

Oracle Data Integrator 12c New Features Overview

Advancing Autonomous Database and Big Data

ORACLE WHITEPAPER | SEPTEMBER 2019



Table of Contents

| Executive Overview | 6 |
|---|----|
| Oracle Data Integrator 12.2.1.4.0 | |
| Oracle Sales Cloud | 7 |
| Oracle Service Cloud | 7 |
| GIT Offline Support | 7 |
| SAP Delta Extraction | 7 |
| Oracle Data Integrator 12.2.1.3.1 | 8 |
| Oracle Object Storage and Oracle Object Storage Classic | 8 |
| Autonomous Databases | 8 |
| Oracle Enterprise Resource Planning (ERP) Cloud | 9 |
| Oracle Data Integrator 12.2.1.3.0 | |
| Big Data | 10 |
| Cloud | 11 |
| Oracle Data Integrator 12.2.1.2.6 | 12 |
| Big Data | 12 |
| Cloud | 13 |
| Lifecycle Management | 13 |
| Developer Productivity | 14 |
| Oracle Data Integrator 12.2.1.1.0 | |
| Hyperion Essbase and Hyperion Planning Knowledge Modules | 15 |
| Integrated Capture/Delivery support in GoldenGate Knowledge Modules | 15 |
| Support for Cubes and Dimensions | 15 |

| | Big Data Configuration Wizard | 15 |
|----|--|----|
| Эr | racle Data Integrator 12.2.1.0.0 | 16 |
| | Enhanced Big Data Support | 16 |
| | Lifecycle Management of Oracle Data Integrator Objects | 16 |
| | ODI Exchange for Sharing Global Oracle Data Integrator Objects | 17 |
| | Oracle Connectivity Enhancements | 18 |
| | Enhanced Integration with Oracle Enterprise Data Quality | 18 |
| | Complex File Enhancements | 18 |
| | Complex File, File, LDAP, JMS Queue XML, JMS Topic XML, and XML | 19 |
| | Pre / Post Processing for XML and Complex JDBC Drivers | 19 |
| | Improved Web Service Support | 19 |
| | Cancelling of Import/Export and Reverse Engineering Operations | 19 |
| | Support for Analytic or Window Functions | 19 |
| | Ability to View the List of Users Connected to the Studio/Repository | 19 |
| | Oracle Data Integrator Console Enhancements | 19 |
| | Oracle Data Integrator 12.1.3.0.1 | 20 |
| | Feature Highlights: Enhanced Big Data Support | 20 |
| | Introduction of Spark and Pig | 20 |
| | Orchestration of Oracle Data Integrator Jobs using Oozie | 20 |
| | Enhanced Hive Driver and Knowledge Modules | 20 |
| | Retrieval of Hadoop Audit Logs | 21 |
| | HDFS access in Oracle Data Integrator File Tools | 21 |
| | New Flatten and Jagged Components | 21 |

| 0 | racle Data Integrator 12.1.3.0.0 | 22 |
|----------------|---|----|
| Fe | Feature Highlights: Expanded Big Data Support | |
| | Sqoop Integration | 22 |
| | HBase Integration | 22 |
| | Hive Append Optimization | 22 |
| Se | ecurity Enhancements | 22 |
| | Oracle Data Integrator Federal Information Processing Standard (FIPS) | |
| | Compliance | 22 |
| D | esign-Time Experience and Productivity | 22 |
| | New Model and Topology Objects Wizard | 22 |
| | Improved Control for Scenario and Load Plan Concurrent Execution | 23 |
| Other Features | | 23 |
| | Oracle Data Integrator XML Driver Enhancements | 23 |
| | JSON Support | 23 |
| | Multi-threaded Target Table Load | 24 |
| | Documentation Changes | 24 |
| Or | racle Data Integrator 12.1.2.0.0 | |
| Fe | eature Highlights: Design-Time Experience and Productivity | 25 |
| | Design-Time Experience | 25 |
| | Reusable Mappings | 25 |
| | Multiple Target Support | 25 |
| | Step-by-Step Debugger | 26 |
| | Knowledge Module Architecture | 27 |

| Ru | Runtime Performance Enhancements | |
|----|---|----|
| | Lower Session Overhead and Enhanced Parallelism | 27 |
| | Superior Oracle GoldenGate Integration | 28 |
| | Oracle Architecture for Enterprise-Scale Deployment | 28 |
| | Oracle Platform Security Services Integration | 29 |
| Un | ified Administration and Monitoring | 29 |
| | Migrating to Oracle's Strategic Data Integration Platform | 30 |
| | Oracle Warehouse Builder Integration | 30 |
| | XML Driver Enhancements | 31 |
| | Unique Repository IDs | 31 |
| | Studio/Repository Performance Improvements | 31 |

Executive Overview

Oracle is a leader in the Data Integration market, with the industry's most comprehensive fully-integrated offering for data integration, including Oracle Data Integration Platform Cloud, Oracle Data Integrator Cloud Service, Oracle Data Integrator, Oracle GoldenGate, Oracle Enterprise Data Quality, Oracle Enterprise Metadata Management and Oracle Stream Analytics. Oracle's Data Integration solutions provide continuous access to timely, trusted, and heterogeneous data across the enterprise to support both analytical and operational data integration on-premises and in the cloud.

Oracle Data Integrator (ODI) is a best-of-breed data integration platform focused on fast bulk data movement and handling complex data transformations. It provides high-performance data movement and transformation among enterprise platforms with its open and integrated E-LT architecture and extended support for Big Data. Oracle Data Integrator is critical to leveraging data integration initiatives on-premise or in the cloud, such as Big Data management, Service Oriented Architecture and Business Intelligence. An easy-to-use user interface combined with a rich extensibility framework helps Oracle Data Integrator improve productivity, reduce development costs and lower total cost of ownership for data-centric architectures.

Oracle Data Integrator is fully integrated with Oracle Cloud, Oracle Autonomous Data Warehouse, Oracle Autonomous Transaction Processing, Oracle ERP Cloud, Oracle Fusion Middleware, Oracle GoldenGate, Oracle Enterprise Data Quality, Oracle Database, Oracle Big Data Appliance, and Oracle Exadata amongst other technologies to put data at the center of your enterprise. Oracle Data Integrator is open and standards-based to work with 3rd party applications as well as Oracle's applications. Oracle Data Integrator's E-LT architecture generates native code for disparate RDBMS or big data engines. The E-LT architecture extracts data from the disparate sources, loads it into a target, and executes transformations using the power of the database or Big Data technologies. By leveraging existing databases and big data infrastructures, Oracle Data Integrator provides unparalleled efficiency by reducing network traffic and transforming data in the server containing the target data, the E-LT architecture delivers the highest possible performance.

Oracle Data Integrator also brings critical big data integration capabilities to enterprises. It provides heterogeneity to Big Data by extending support to the most commonly used Hadoop standards. Through its decoupled design and implementation principles, Oracle Data Integrator mappings seamlessly can switch between underlying Big Data technologies. This approach provides quicker implementation and makes development for big data more efficient, removing the need for hand coding, reducing the requirement for specialist skills and future-proofing Big Data investments.

Oracle Data Integrator is the strategic data integration platform for Oracle. The 12c version of Oracle Data Integrator pushes this state of the art technology in data integration further ahead of the rest of the industry. This whitepaper describes in detail some of the new features and capabilities offered in the Oracle Data Integrator 12c platform.

Oracle Data Integrator 12.2.1.4.0

Oracle Data Integrator 12c (12.2.1.4.0) introduces new functionality in the following areas:

Oracle Sales Cloud

Oracle Data Integrator integrates with Oracle Sales Cloud. The Oracle Sales Cloud JDBC drivers are now packaged with Oracle Data Integrator. It allows you to create an Oracle Sales Cloud Technology and a Data Server, which can be used to reverse engineer the Oracle Sales Cloud Model and create ETL mappings.

For more information, see Oracle Sales Cloud chapter in the Connectivity and Knowledge Modules Guide for Oracle Data Integrator guide.

Oracle Service Cloud

Oracle Data Integrator integrates with Oracle Service Cloud. The Oracle Service Cloud JDBC drivers are now packaged with Oracle Data Integrator. It allows you to create an Oracle Service Cloud Technology and a Data Server, which can be used to reverse engineer the Oracle Service Cloud Model and create ETL mappings.

For more information, see Oracle Service Cloud chapter in the Connectivity and Knowledge Modules Guide for Oracle Data Integrator guide.

GIT Offline Support

Oracle Data Integrator (ODI) integration with Version Control System (VCS) has been extended by Git offline mode support. Now you can perform VCS operation in local Git repository and push to the remote Git Repository on need basis or when the remote Git Repository is reachable.

For more information, see Integrating ODI with Version Control Systems chapter in the Developing Integration Projects with Oracle Data Integrator guide.

SAP Delta Extraction

SAP delta extraction provides advanced table join support. This can be used to join with SAP CHDR table for delta extraction.

For more information, see Using Advanced Join Handling for Change Detection with SAP CDHDR table chapter in the Getting Started with SAP ABAP ERP Adapter for Oracle Data Integrator guide.

Oracle Data Integrator 12.2.1.3.1

Oracle Data Integrator 12c (12.2.1.3.1) introduces new functionality in the following areas:

- Integration with Oracle Object Storage and Oracle Object Storage Classic
- Optimized Knowledge Modules (KMs) for Oracle Autonomous Data Warehouse Cloud and Oracle Autonomous Transaction Processing
- Support for Oracle Enterprise Resource Planning (ERP) Cloud

These key investment areas ensure that Oracle Data Integrator (ODI) will continue to accompany customers throughout their technological transformation and modernization process.

Oracle Object Storage and Oracle Object Storage Classic

Oracle Objet Storage and Object Storage Classic offers fast, reliable and secure cloud storage and now Oracle Data Integrator can seamlessly integrate with them on Oracle Cloud Infrastructure (OCI).

ODI comes with a set of Knowledge Modules (KMs) and ODI Tools that can used in Mappings and Packages to connect to Oracle Object Storage and Object Storage Classic for uploading, downloading and deleting files/objects onto/from local directory or the Hadoop Distributed File System (HDFS).

Autonomous Databases: Autonomous Data Warehouse and Autonomous Transaction Processing

Oracle Data Integrator now comes with optimized Loading and Integration Knowledge Modules (KMs) that are certified with Oracle Autonomous databases:

- Oracle Autonomous Data Warehouse Cloud (ADW)
- Oracle Autonomous Transaction Processing (ATP)

Oracle Data Integrator seamlessly integrates with ADW and ATP. By integrating ODI with Autonomous databases, you can get the full performance of Oracle Data Integrator and Oracle databases, in a fully-managed environment that is tuned and optimized for various workloads.

The same set of Knowledge Modules are used for both ADW and ATP and also leverage the new native integration with Oracle Object Storage and Oracle Object Storage Classic.

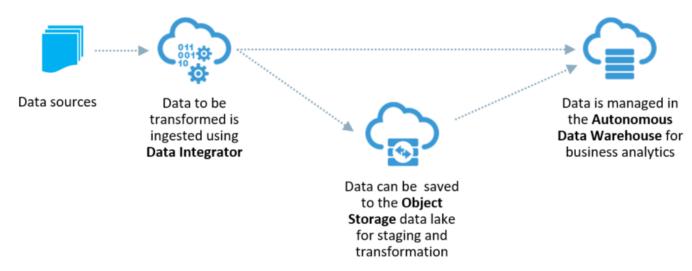


FIG: Oracle Data Integrator with Autonomous Data Warehouse Cloud

In addition to being able to load data directly into ADWC or ATP, Oracle Data Integrator users can benefit from the native integration between Oracle Autonomous Data Warehouse and Oracle Object Storage to enable extremely fast data transfer into ADWC or ATP. ODI can then automate the complete loading process of Oracle Autonomous Data Warehouse Cloud (ADW) and Oracle Autonomous Transaction Processing (ATP).

Oracle Enterprise Resource Planning (ERP) Cloud

This release also adds a new Technology and Knowledge Modules for Oracle Enterprise Resource Planning (ERP) Cloud, a suite of cloud applications for finance, project management, procurement, risk management and other core day-to-day activities important in every business, regardless of size, industry or geography.

Oracle Data Integrator (ODI) seamlessly integrates with Oracle Enterprise Resource Planning (ERP) Cloud and helps organizations integrate their ERP data into their data warehouses, data marts or data lakes. This native integration also allows ODI users to load data into Oracle ERP Cloud.

Oracle Data Integrator 12.2.1.3.0

Oracle Data Integrator 12c (12.2.1.3.0) introduces enhancements in the following areas:

- Big Data
- Cloud

Cloud and Big Data remain key investment areas and ensure that Oracle Data Integrator will continue to accompany customers throughout their technological transformation and modernization process.

Big Data

Big Data continues to evolve within Oracle Data Integrator with advances including:

- Spark Knowledge Modules Improvements— In this release, the focus has been on generating high performing and easily
 readable Spark code that will stand up to any hand written scripts. In addition, the Spark Knowledge Modules (KMs) are now
 leveraging the latest Apache Spark 2.x features such as Dataframes. SparkSQL is also leveraged where applicable to speed
 up the ODI processes running on Spark
- Spark KMs support in Knowledge Module Editor— The Spark KMs are now fully supported in the Knowledge Module editor and can be tailored to your own specific requirements

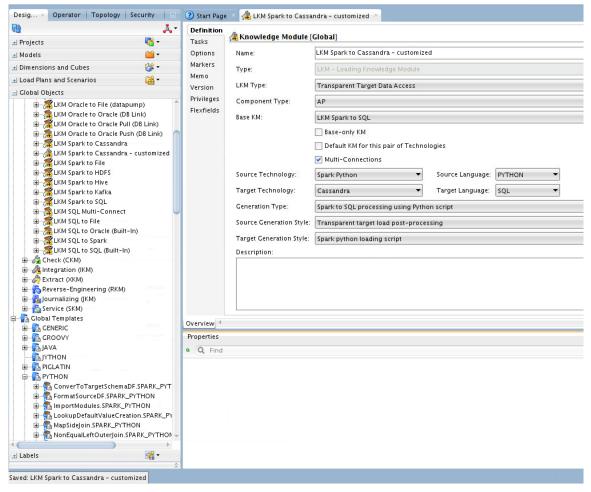


FIG: SPARK KMs IN KNOWLEDGE MODULE EDITOR

- Hadoop Complex Types Enhancements— ODI 12.2.1.2.6 introduced Complex Types support for Hive and HDFS (batch mode) technologies. ODI 12.2.1.3.0 further improves this functionality for Spark Streaming with Complex Types support in Apache Kafka as well as in HDFS
- Big Data Configuration Wizard— The Big Data Configuration Wizard has been improved with newer templates for the latest Cloudera's Distribution including Apache Hadoop (CDH) releases

Cloud

Cloud improvements include:

 Certification with Salesforce.com— Oracle Data Integrator is fully certified with Salesforce.com and now includes a JDBC driver for this technology out of the box

Oracle Data Integrator 12.2.1.2.6

Feature highlights continue to further extend our capabilities in four major areas: Big Data, Cloud, Lifecycle Management and Developer Productivity.

Big Data

Spark Streaming Support: Oracle Data Integrator (ODI) now supports Spark Streaming to fully enable the creation of Big Data streaming jobs easily without requiring end users to write a single line of code. In addition to Spark Streaming ODI already supports Hive, Pig and batch Spark when it comes to data processing. Through its unique decoupling of the Logical and Physical design of Mappings Oracle Data Integrator is the only Data Integration tool on the market giving developers the flexibility to design Mappings with a generic business logic and then generate code for as many data processing technologies (Hive, Spark, Spark Streaming etc.) as they want. This unique capability also helps future proof Data Integration processes.

Support for Apache Kafka and Apache Cassandra: Apache Kafka and Cassandra are certified with the latest version of Oracle Data Integrator as both sources and targets.

Hadoop Complex Types and Storage Format: This release further extends the market leading Hadoop support in ODI with the ability to natively access data stored in various formats such as Avro, Parquet or JSON. In addition, new features were added to leverage complex types or nested types in Mappings such as Array, Struct or Map.

Enhancements to Big Data Configuration Wizard: The Big Data Configuration Wizard introduced with ODI 12.2.1.1.0 has been improved to support new Hadoop technologies such as Kafka and Cassandra. It also now helps users configure Oracle Data Integrator with Hadoop clusters secured with Kerberos.

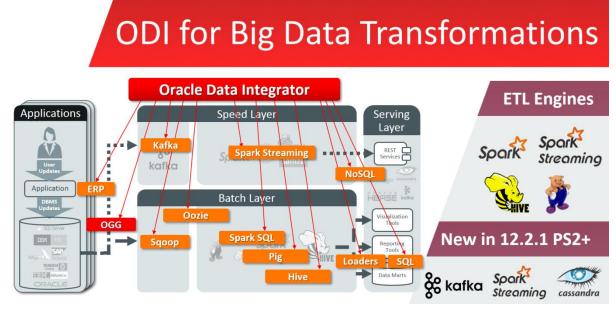


FIG: ODI FOR BIG DATA

Cloud

RESTful Service Support: Oracle Data Integrator can now invoke RESTful Service. A RESTful Service connectivity, resource URI, methods and parameters can be configured in Topology configurations like any other data source connectivity. There are a number of parameters supported providing maximum flexibility to support widespread RESTful services. Data chunking and pagination are also supported for uploading or downloading larger payloads.

Business Intelligence Cloud Service (BICS) Knowledge Modules: Business Intelligence Cloud Service is now supported out of the box in Oracle Data Integrator. You can define Business Intelligence Cloud Service connectivity in Topology, reverse engineer metadata and load data into it just like any other target data server.

Lifecycle Management

Git Support and improvements to Lifecycle Management functionality: In addition to Apache Subversion, Oracle Data Integrator now also support Git as an external version control system. A unified user experience is provided when using either of the supported version control systems. There are many advanced operations added for improved lifecycle management needs. You can now view all the pending changes and create version for set of objects from a single place. You can regenerate scenarios while creating versions or creating deployment archives ensuring scenarios always correspond to the current object version. You can selectively populate objects from the branch or a tag. The SDK APIs are also enhanced to support various operations needed for continuous integration.

Enhanced Merge Capability: The merge capabilities have been enhanced to auto-merge changes based upon three-way merge and object change detection. Conflict resolution is simplified with brand new support to pick and choose properties or objects from source or target in the Merge user interface.

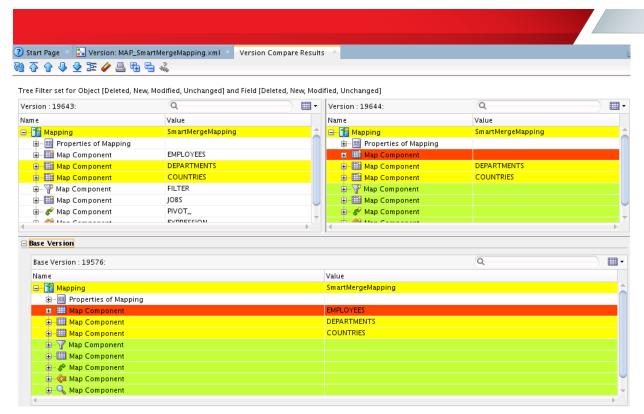


FIG: ENHANCED MERGE FUNCTIONALITY

Developer Productivity

Enhanced Knowledge Module Framework: There are exciting new features added in the Knowledge Modules development framework to maximize flexibility and minimize maintenance. You can now inherit steps from a Knowledge Module into another Knowledge Module and override steps like in object oriented programming languages. There are brand new template languages and syntaxes introduced providing greater control over the generated code. Furthermore, several other productivity enhancements were added such as syntax highlighting, auto-complete, folding code blocks and more.

Oracle Data Integrator 12.2.1.1.0

Hyperion Essbase and Hyperion Planning Knowledge Modules

Hyperion Essbase and Hyperion Planning Knowledge Modules have been made available out of the box with Oracle Data integrator and support the latest version (11.1.2.4) of these Hyperion Applications.

Integrated Capture/Delivery Support in GoldenGate Knowledge Modules

The GoldenGate Journalization Knowledge Modules (JKMs) for Oracle databases have been updated and now support Integrated Capture and Delivery. This updated functionality can improve performance and provides better scalability and load balancing.

Support for Cubes and Dimensions

Core ETL – ELT enhancements have been made; where ODI now provides support for two types of dimensional objects: Cubes and Dimensions. Users can create and use Cubes and Dimensions objects directly in Mappings to improve developer productivity with out of the box patterns that automate the loading of dimensional objects. This also allows for improved Type 2 Slowly Changing Dimensions and brand new Type 3 Slowly Changing Dimensions support with ODI.

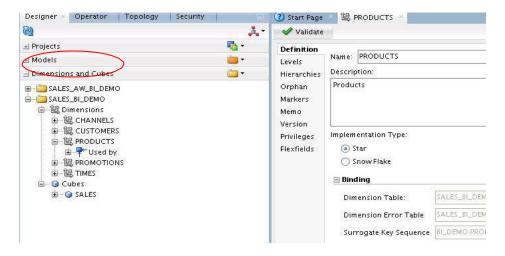


FIG: CUBES AND DIMENSIONS IN ODI STUDIO

Big Data Configuration Wizard

A brand new Big Data Configuration wizard is now available in the ODI Studio Gallery and provides a single entry point to configure the Topology objects for Hadoop technologies such as Hadoop, Hive, Spark, Pig, Oozie, etc.

Oracle Data Integrator 12.2.1.0.0

Feature Highlights: Lifecycle Management and Enhanced Big Data Support

Enhanced Big Data Support

Oracle Data Integrator's previous release brought together a series of advanced Big Data updates and features that Oracle Data Integration is rolling out for customers to help take their Hadoop projects to the next level with support for Apache Spark, Apache Pig, and orchestration using Oozie. With this release, we continue to enhance this functionality by allowing users to choose between Task and Session execution modes for Oozie workflow generation. Users can choose between Task and Session execution modes for Oozie workflow generation. The new Session mode allows support for transactions, scripting, and loops in packages. Oracle Data Integrator will automatically choose the correct mode based on the executed object, or the mode can be manually selected.

Lifecycle Management of Oracle Data Integrator Objects

Oracle Data Integrator introduces release management capabilities to provide a distinction between development and deployment environments. You can create deployment artifacts from a development environment, which can be deployed in a Quality Assurance environment for testing and then delivered to the production environment. The deployment artifacts can be created using the Oracle Data Integrator Studio or from the command line.

Oracle Data Integrator is integrated with Subversion and this provides the ability to control Oracle Data Integrator Objects in Subversion. Using this integration, you can create tags, to take a snapshot of Oracle Data Integrator object versions. You can also create branches for parallel development from distributed locations or for parallel development for multiple releases.

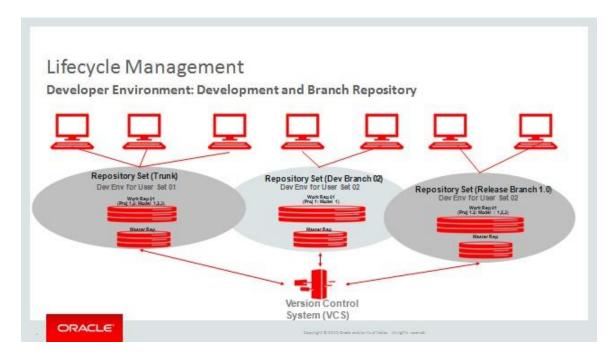


FIG: LIFECYCLE MANAGEMENT

ODI Exchange for Sharing Global Oracle Data Integrator Objects

Oracle Data Integrator is creating more of a community feel by providing a place to browse, download, and install global ODI Objects made available by Oracle or other Oracle Data Integrator users through Official or Third-Party Update Centers. This functionality is available for Global Knowledge Modules, Global User Functions, and Mapping Components. The Check for Updates menu item in the Help menu in the Oracle Data Integrator Studio enables you to connect to the Update Centers and obtain those Global Oracle Data Integrator Objects.

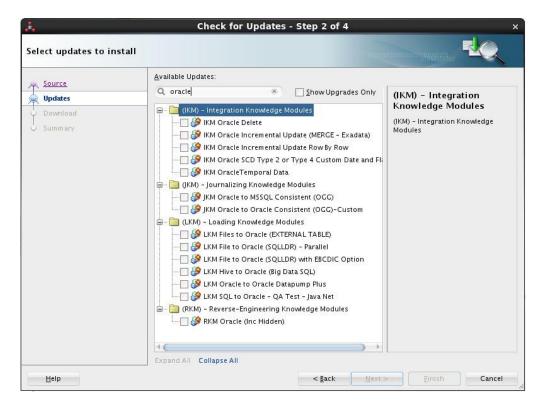


FIG: GLOBAL ORACLE DATA INTEGRATOR OBJECTS

Oracle Connectivity Enhancements

Oracle Data Integrator's repertoire of Knowledge Modules is continually improved. A Knowledge Module to perform Partition Exchange Loading is available to allow users to swap partitions as needed. Improvements have also been made to the Loading Knowledge Modules using External Tables, which allows the ability to load more than one file at a time. Data Pump Knowledge Modules have also been improved.

Enhanced Integration with Oracle Enterprise Data Quality

A new Oracle Enterprise Data Quality (EDQ) technology is available in Topology and allows the creation of data servers, physical schemas, and logical schemas for EDQ. In addition, the OdiEnterpriseDataQuality tool was improved to support EDQ data servers through Contexts and Logical Schemas.

Complex File Enhancements

Oracle Data Integrator's Native Format builder utility is now included with Oracle Data Integrator Studio and allows you to create nXSD files without having to leave the Oracle Data Integrator user interface.

Complex File, File, LDAP, JMS Queue XML, JMS Topic XML, and XML Technology Enhancements

Oracle Data Integrator's JDBC properties for Complex File, File, LDAP, JMS Queue XML, and XML technologies are now displayed at the Data Server level along with default values where applicable and a description of those properties, in an effort to enhance usability.

Pre / Post Processing for XML and Complex JDBC Drivers

Oracle Data Integrator provides the ability to customize the way data is fed into XML and Complex File drivers. This adds support for intermediate processing stages that may be added for processing data as it has either been received from the external endpoint using Oracle Data Integrator, or it is written out to an external endpoint. This also provides support for complex configuration of intermediate processing stages as part of the configuration of data servers that use ODI XML or Complex File JDBC drivers.

Improved Web Service Support

Oracle Data Integrator provides a new SOAP Web Service technology that is available in Topology and allows the creation of data servers, physical and logical schemas for Web Services. Oracle Web Service Management (OWSM) policies can now be attached to Web Service data servers. Additionally, the OdilnvokeWebService tool is enhanced to support Web Services data servers through Contexts and Logical Schemas.

Cancelling of Import/Export and Reverse Engineering Operations

Oracle Data Integrator now provides the ability to cancel import/export and reverse engineering operations that may run for a long time.

Support for Analytic or Window Functions

Oracle Data Integrator supports Analytic or Window functions out of the box at the Mapping Level. Analytic functions such as PERCENT_RANK, LAST, FIRST, or LAG can be used at the Mapping Expression level in any component.

Ability to View the List of Users Connected to the Studio/Repository

The Review User Activity menu item has been added to the Security Menu. This allows viewing, purging, and saving of user activity record in the User Connections dialog. This is available in both the Oracle Data Integration Studio and the Oracle Data Integrator Console.

Oracle Data Integrator Console Enhancements

The overall look and feel of the Oracle Data Integrator Console has been improved. Additionally, security tasks such as creating users or profiles can now be performed using the Console. Release Management activities can be performed through the Console, and functionality related to Topology tasks has been enhanced.

Oracle Data Integrator 12.1.3.0.1

Feature Highlights: Enhanced Big Data Support

Introduction of Spark and Pig

Oracle Data Integrator introduces the execution of mappings using Spark or Pig. Oracle Data Integrator allows the defining of mappings through a logical design which is independent of the implementation language. Users can select for Hadoop-based transformations between Hive, Spark, and Pig as the generated transformation code, allowing users to pick the best implementation based on the environment and use case:

- Spark: Oracle Data Integrator mappings can generate PySpark, which exposes the Spark programming
 model in the Python language. Apache Spark is a transformation engine for large-scale data processing. It
 provides fast in-memory processing of large data sets. Custom PySpark code can be added through userdefined function or the table function component.
- Pig: Oracle Data Integrator mappings can leverage Pig Latin as a transformation language and execution
 engine. Apache Pig is a platform for analyzing large data sets in Hadoop and uses the high-level
 language Pig Latin for expressing data analysis programs. Any Pig transformation can be executed either
 in local or map-reduce mode. Custom Pig code can be added through user-defined function or the table
 function component.

Orchestration of Oracle Data Integrator Jobs using Oozie

Oracle Data Integrator supports both the traditional Oracle Data Integrator Agent and Apache Oozie as the orchestration engine for Jobs – including: mappings, packages, scenarios, and procedures. Apache Oozie allows native execution on a Hadoop infrastructure without installing an Oracle Data Integrator agent for orchestration.

Users can utilize Oozie tooling to schedule, manage, and monitor the Jobs. Oracle Data Integrator uses Oozie's native actions to execute Hadoop processes and conditional branching logic.

Enhanced Hive Driver and Knowledge Modules

Oracle Data Integrator includes the WebLogic Hive JDBC driver that provides a number of advantages such as full JDBC compliance and improved performance. All Hive Knowledge Modules have been rewritten to take advantage of this new driver. Also, Knowledge Modules whose main purpose is to load from a source are now provided as Loading Knowledge Modules, enabling them to be combined in a single mapping with other Loading Knowledge Modules. A new class of "_direct load_" Loading Knowledge Modules also allows loading of target without intermediate staging. Additionally, the table function component has been extended to support Hive constructs.

New/Updated Hive Component Style Knowledge Modules:

- LKM SQL to Hive SQOOP
- LKM SQL to File SQOOP Direct
- LKM SQL to HBase SQOOP Direct
- LKM HBase to SQL SQOOP
- LKM File to SQL SQOOP
- LKM Hive to SQL SQOOP

- LKM File to Hive LOADDATA
- LKM File to Hive LOAD DATA Direct
- LKM HBase to Hive HBASE-SERDE
- LKM Hive to HBase Incremental Update HBASE-SERDE Direct
- IKM Hive Append
- LKM Hive to File Direct
- LKM Hive to Oracle OLH-OSCH
- LKM File to Oracle OLH-OSCH Direct
- LKM Hive to Oracle OLH-OSCH
- LKM Hive to Oracle OLH-OSCH Direct
- XKM Hive Sort

Retrieval of Hadoop Audit Logs

Oracle Data Integrator integrates results from Hadoop Audit Logs in Operator tasks for executions of Oozie, Pig, and other tasks. The log results show Map-Reduce statistics and provide a link back to the full Hadoop statistics in native web consoles.

HDFS access in Oracle Data Integrator File Tools

Oracle Data Integrator's file based tools which are leveraged in packages and procedures have been enhanced to include HDFS file processing. This includes copying, moving, appending, and deleting files, detecting file changes, managing folders, and transferring files using FTP directly into HDFS.

New Flatten and Jagged Components

Oracle Data Integrator introduces the Flatten component for mappings which allows complex sub-structures to be processed as part of a flat list of attributes.

Oracle Data Integrator also introduces the Jagged component that converts Key-Value lists into named attributes for further processing.

Oracle Data Integrator 12.1.3.0.0

Feature Highlights: Expanded Big Data Support

Sqoop Integration

Oracle Data Integrator introduces loading of the following sources and targets using Apache Sqoop:

- From relational databases to HDFS, Hive, and HBase through Knowledge Module IKM File- Hive to SQL (SQOOP)
- From HDFS and Hive to relational databases through Knowledge Module IKM SQL to Hive- HBase-File (SQOOP)

Sqoop enables load and unload mechanisms using parallel JDBC connections in Hadoop Map-Reduce processes.

HBase Integration

Oracle Data Integrator supports Apache HBase through a new technology and the following Knowledge Modules:

- LKM HBase to Hive (HBase-SerDe)
- IKM Hive to HBase Incremental Update (HBase-SerDe)
- RKM HBase

Hive Append Optimization

Knowledge Modules writing to Hive support the Hive 0.8+ capability and can append data to the existing data files rather than copying existing data into a new appended file.

Security Enhancements

Oracle Data Integrator Federal Information Processing Standard (FIPS) Compliance

Oracle Data Integrator leverages Advanced Encryption Standard (AES) as the standard encryption algorithm for encrypting Knowledge Modules, procedures, scenarios, actions, and passwords. You can configure the encryption algorithm and key length to meet requirements. Passwords and other sensitive information included in repository exports are now encrypted and secured by a password.

Design-Time Experience and Productivity

New Model and Topology Objects Wizard

The Create New Model and Topology Objects dialog in the Designer Navigator provides the ability to create a new model and associate it with new or existing topology objects, if connected to a work repository. This dialog enables you to create topology objects without having to use Topology editors unless more advanced options are required.

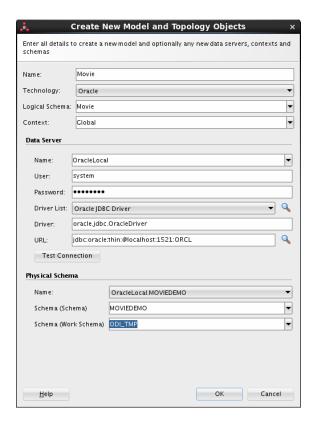


FIG: CREATE NEW MODEL AND TOPOLOGY OBJECTS WIZARD

Improved Control for Scenario and Load Plan Concurrent Execution

Improved control allows the ability to limit concurrent executions in a scenario or load plan and force a concurrent execution to either wait or raise an execution error.

Other Features

Oracle Data Integrator XML Driver Enhancements

The following XML Schema support enhancements have been added:

- Recursion: Oracle Data Integrator supports recursion inside XML Schemas.
- any, anyType, and anyAttribute: Data defined by these types is stored in string type columns with XML markup from the original document.
- Metadata annotations can be added inside an XML Schema to instruct the ODI XML Driver which table name, column name, type, length, and precision should be used.

JSON Support

The Oracle Data Integrator Complex File Driver introduces the ability to read and write files in JSON format. The JSON structure is defined through an nXSD schema.

Multi-threaded Target Table Load

Oracle Data Integrator provides the functionality to load a target table using multiple parallel connections. This capability is controlled through the Degree of Parallelism for Target property available at the data server level in Topology Navigator.

Documentation Changes

The information that was previously available in the Oracle Data Integrator Developer's Guide is now reorganized. The following new guides have been added to the Oracle Data Integrator documentation library:

- Understanding Oracle DataIntegrator
- Administering Oracle Data Integrator
- Oracle Data Integrator Tool Reference

Oracle Data Integrator 12.1.2.0.0

Feature Highlights: Design-Time Experience and Productivity

Design-Time Experience

Oracle Data Integrator 12c introduces superior productivity with a new flow-based declarative user interface. The enhanced user experience is simple yet powerful and comprehensive. The Oracle Data Integrator Studio client is entirely redesigned in this release to enhance user experience and productivity. The new user interface blends the previous declarative design model approach with a flow-based methodology. This blend simplifies common data integration design and deployment use cases, shortening implementation times. It combines the simplicity and ease-of-use of the declarative approach with the flexibility and extensibility of configurable flows. Data integration designers describe source and target data formats and data integration processes. The business user or the developer can focus on describing what to do, not how to do it. Mappings (the successor of the Interface concept in Oracle Data Integrator 11g) connect sources to targets through a flow of components such as: Join, Filter, Aggregate, Set, Split, etc. Oracle Data Integrator Enterprise Edition generates, deploys and manages the code required to implement those processes across the various source and target systems.

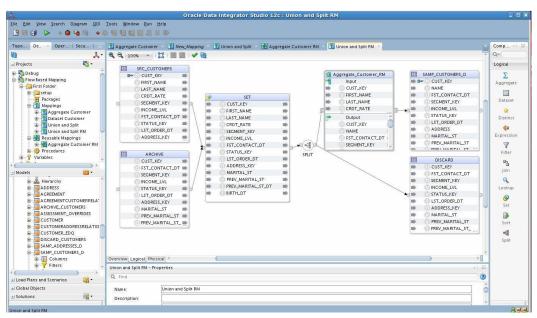


FIG: ORACLE DATA INTEGRATOR FLOW-BASED DECLARATIVE USER INTERFACE

Reusable Mappings

The paradigm is also enriched with the ability to seamlessly reuse mapping logic during development, giving developers a simpler and more efficient technique for providing solutions to their completion. Reusable mappings can be used to encapsulate flow sections that can then be reused in multiple mappings. A reusable mapping can have input and output signatures to connect to an enclosing flow and it can contain sources and targets that are encapsulated inside the reusable mapping.

Multiple Target Support

Oracle Data Integrator's flow-based declarative user interface allows for a mapping to load multiple targets as part of a single flow. Whether one or multiple targets, the order of target loading can be specified, and the Split component can be optionally used to route rows into different targets, based on one or several conditions.

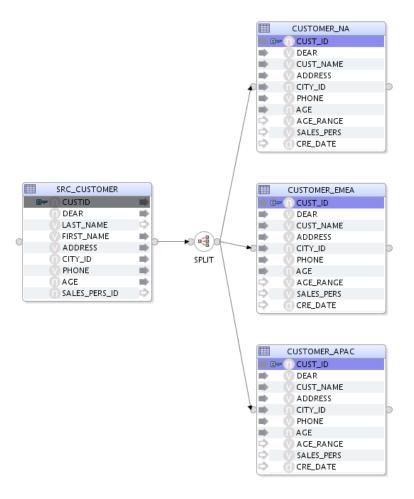


FIG: MULTIPLE TARGET MAPPING

Step-by-Step Debugger

Oracle Data Integrator has introduced a step-by-step debugger. Mappings, Packages, Procedures, and Scenarios can be debugged in a step-by-step manner. Users are able to manually traverse task execution within these objects and set breakpoints to interrupt execution at pre-defined locations.

Values of variables can be introspected and changed during a debugging session, and data of underlying sources and targets can be queried, including the content of uncommitted transactions for better insight.

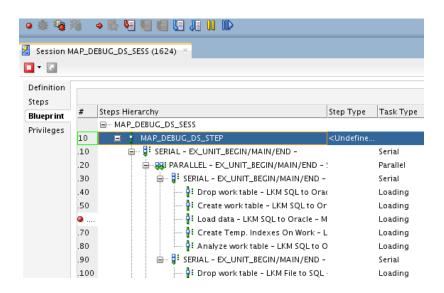


FIG: DEBUGGER - CONTROLLING THE EXECUTION FLOW

Knowledge Module Architecture

Oracle Data Integrator has introduced a new style of Knowledge Module, called Component-Style Knowledge Module in addition to Template-Style Knowledge Modules available from Oracle Data Integrator 11g. This new style of Knowledge Module provides an extensible component framework that improves the overall mapping design, where for example users are able to declare the transformation order. These also improve reusability as they can be plugged together; in addition to helping avoid code and data duplication as well as providing improved Oracle connectivity with database links.

Runtime Performance Enhancements

Lower Session Overhead and Enhanced Parallelism

Oracle Data Integrator 12c introduces improved runtime execution to enhance performance. Various changes have been made to reduce overhead of session execution, including the introduction of blueprints, which are cached execution plans for sessions. Additionally, improved parallelism functionality increases Oracle Data Integrator's already high performance architecture. The improvement comes from loading sources in parallel into the staging area. Parallelism of loads can be customized in the physical view of a map. Users also have the option to use unique names for temporary database objects, allowing parallel execution of the same mapping.

Superior Oracle GoldenGate Integration

High performance E-LT capabilities integrated with Oracle GoldenGate enable faster and more efficient loading and transformation of real-time data into a data warehouse. Customers can now easily configure and deploy real-time data warehousing solutions without impacting source systems or batch window dependencies.

The integration of Oracle GoldenGate as a source for the Change Data Capture (CDC) framework inside of Oracle Data Integrator has been improved in the following areas:

- Oracle GoldenGate Gate source and target systems are now configured as data servers in Oracle Data Integrator's Topology. Extract and replicate processes are represented by physical and logical schemas. This representation in Topology allows separate configuration for multiple contexts, following the overarching philosophy around contexts.
- Most Oracle GoldenGate parameters can be added to extract and replicate processes in the physical schema configuration. The user interface provides support for selecting parameters from lists. This minimizes the need for the modification of the Oracle GoldenGate parameter files after generation.
- A single Oracle Data Integrator mapping can be used for journalized Change Data Capture load and bulk load of a target. This is enabled through the Oracle GoldenGate Journalizing Knowledge Module using the source model as opposed to the Oracle GoldenGate replication target, as well as configuration of journalizing in a mapping as part of a deployment specification. Multiple deployment specifications can be used in a single mapping for a journalized load and bulk load.
- Oracle GoldenGate parameter files can now be automatically deployed and started to source and target Oracle GoldenGate instances through the JAgent technology.

Oracle Architecture for Enterprise-Scale Deployment

Standalone Agent Management with WebLogic Management Framework Oracle Data Integrator standalone agents are managed through the WebLogic Management Framework. The following advantages comes as a result:

- User-interface driven configuration through the Configuration Wizard
- Multiple configurations can be maintained in separate domains
- Node Manager can be used to control and automatically restart agents

Oracle Platform Security Services Integration

Oracle Data Integrator is able to use the authorization model in Oracle Platform Security Services (OPSS) to control access to resources. Enterprise roles can be mapped into Oracle Data Integrator roles to authorize enterprise users across different tools



FIG - ORACLE PLATFORM SECURITY SERVICES INTEGRATION

Unified Administration and Monitoring

Oracle introduces Management Pack for Oracle Data Integrator, which leverages Oracle Enterprise Manager Cloud Control's advanced management capabilities, to provide an integrated and top-down solution for your Oracle Data Integrator environments. Management Pack for Oracle Data Integrator provides a consolidated view of your entire Oracle Data Integrator infrastructure enabling users to monitor and manage all their components centrally from Oracle Enterprise Manager Cloud Control. Key features of the pack include the following:

- Manage multiple Oracle Data Integrator domains from a single location
- Monitor Oracle Data Integrator components availability and performance out-of-thebox; access historical data, track logs, and receive notifications of potential problems
- Trace end-to-end Oracle Data Integrator Sessions activity, review execution statistics and drill- down from a particular Step or Task into a detailed report of Oracle databases activity
- Control Service Level Agreements (SLA) with robust and scalable alerting capabilities
- Obtain real-time and historical in-depth performance statistics for the Oracle Data

Integrator Standalone and JEE Agents

Discover and model dependencies between Oracle Data Integrator and various components such as databases or other Oracle Fusion Middleware products automatically using the Oracle Enterprise Manager Cloud Control framework

Management Pack for Oracle Data Integrator supports both 11g (11.1.1.7.0 and higher) and 12c versions of Oracle Data Integrator.



FIG: MANAGEMENT PACK FOR ORACLE DATA INTEGRATOR HOMEPAGE

Migrating to Oracle's Strategic Data Integration Platform

Oracle Warehouse Builder Integration

Oracle Data Integrator provides improved interoperability with Oracle Warehouse Builder, providing the foundation for customers to start migrating to Oracle's strategic Oracle Data Integrator platform. Oracle Warehouse Builder (OWB) jobs can now be executed in Oracle Data Integrator through the OdiStartOwbJob tool. The Oracle Warehouse Builder repository is configured as a data server in Topology. Users will thus also find all the Oracle Warehouse Builder job execution details displayed as a session in Operator, Oracle Data Integrator Console, and Enterprise Manager.

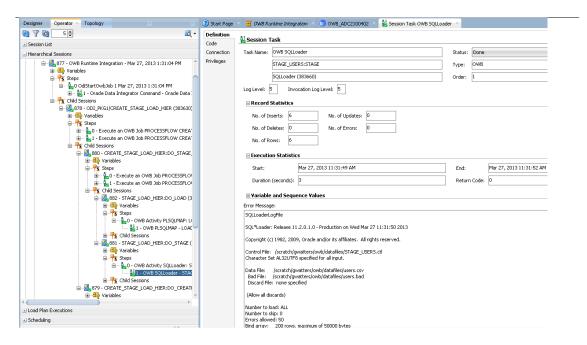


FIG - ORACLE WAREHOUSE BUILDER INTEGRATION - RUNTIME

Other Features

XML Driver Enhancements

Support for XML Schema constructs further includes:

- List and Union: List or Union based elements are mapped into VARCHAR columns.
- Substitution Group: Elements based on substitution groups create a table each for all types of the substitution group.
- Mixed Content: Elements with mixed content map into a VARCHAR column that contains text and markup content of the element.
- Annotation: Content of XML schema annotations are stored in the table metadata.

Unique Repository IDs

Master and Work Repositories now use unique IDs following the GUID convention. This avoids collisions during import of artifacts and allows simpler management and consolidation of multiple repositories in any given project or organization.

Studio/Repository Performance Improvements

Studio/Repository performance improvements have been made, reducing the chattiness for remote repository access, and optimizing database access.

Conclusion

With the Oracle Data Integrator 12c releases Oracle introduced several new enhancements such as a redesigned declarative flow-based user interface, reusable mappings, runtime performance enhancements, Oracle GoldenGate integration improvements, Cloud, Big Data and XML improvements, Oracle Warehouse Builder integration. The ODI 12c releases continue to improve Oracle's strategic Data Integration platform while preserving the key product differentiators: Declarative Design, Knowledge Modules, Hot-Plug-ability, and E-LT architecture.



Oracle Corporation, World Headquarters

500 Oracle Parkway Redwood Shores, CA 94065, USA Worldwide Inquiries Phone: +1.650.506.7000 Fax:+1.650.506.7200

CONNECT WITH US



blogs.oracle.com/oracle



facebook.com/oracle



oracle.com

Copyright © 2019, Oracle and/or its affiliates. All rights reserved. This document is provided for information purposes only, and the contents hereof are subject to change without notice. This document is not warranted to be error-free, nor subject to any other warranties or conditions, whether expressed orally or implied in law, including implied warranties and conditions of merchantability or fitness for a particular purpose. We specifically disclaim any liability with respect to this document, and no contractual obligations are formed either directly or indirectly by this document. This document may not be reproduced or transmitted in any form or by any means, electronic or mechanical, for any purpose, without our prior written permission.

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

Intel and Intel Xeon are trademarks or registered trademarks of Intel Corporation. All SPARC trademarks are used under license and are trademarks or registered trademarks of SPARC International, Inc. AMD, Opteron, the AMD logo, and the AMD Opteron logo are trademarks or registered trademarks of Advanced Micro Devices. UNIX is a registered trademark of The Open Group. 0116

White Paper: Oracle Data Integrator 12c New Features September 2019 Author: Oracle

