

Oracle GoldenGate 12c: Integrate Big Data

The course covers the Oracle GoldenGate 12c (12.3) Java adapter and Big Data/NoSQL modules which allow relational databases to integrate with the Hadoop ecosystem and NoSQL databases such as MongoDB and Cassandra.

Each Hadoop component or NoSQL database is adequately introduced, the various parameters and configuration options are explained and where appropriate best practices are demonstrated.

Lab practices complete the didactic experience of students who have the chance to configure and perform data replication between an Oracle RDBMS (relational database management system) and the various technologies supported by the Oracle GoldenGate for Big Data software.

Learn To

The course shows how to integrate more traditional relational databases with the current wave of new technologies such as Big Data (in particular, Hadoop) and the NoSQL databases MongoDB and Cassandra.

Hadoop is defined as an Ecosystem consisting of many components and technologies. The course covers several of the most important Hadoop components, such as HDFS, HBase, Flume and Kafka. Students have access to a lab environment where they can configure data replication from an Oracle RDBMS (relational database management system) to the Hadoop system using the various Hadoop technologies, therefore validating in practice what learned at a more theoretical level from the instructor and the slides.

The NoSQL databases MongoDB and Cassandra are also extensively covered by the course. Again, lab practices are designed to give students a practical exposure to the concepts introduced by the instructor.

Other technologies explained in the course are the Generic JDBC Handler, which allows for integration with all systems accessible through a JDBC driver (like spreadsheets and flat files) and Pluggable Formatters, which allow for information stored in the Oracle GoldenGate trail files to be sent to Hadoop or NoSQL databases using the Oracle GoldenGate 12c Data Handlers.

Prerequisites

Suggested Prerequisite

- Students should be familiar with the Oracle GoldenGate 12c product
- While not required, a generic knowledge of SQL would help.

Required Prerequisite

- Oracle GoldenGate 12c: Fundamentals for Oracle

Audience

- Configuration Consultant
- Data Warehouse Administrator
- Data Warehouse Administrator
- Data Warehouse Administrator
- Database Administrators
- Database Administrators
- Database Designers
- Database Designers
- System Integrator
- System Integrator
- Data Warehouse Analyst
- Data Warehouse Analyst
- Technical Consultant

- Data Warehouse Developer
- Data Warehouse Developer

Course Objectives

Course Topics

Oracle GoldenGate for Big Data 12c: Product Overview

- Oracle GoldenGate Adapter Architecture - the Overall Framework
- Configuring Message Capture
- Message Parsing
- Java Adapter and Oracle GoldenGate for Big Data: Compare and Contrast
- Oracle GoldenGate for Big Data
- Oracle GoldenGate Java Adapter: Message Capture and Delivery
- Java Handlers
- Configuring Message Delivery

Configuring and Using the Hadoop HDFS Handler

- HDFS Handler Supported Formats
- Relevant HDFS Handler Configuration Parameters
- HDFS Partitioning
- HDFS Handler Performance Considerations
- HDFS Handler Best Practices
- HDFS Metadata Change Events
- HDFS Overview
- Hadoop Sequence Files

Configuring and Using the Hadoop HBase Handler

- HBase Handler Best Practices
- Relevant HBase Handler Configuration Parameters
- HBase Column Family
- HBase Metadata Change Events
- HBase Row Key
- HBase Handler Functionality
- HBase Overview

Configuring and Using the Apache Flume Handler

- Metadata Change Events
- Mapping Transactions to Transactional Flume Events
- Relevant Flume Handler Configuration Parameters
- Flume Overview
- Flume Handler Schema Propagation
- Mapping Operations to Flume Events
- Mapping Transactions to Operational Flume Events
- Flume Handler Setup

Configuring and Using the Apache Kafka Handler

- Kafka Handler Advanced Configuration
- Kafka Handler Best Practices
- Kafka Handler Schema Propagation
- Relevant Kafka Handler Configuration Parameters
- Kafka Core APIs
- Kafka Producer Configuration File

- Kafka Handler Setup
- Kafka Overview

Configuring and Using the Cassandra Handler

- Cassandra Handler: Supported Data Types
- Cassandra Overview
- ACID Properties Versus Eventual Consistency
- Cassandra Handler Operation Processing
- Cassandra Handler Best Practices
- Relevant Cassandra Handler Configuration Parameters
- Cassandra Primary Keys
- Schema, Table, and Column Mapping

Configuring and Using the MongoDB Handler

- Relevant MongoDB Handler Configuration Parameters
- MongoDB Undo Handling
- MongoDB Java Driver
- MongoDB-Specific Considerations: Document Key Column
- MongoDB Bulk Write
- MongoDB Overview
- MongoDB-Supported Data Types

Configuring and Using the Generic JDBC Handler

- JDBC Handler Single Operation Mode
- JDBC Handler: Overview
- JDBC Handler-Supported Targets
- Relevant JDBC Handler Configuration Parameters
- JDBC Handler Error Handling

Configuring and Using Pluggable Formatters and the Metadata Provider

- Row Formatters
- Operation Formatters
- Metadata Provider Types
- Pluggable Formatters: Overview
- Metadata Provider: Overview
- Metadata Provider Configuration
- Metadata Provider: When Is It Necessary?
- Formatter Configuration