Oracle Database 19c: Data Warehousing Techniques LVC

Oracle Database

DURATION	MODULES	COURSE CODE
3 Days	**11 Lectures**	**—**

Course Overview

The course covers the following topics: Data warehousing concepts Data warehousing techniques related to data extraction, transformation, and loading phases Various data capture methods for data refresh Summary management and implement materialized views Role of warehouse metadata Data warehouse implementation considerations

What You Will Learn

Course Overview

- Objectives
- Introduce Yourself
- Lesson Agenda
- Course Objectives
- Prerequisites and Suggested Prerequisites
- Lesson Agenda
- Course Road Map
- Lesson Agenda
- Use Case: NuMart Electronics
- Course Persona: Amy and John
- Lesson Agenda
- Course Environment
- Sample Database Schemas Used in the Course
- SRC_APAC and SRC_USA Logical Schema
- Sales Persons’ Details File Structure
- The Target Data Warehouse Schema (TRG_SALES)
- Class Account Information
- Lesson Agenda
- Oracle 19c Data Warehousing Documentation
- Continuing Your Education: Recommended Follow-Up Classes
- Resources to Bookmark

- Summary
- Practices

Data Warehousing Concepts

- Objectives
- John Begins His Training
- Pre-Assessment
- Pre-Assessment Question 1
- Pre-Assessment Question 2
- Pre-Assessment Question 3
- Lesson Agenda
- Data Warehouse Characteristics
- Subject Oriented
- Integrated
- Time Variant
- Nonvolatile
- Lesson Agenda
- Data Warehouse Architectures
- Basic Data Warehouse
- Time to Ponder
- Data Warehouse Architecture with Staging Area
- Time to Ponder
- Data Warehouse Architecture with Staging Area and Data Marts
- Lesson Agenda
- Types of Data Marts
- Dependent Data Mart
- Independent Data Mart
- Hybrid Data Mart
- Lesson Agenda
- Data Warehouse Schemas
- Star Schema Model
- Star Dimensional Modeling
- Factless Fact Tables
- Time to Ponder
- Snowflake Schema Model
- Third Normal Form (3NF)
- Lesson Agenda
- Extraction, Transformation, and Loading (ETL) Process
- ETL: NuMart Electronics Use Case
- Extraction, Loading, and Transformation (E-LT)
- ETL Versus E-LT
- Lesson Agenda
- Data Warehousing: On-Premises Tools
- Data Enrichment Tools: On Premises
- ODI Studio: Graphical User Interface to ODI

- Oracle Golden Gate: Solution for Real-Time Data Integration
- Model Enrichment Tools: On Premises
- Oracle Database Supported Features for ETL
- Business Enrichment Tools: On Premises
- Data Warehousing: Cloud Services
- Autonomous Database Cloud for Data Warehousing
- Quiz
- Summary
- Practices

Data Extraction Techniques

- Course Road Map
- Objectives
- Lesson Agenda
- Can You Identify the Data Sources for NuMart Electronics’s Warehouse?
- Examining Data Sources
- Production Data
- Archive Data
- Internal Data
- External Data
- Extracting Data
- Mapping Data
- Quiz
- Lesson Agenda
- Can You Locate Some Factors that Influence Data Extraction?
- Extraction Methods
- Logical Extraction Methods
- Incremental Extraction: Change Data Capture
- Change Tracking Methods in Oracle Database
- Physical Extraction Methods
- Extraction Techniques
- Brainstorming Activity
- Lesson Agenda
- Considerations in Designing Data Extraction Process
- Maintaining Extraction Metadata
- Possible Failures During Data Extraction
- Maintaining ETL Quality
- Lesson Agenda
- Which of These Components Cease to Exist with the Proposed
- Data Warehouse?
- Proposed NuMart Electronics’s Data Warehouse System
- Data Integration and Oracle Data Integrator(ODI)
- ODI Key Components
- Data Extraction Using ODI Studio
- Define the Topology—Physical Architecture

- Define the Topology—Logical Architecture
- Set Up a New ODI Project
- Create Models and Extract Data by Reverse Engineering
- Check the Quality of the Extracted Data
- Summary
- Practices

Introduction to Data Transformation

- Course Road Map
- Objectives
- Lesson Agenda
- John Tries to Load Data—Will He Succeed?
- What Is Data Transformation?
- Data Staging Models: Remote Staging Model
- Data Staging Models: On-Site Staging Model
- Multistage Data Transformation
- Pipelined Data Transformation
- Quiz
- Lesson Agenda
- Observe and Answer
- Transformation Routines
- Quiz
- Lesson Agenda
- What Is the Key to a Successful Warehouse Implementation?
- Benefits of Data Quality
- Quality: Standards and Improvements
- Myths and Facts About Data Quality
- Data Quality: Solutions and Management
- Lesson Agenda
- Examine the Sample Data and Answer the Question
- Multiple Local Standards Challenge
- Challenge #1: Multiple Local Standards
- Other Challenges with Data Transformation
- Challenge #2: Multipart Keys
- Challenge #3: Multiple Files or Sources
- Challenge #4: Missing Values
- Challenge #5: Duplicate Values
- Challenge #6: Element Names
- Challenge #7: Element Meanings
- Challenge #8: Input Formats
- Challenge #9: Referential Integrity Constraints
- Challenge #10: Name and Address
- Solving Name and Address Challenge for a Sample Address
- Quiz
- Summary

Data Transformation Techniques

- Course Road Map
- Objectives
- Lesson Agenda
- John Identifies the Transformations Required for NuMart’s Data Warehouse
- Transformation Techniques: Merging Data
- Transformation Techniques: Adding a Date Stamp
- Transformation Techniques: Adding Keys to Data
- Quiz
- Lesson Agenda
- Can You Identify These?
- Transformation Timing and Location
- Designing Transformation Processes
- Transformation Tools
- Summarizing Data
- Maintaining Transformation Metadata
- Data Ownership and Responsibilities
- Choosing a Transformation Point
- Monitoring and Tracking
- Lesson Agenda
- Transformation Mechanisms Using SQL
- Example: Application of the MERGE Statement in Data Warehousing
- Multitable INSERT Statements
- Transformation Mechanisms Using PL/SQL and Table Functions
- Data Transformation Using ODI Studio
- Summary
- Practice 5-1: Creating ODI Mapping with Simple Transformations
- Practice 5-2: Creating ODI Mapping with Complex Transformations

Data Loading Techniques

- Course Road Map
- Objectives
- Lesson Agenda
- Identify the NuMart’s Data Transportation Needs
- Data Transportation and Loading in a Data Warehouse
- Transportation Options in a Data Warehouse
- Transportable Tablespaces
- Types of Data Loading
- Initial Load Versus Refresh
- Quiz
- Lesson Agenda
- Data Refresh in a Simple Extract Processing Environment
- Data Refresh in a Warehouse Processing Environment
- Lesson Agenda

- Unscramble Challenge
- Building the Loading Process
- Data Granularity
- Loading Techniques
- Loading Technique Considerations
- Quiz
- Lesson Agenda
- John Evaluates Loading Mechanisms for NuMart Electronics
- Loading a Data Warehouse with SQL*Loader
- Load Options with SQL *Loader
- Loading Data Warehouse Using External Tables
- Benefits of Using External Tables
- Example: Creating External Table
- Defining External Tables Using SQL*Loader
- Creating and Loading an External Table Using ORACLE_DATAPUMP
- Other Loading Techniques Provided by Oracle
- Data Loading Using ODI Studio
- Quiz
- Lesson Agenda
- What Next?
- Postprocessing of Loaded Data
- Indexing and Sorting Data
- Unique Indexes
- Creating Derived Keys
- Summary Management
- Filtering Data
- Verifying Data Integrity
- Steps for Verifying Data Integrity
- Standard Quality Assurance Checks
- Summary
- Practice 6-1: Loading Data from a Single Data Source
- Practice 6-2: Loading Data from Multiple Data Sources

Database Sizing, Storage, and Security Considerations

- Objectives
- Lesson Agenda
- Examine This Graph: What do you observe?
- Data Volume in a Data Warehouse Grows Multi-Fold Over Years
- Sizing the Database and Other Storage Requirements
- Estimating the Database Size: Example
- Validating Database Size Assumptions
- Testing Load Sampling
- Lesson Agenda
- Examine This Data Warehouse Query
- Indexing

- Benefits of Indexes for Data Warehouse Applications
- Important Guidelines for Managing Indexes
- Common Indexing Types Supported by Oracle Database
- Quiz
- Lesson Agenda
- Star Query Transformation
- Quiz
- Lesson Agenda
- Examine This Graph: What do you observe?
- Need for Very Large Databases (VLDBs)
- Challenges with VLDBs: Need for Data Partitioning
- What is data partitioning?
- Benefits of Data Partitioning
- Oracle’s Partitioning Strategies
- Single-Level and Composite Partitioning
- Partition Pruning
- Quiz
- Lesson Agenda
- Observe This Scenario
- Parallelism
- Automatic Degree of Parallelism (DOP)
- Operations That Can Be Parallelized
- Quiz
- Lesson Agenda
- Observe This Scenario
- Summary Data
- Materialized Views
- Analytic Views
- Lesson Agenda
- Observe This Scenario
- Data Warehouse Security
- Some More Security Requirements
- Security in Data Warehouses
- Oracle’s Strategy for Data Warehouse Security
- VPD and Oracle Label Security
- Summary

Data Refresh Techniques

- Course Road Map
- Objectives
- Examine This Scenario
- Lesson Agenda
- Data Refresh in a Warehouse Processing Environment
- Lesson Agenda
- Can you answer this?

- Data Refresh in a Data Warehouse
- Lesson Agenda
- Data Refresh Strategy
- Factors Involved in Developing a Refresh Strategy
- Load Window
- Scheduling the Load Window: Stage 1
- Scheduling the Load Window: Stage 2
- Scheduling the Load Window: Stage 3
- Lesson Agenda
- Time to Ponder
- Capturing Changed Data for Refresh
- Change Tracking Methods
- Timestamps
- Triggers
- Partitioning
- Comparison of Database at Different Times
- Using a Database Log
- Lesson Agenda
- Can You Answer This?
- Replacing the Entire Data Warehouse
- Overwriting a Record
- Overwriting a Record: Example and Limitations
- Adding a New Record
- Adding a New Record: Example and Limitations
- Adding a Current Field
- Adding a Current Field: Example and Limitations
- Time to Ponder
- Maintaining History: Techniques
- History Tables and One-to-Many Relationships
- History Tables and One-to-Many Relationships: Example
- Versioning
- Preserving Complete History
- Purging and Archiving Data
- Oracle-Supported Techniques for Purging Data
- Oracle-Supported Techniques for Archiving Data
- Lesson Agenda
- Final Tasks
- Publishing Data
- Quiz
- Lesson Agenda
- Identify the Stages where Changed Data can be Captured
- Lesson Agenda
- Data Warehouse Refresh Mechanisms in Oracle Database
- Lesson Agenda
- Change Data Capture in ODI
- Types of Journalizing and JKM Naming

- Change Data Capture Implementation in ODI
- Lesson Agenda
- CDC with ODI and OGG Integration
- ODI to OGG Integration Knowledge Modules
- Summary
- Practice 8-1: Implementing CDC in ODI for NuMart Use-Case
- Practice 8-2: Integrating ODI and OGG for Change Data Capture

Summary Management

- Course Road Map
- Objectives
- Lesson Agenda
- Observe This Conversation and Answer
- Using Summaries to Improve Query Performance
- Using Summaries to Improve Performance
- Summary Management
- Summary Navigation
- Managing Historical Summary Data in the Warehouse
- Lesson Agenda
- Time to Ponder
- Challenges with Summary Tables
- Components of Summary Management
- Materialized Views
- Determining Which Materialized View to Create
- Using Materialized Views for Summary Management
- Creating a Materialized View Using the CREATE SQL Statement: Example
- Using Summaries Without Materialized Views: Example
- Using Materialized Views for Summary Management: Example
- Quiz
- Lesson Agenda
- Brainstorming Activity
- Refresh Modes for Materialized Views
- Manual Refresh Using the DBMS_MVIEW Package Procedures
- Using the DBMS_MVIEW Package: Available ON DEMAND Refresh Methods
- Refreshing at Scheduled Time: Using the START WITH and NEXT Clauses
- Time to Ponder
- Synchronous Refresh for Materialized Views
- Data Dictionary Views for Materialized View Refresh Statistics
- Managing Materialized View Refresh Statistics
- Real-Time Materialized Views
- Quiz
- Lesson Agenda
- Can you identify the missing components?
- Query Rewrite
- Cost-Based Query Rewrite Process

- Conditions Required for Oracle to Rewrite a Query
- Query Rewrite
- Quiz
- Summary
- Practice 9-1: Creating and Refreshing Materialized Views
- Practice 9-2: Demonstrating Query Rewrite Using a Materialized View

Metadata Management

- Course Road Map
- Objectives
- Lesson Agenda
- Examine These Questions
- Metadata
- Defining Metadata
- Metadata Users
- Lesson Agenda
- Try to Match
- Types of Metadata
- Examining Metadata: ETL Metadata
- Extraction Metadata
- Transformation Metadata
- Loading Metadata
- Examining Metadata: End-User Metadata
- End-User Metadata: Context
- Historic Context of Data
- Types of Context
- Lesson Agenda
- Crossword Challenge
- Metadata Strategy
- Primary Considerations in Metadata Strategy
- Define Goals and Intended Use of Metadata
- Identify Target Users
- Choose Metadata Tools and Techniques
- Choose Metadata Location
- Manage Metadata
- Integrate Multiple Sets of Metadata
- Manage Changes to Metadata
- Additional Metadata Content and Considerations
- Lesson Agenda
- Common Warehouse Metamodel
- Oracle Enterprise Metadata Management (OEMM)
- Manage Metadata with Oracle Data Integrator
- Reverse Engineering
- Summary

Data Warehouse Implementation Considerations

- Course Road Map
- Objectives
- Lesson Agenda
- Can you identify these phases of a Data Warehouse Project?
- Planning
- Project Management Plan
- Design and Build
- Requirements Specification or Definition
- Logical, Dimensional, and Physical Data Models
- Data Warehouse Architecture
- ETL Considerations
- Reporting Considerations
- Security Considerations
- Metadata Management
- Testing and Implementation
- Testing the Implementation
- Testing and Implementation
- Post-Implementation Change Management
- Summary