

OWB Users, Enter The New ODI World

Kulvinder Hari
Oracle

Introduction

Oracle Data Integrator (ODI) is a best-of-breed data integration platform focused on fast bulk data movement and handling complex data transformations. The 11g version of ODI advances this state of the art technology even further ahead of the rest of the industry.

Oracle Data Integrator is fully integrated with the Oracle technology stack, including Oracle Database, Exadata Database Machine, Exalogic, Big Data Appliance, WebLogic Server, Business Intelligence, and Oracle Applications. ODI is the strategic data integration platform for Oracle.

The 11g version of ODI has pushed the state of the art technology in data integration further ahead of the rest of the industry. Since the first ODI 11.1.1.3.0 release in July 2010, Oracle has continued its investments on this strategic data integration platform with the 11.1.1.5.0, 11.1.1.6.0 and 11.1.1.7.0 releases.

One of the key product objectives with Oracle Data Integrator 11g is to maximize ETL developer productivity and efficiency by providing new features that help quickly deliver high quality integration processes and meet project requirements ahead of schedule and under project budget. .

Oracle Data Integrator – ETL or ELT

The ODI architecture is organized around a modular repository, which is accessed in client-server mode by components such as the ODI Studio and execution Agents that are written entirely in Java. The architecture also includes a web-based application, the ODI Console, which enables users to access information through a Web interface and an extension for Oracle Fusion Middleware Control Console.

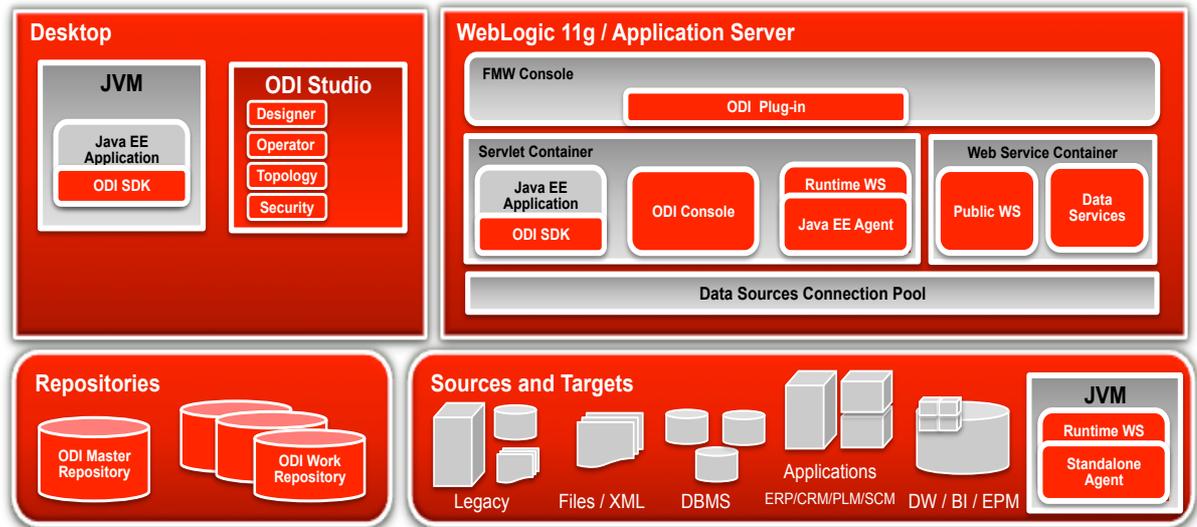


Figure 1 ODI Architecture

ODI Studio provides four graphical Navigators for managing ODI artifacts: Designer, Operator, Topology and Security. ODI Studio offers an easy to use graphical user interface and can be installed on various platforms such as Microsoft Windows, Linux and Mac OS.

At runtime, the Agent coordinates the execution of the ODI scenarios. It retrieves the code stored in the ODI repository, connects to the various source and target systems and orchestrates the overall data integration process. There are two types of ODI Agents:

- The Standalone Agent can be installed on the source or target systems and requires a Java Virtual Machine.
- The Java EE Agent is deployed on Oracle WebLogic Server and can benefit from the application server layer featured such as clustering for High Availability requirements.

With the Extract-Load Transform (E-LT) architecture, the Agent rarely performs any transformation.

It simply retrieves code from the ODI repository and then requests database servers, operating systems, or scripting engines to execute that code.

Although the Agent can act as a transformation engine, it is rarely used for that purpose. Agents are installed at tactical locations in the information system to coordinate the integration processes and leverage existing systems.

The Repository consists of a Master Repository and typically several Work Repositories. These repositories are sets of tables stored in relational database management systems such as Oracle, Microsoft SQL Server, IBM DB2 and others. All objects that the ODI modules configure, develop, or use are stored in one of these repositories, and are accessed in client-server mode by the various components of the architecture.

The Master Repository contains the security information (user profiles and privileges), the topology information (definitions of technologies and servers), and the source code for all versions of all ODI objects ever versioned.

Project objects are stored in a Work Repository. Several Work Repositories can coexist in the same installation.

A work repository stores information for

- Models (i.e. metadata)—including datastores, columns, data integrity constraints, cross references, data lineage and impact analysis
- Projects—including interfaces, packages, procedures, folders, knowledge modules, and variables
- Runtime information—including scenarios, load plans, scheduling information, and logs

Users manage the content of a work repository with the Designer and Operator Navigators in ODI Studio.

Oracle Data Integrator also includes the following components and features:

Knowledge Modules make it possible to quickly and easily integrate technologies, databases, and applications. They exist for a large range of platforms, including Oracle, Exadata, Big Data, Oracle Applications, Oracle GoldenGate, Teradata, Netezza, SAP ERP, SAP BW and XML.

- The Changed Data Capture (CDC) feature tracks changes in source systems and reduces the volume of processed data by extracting only the changed data. ODI can use native database features to capture changes or Oracle GoldenGate.
- The integration with Oracle Business Intelligence Enterprise Edition provides business users with report-to-source data lineage.

- Advanced Web Services capabilities offer best-in-class integration with Service-Oriented architecture (SOA) environments
- The Load Balancing feature enables large volumes of data to be processed by automatically balancing the workload between several Agents.
- Advanced Version Management provides an interface to manage, safeguard, and replicate revisions of units of work, even in the largest development environments.

Oracle Data Integrator – Addressing Today’s Data Challenges

Oracle Data Integrator provides a fully unified solution for building, deploying, and managing data-centric architectures for operational and analytical environments. In addition, it combines all the elements of data integration—real-time data movement, transformation, synchronization, data quality, data management, and data services—to ensure that information is timely, accurate, and consistent across complex systems.

Oracle Data Integrator provides the key components to meet today’s sophisticated enterprise data integration requirements, addressing both core ETL requirements as well as emerging trends in real-time, data quality, data management, data services federation and big data.

Oracle Data Integrator – The Way Forward

Oracle Data Integrator delivers unique next-generation, Extract Load and Transform (E-LT) technology that improves performance, reduces data integration costs, even across heterogeneous systems. Unlike conventional ETL tools, Oracle's solution doesn't require separate hardware for the ETL engines. These ETL engines can be deployed directly on the target or even the source database.

Oracle Data Integrator Enterprise Edition can be optimized for improved query performance together with Oracle Exadata, improving performance by as much as 10 times traditional ETL approaches.

In addition, Oracle Data Integrator provides capabilities for data services between Oracle SOA Suite, and rich metadata management to enable improved data lineage across Oracle BI / EPM solutions.

In addition, Oracle Data Integrator provides capabilities for data virtualization together with Oracle Data Service Integrator. Data services can be integrated across Oracle SOA Suite and Oracle Business Process Management with rich metadata management to enable improved data lineage across Oracle BI / EPM solutions.

Oracle Data Integrator is the way forward.....