 5-49
- Lesson Agenda 5-50
- Setting Loading Type 5-51
- Target Load Ordering 5-52
- Overriding the Default Load Order 5-53
- Lesson Agenda 5-54
- Validating the Mapping 5-55
- Generating the Mapping 5-56
- Quiz 5-57
- Summary 5-58
- Practice 5-3 Overview: Defining ETL Mappings for Staging Data 5-59

6 Using the Data Transformation Operators

- Objectives 6-2
- Lesson Agenda 6-3
- Component Palette 6-4
- Using a Joiner 6-5
- Setting the Join Condition 6-6
- Setting the Joiner Input Role 6-7
- Lookup Operator 6-8
- Lookup Condition 6-9
- Handling Multiple Match Rows 6-10
- Using the Subquery Filter Operator 6-12
- Sequence Operator 6-13
- Set Operation Operator 6-14
- Using the Splitter Operator for Multitable Insert 6-15
- Quiz 6-16
- Pivot and Unpivot Operators 6-17
- Wizard-Based Pivot Definition 6-18
- Resulting Pivot Code and Mapping 6-19
- Aggregator Operator 6-20
- Using a Constant Operator 6-22
- Using the Expression Builder to Define a Constant Operator 6-23
- Using the Transformation Operator 6-24
- Transformation Library 6-25
- Transformation Details 6-26
- Pre/Post Processing Operators 6-27
- Pre/Post Mapping Operator: Example 6-28

- Lesson Agenda 6-29
- Deploying Database Objects and Mappings 6-30
- Deploying in Projects Navigator Panel 6-31
- Executing a Mapping 6-32
- Quiz 6-33
- Summary 6-34
- Practice 6-1 Overview: Using the Joiner, Sequence, and the Transformation Operator in a Mapping 6-35
- Practice 6-2 Overview: Using the Subquery and the Lookup Operator 6-36

7 Cleansing and Match-Merging Name and Address Data

- Objectives 7-2
- Lesson Agenda 7-3
- Integrating Data Quality into ETL 7-4
- Name and Address Data Cleansing 7-5
- Name and Address Server 7-7
- Name and Address Software Providers 7-8
- Reviewing MAP_NA Mapping 7-9
- Examine the SOURCE Table Data 7-10
- Settings in the Name and Address Operator 7-11
- Deploying and Executing the Mapping 7-15
- CLEANSSED Name and Address Data 7-16
- Quiz 7-17
- Lesson Agenda 7-18
- Consolidating Data Using the Match Merge Operator 7-19
- Match Merge Operator 7-20
- Match and Merge Rules 7-21
- Using the Match Merge Operator in a Mapping 7-22
- Defining the Match Bins 7-26
- Defining the Match Rules 7-27
- Defining an Active Custom Rule 7-29
- Defining the Merge Rules 7-30
- Target Tables 7-31
- Merged Output 7-32
- Quiz 7-33
- Summary 7-34
- Practice 7-1 Overview: Cleansing and Match-Merging Name and Address Data 7-35

8 Using Process Flows

- Objectives 8-2
- Lesson Agenda 8-3
- Data Warehouse Implementation: Typical Steps 8-4
- Process Flow Concepts 8-5
- Creating a Process Flow Module 8-6
- Creating an Oracle Workflow Location 8-7
- Creating a Process Flow Package 8-8
- Creating a Process Flow 8-9
- Lesson Agenda 8-11
- Types of Activities 8-12
- Adding a Fork Activity 8-14

- Adding a Mapping Activity 8-15
- Adding an And Activity 8-16
- Adding End Activities 8-17
- Lesson Agenda 8-18
- Creating Transitions Between Activities 8-19
- Transition Conditions 8-20
- Lesson Agenda 8-21
- Manual Activity 8-22
- SQLPLUS Activity 8-23
- Email Activity 8-24
- Some More Activities 8-25
- Lesson Agenda 8-26
- Validating the Process Flow 8-27
- Generating the Process Flow Package 8-28
- Quiz 8-29
- Summary 8-30
- Practice 8-1 Overview: Creating a Simple Process Flow 8-31

9 Deploying and Reporting on ETL Jobs

- Objectives 9-2
- Lesson Agenda 9-3
- Data Warehouse Implementation: Typical Steps 9-4
- Logical Versus Physical 9-5
- Physical Implementation 9-6
- Configuring for Physical Implementation 9-7
- Setting Object Configuration 9-8
- Deployment Concepts 9-9
- Registering Locations 9-10
- Unregistering Locations 9-11
- Configurations and Control Centers 9-12
- Active Configuration 9-13
- Control Center and Locations 9-14
- Quiz 9-15
- Lesson Agenda 9-16
- Invoking the Control Center Manager 9-17
- Control Center Manager Components 9-18
- Control Center Manager Interface 9-19
- Various Deploy Actions 9-21
- Deploy Options and Preferences 9-22
- Deployment Preferences 9-23
- Control Center Jobs 9-25
- Job Details Window 9-26
- OWB Automatically Creates a Connector 9-27
- Quiz 9-28
- Lesson Agenda 9-29
- Data Warehouse Implementation: Typical Steps 9-30
- Executing Mappings 9-31
- Viewing Data 9-32
- Practice 9-1 Overview: Deploying, Loading, and Viewing Data 9-33
- Lesson Agenda 9-34
- Repository Browser 9-35

- Common Repository Browser Tasks 9-36
- Starting OWB Browser Listener 9-37
- Starting the Repository Browser 9-38
- Logging In to the Repository Browser 9-39
- Browsing Design Center Reports 9-40
- Browsing Control Center Reports 9-42
- Types of Control Center Reports 9-43
- Object Summary Report 9-44
- Deployment Report 9-45
- Execution Summary Report 9-46
- Execution Report 9-47
- Quiz 9-48
- Summary 9-49
- Practice 9-2 Overview: Viewing Design and Control Center Reports 9-50

10 Using the Mapping Debugger

- Objectives 10-2
- Lesson Agenda 10-3
- Mapping Debugger: Overview 10-4
- Mapping Debugger 10-5
- Lesson Agenda 10-6
- Warning: Not All Test Data Found 10-7
- Test Data Tab Shows Tables Not Found 10-8
- Operator Debug Binding Dialog Box 10-10
- Direct Access to Database Entity 10-11
- Define the Other Source Table 10-12
- Replace Target with a Debugging Table 10-13
- Create a Copy of the Target Table 10-14
- Reinitialize the Debug Session 10-15
- Reinitializing the Session 10-16
- Set Breakpoints and Watch Points 10-17
- Watch Panel For Non-Data-Based Operators 10-18
- Quiz 10-19
- Summary 10-21
- Practice 10-1 Overview: Using the Mapping Debugger 10-22

11 Enhancing ETL Performance

- Objectives 11-2
- Lesson Agenda 11-3
- Devise a Performance Strategy Early 11-4
- Performance Tuning at Various Levels 11-5
- ETL Design: Mappings 11-6
- Performance-Related Parameters in ETL Design 11-7
- Pros and Cons of Operating Modes 11-8
- Configuring Mappings for Operating Modes 11-9
- DML Error Logging 11-11
- Commit Control 11-12
- Setting Default Audit Level 11-13
- Additional Run-Time Parameters for Mappings 11-14
- Enable Partition Exchange Loading (PEL) for Targets 11-15
- Best Practices Tips: ETL Design Factors That Impact Performance 11-17

Quiz 11-19
Performance-Related Parameters in Schema Design 11-21
Benefits of Using Schema Design Capabilities in Warehouse Builder Design Center 11-22
Indexing 11-23
Configuring Properties of Indexes 11-24
Index Performance Considerations: Drop Indexes Before the Load Process 11-25
Index Performance Considerations: Drop Indexes Before the Load Process (Example Mapping) 11-26
Constraints Management 11-27
Configuring Constraints Individually 11-28
Defining Partitions in Warehouse Builder 11-30
Parallelism 11-32
Enable Parallel DML 11-33
Setting Tablespace Properties 11-34
Minimum Error Logging 11-35
Gathering Statistics 11-36
Analyze Table Statements Property 11-37
Gathering Schema Statistics 11-38
Quiz 11-39
Summary 11-40
Practice 11-1 Overview: Configuring Performance Parameters for Mappings and Tables 11-41

12 Managing Backups, Development Changes, and Security

Objectives 12-2
Lesson Agenda 12-3
Metadata Loader Utilities (MDL) 12-4
Uses of Metadata Loader Utilities 12-5
Exporting Objects 12-6
Importing Objects 12-8
Accessing the MDL Utilities 12-9
Metadata Loader Log File 12-10
Importing Metadata from Previous Versions 12-11
Upgrading the Repository 12-12
Lesson Agenda 12-13
Manage Life-Cycle Changes 12-14
Managing Metadata Changes by Using Snapshots 12-16
Creating Metadata Snapshot 12-17
Using Change Manager 12-19
Exporting and Importing Snapshots 12-20
Quiz 12-21
Lesson Agenda 12-22
Version Management of Design Objects 12-23
Version Management: Using Generation Code Comments 12-25
Version Management: Using Annotations in Export/Import 12-26
Version Management: Using Collections and Snapshots to Version-Control 12-27
Quiz 12-29
Practice 12-1 Overview: Managing Metadata Changes, Code, and MDL File Versions 12-30
Lesson Agenda 12-31
Warehouse Builder Users 12-32
Warehouse Builder User Considerations 12-33
Graphical UI for Security Management 12-34

Workspace Users Roles and Privileges 12-35
EVERYONE Role 12-37
ADMINISTRATOR Role 12-38
Object-Level Security 12-39
Access Privilege Policies 12-40
Setting Security Parameters 12-41
Quiz 12-43
Summary 12-44
Practice 12-2 Overview: Managing Security 12-45

13 Integrating with Oracle Business Intelligence Enterprise Edition (OBI EE)

Objectives 13-2
Lesson Agenda 13-3
Business Justification: Tools Integration 13-4
Integrating with OBI EE 13-5
Integration with OBI SE (Discoverer) 13-6
Lesson Agenda 13-8
Transfer BI Metadata to OBI EE Server 13-9
Setting Up the UDML File Location 13-10
Creating a New Catalog Folder 13-11
Deriving the BI Metadata (OBI EE) 13-12
Deriving the BI Metadata 13-13
Deploying the BI Module 13-16
Converting the UDML File for OBI EE 13-17
Oracle BI Admin and Answers Tool 13-18
Quiz 13-19
Summary 13-20
Practice 13-1 Overview: Derive, Deploy, and Transfer Business Intelligence Metadata into OBI EE Server 13-21
Practice 13-2 Overview: Invoking the Demo titled “Reporting Using the OBI EE Answers” 13-22

A Practices and Solutions

B Creating Experts

Objectives B-2
Perception: Power and Complexity of OWB Challenges New Users B-3
Requirement: Harnessing OWB Power and Complexity for New Users B-4
Solution: OWB “Experts” Give Directed Guidance and Knowledge Management B-5
Major Benefits of Experts B-6
Profile of Designers and Users of Experts B-7
Framework of Experts B-8
Task Sequence: No User Complexity B-9
Creating an Expert B-10
Defining an Expert B-11
Expert Editor B-12
Defining the Tasks B-13
Starting an Expert B-14
Creating Your Own Custom Dialog B-15
Use Cases B-16
Scenario: ROLAP to MOLAP in Five Easy Steps B-17
ROLAP to MOLAP in Five Easy Steps B-18

Scenario: Expert for Creating External Table B-19
Summary B-20

C Using Diagnosis and Debugging Techniques

Objectives C-2

Mandatory Reading C-3

Collecting Information Before Contacting Oracle Support C-4

Sequence Used by Oracle Support Representatives to Process Calls C-6

Timing Your Service Request C-7

Activating Debugging and Logging for Full Java Debug Trace C-8

Activating Debugging and Tracing C-9

Activating Debugging and Logging C-10

Activating Tracing C-11

Using the `service_doctor.sql` Script C-12

Using SQL*Plus Scripts to Test Availability of the Control Center Service C-13

Killing a Job If a Mapping Hangs C-14

Troubleshooting Heterogeneous Sources and Targets C-15

Troubleshooting and Diagnosing Errors in Control Center Agent (CCA) C-16

Encouraging Designers to Use the Mapping Debugger C-18

Run-Time Views and Utilities C-19

Public Run-Time Views C-20

Public Run-Time Utilities C-21

Online Warehouse Builder Resources C-25

Summary C-26

