An Oracle White Paper
May 2010

Oracle Warehouse Builder 11gR2 New Features: Data Integration and Data Warehousing

ORACLE

Introduction

This paper provides an introduction to Oracle Warehouse Builder Enterprise ETL 11gR2 (OWB-EE), a new feature overview, and its place in Oracle's data integration roadmap.

Oracle Data Integrator (ODI) is the long-term strategic data integration offering from Oracle. Warehouse Builder Enterprise ETL 11gR2 (OWB-EE) updates the most widely adopted data warehouse design, ETL and data quality offering for the Oracle database with ODI-based data integration infrastructure. This delivers the benefits of fast, flexible, heterogeneous data integration to today's Warehouse Builder customers.

Customers can update existing designs and build new ones using the new infrastructure, and get the Oracle's latest data integration technology within a familiar tool in a form that preserves investments in existing designs and skills. OWB-EE 11gR2 lowers TCO, shortens time to value and improves ROI for Oracle data warehouse customers.

A future release, ODI 12g, will merge functionalilty from OWB and deliver a superset of the capabilities of today's products. ODI 10g/11g designs will of course be upgradable to ODI 12g, and ODI 12g will also offer a smooth migration path for OWB-EE 11gR2 designs. Customers who license ODI-EE can implement either tool today and know that their strategic investments in data integration will be protected tomorrow.

This paper provides an overview of the major new data integration, data warehousing and business intelligence features of OWB 11gR2, which serve to bring OWB more in line with the larger ODI-EE offering.

Data Integration Trends

Data integration requirements have been evolving beyond the long-established capabilities of classical ETL for data warehousing. Primary drivers of change include:

- Larger data volumes, and more heterogeneous sources and targets
- Operational requirements that shrink or eliminate batch loading windows, driving companies to near-real-time and event-driven integration
- The trend towards service-oriented architectures (SOA) requires support for closer integration of ETL and data quality processes for the warehouse with systems previously considered out of scope
- Data quality issues multiply, ranging from "garbage-in" problems with single sources to subtler data and metadata consistency and integrity issues across sources
- Economic pressures make time to value and ROI more important than ever

At the same time, newer ETL/DI use cases have been emerging alongside the well-established data warehousing and BI use cases: construction of master data stores, database and application consolidation and migration, and modernization initiatives.

Oracle Data Integrator Enterprise Edition

Oracle Data Integrator, Enterprise Edition (ODI-EE) offers a comprehensive solution for these current and emerging data integration requirements. Two tools are included in the core ODI-EE offering:

- Oracle Data Integrator (ODI), a flexible, general-purpose tool for heterogeneous ETL and data integration;
- Oracle Warehouse Builder Enterprise ETL (OWB-EE), the widely-adopted solution for end-to-end Oracle data warehousing, including warehouse design and BI tool integration.

Serving the needs of the emerging broader data integration market, ODI is the basis for Oracle's long-term data integration offering. It supports uniquely flexible heterogeneous source and target connectivity, performs EL-T processing in multiple vendors' databases, supports SOA, CDC and real-time data integration, and integrates with Oracle's middleware stack. Customers seeking a data integration tool that can address current and emerging use cases across the enterprise, with Oracle and non-Oracle databases, can standardize on Oracle Data Integrator.

However, a large Oracle Database customer base already has a large investment in skills and data warehouse ETL and data quality designs in Warehouse Builder, and is affected by these new emerging data integration requirements. To protect those investments while serving the evolving

requirements of that data warehousing customer base, OWB-EE 11gR2 refreshes OWB around the edges with ODI-based heterogeneous source connectivity and CDC, real-time mappings for event-based data integration, and basic SOA integration capabilities.

OWB-EE 11gR2 also improves upon OWB 10.2 and 11.1 in other areas that make it a better data warehousing tool than previous releases: better integration with Oracle database data warehousing features, more advanced data warehouse ETL design, updated business intelligence tool integration, and improved developer and administrator productivity.

These capabilities align OWB-EE with the overall Oracle Data Integration value proposition and with fundamental ODI technologies, without abandoning OWB's core focus on Oracle data warehousing and without requiring extensive changes to existing designs or skills. Customers can extend existing designs and take on new projects with the updated OWB-EE if this is a better fit for their situation. Oracle plans to support migration of designs based on OWB-EE to ODI 12g with no significant redesign or loss of functionality.

OWB-EE 11gR2 Data Integration Enhancements

The most important enhancements in OWB 11gR2 are the new data integration capabilities:

- Code template mappings using ODI knowledge modules
- Native heterogeneous connectivity
- Changed Data Capture mappings
- Advanced Queue Support in Mappings
- SOA Integration: Publishing and Consuming Web Servicess

Code Template Mappings Using ODI Knowledge Modules

Key data integration technologies from Oracle Data Integrator have been added to Warehouse Builder, including:

- Built-in connectivity for heterogeneous source databases, generic JDBC support, and new support for XML and flat files
- High-performance native bulk data movement for DB2 and SQL Server sources
- Extensibility for new source platforms and data movement methods
- Change Data Capture support using a variety of technologies
- Data quality controls during data loads

These features are based on ODI's knowledge module technology, which uses user-modifiable code templates for ETL code generation. Knowledge modules make it possible to generate code

for heterogeneous platforms, provide fine-grained control over generated code, and support high-performance data movement methods like bulk extraction and loading.

In OWB 11gR2, code template-based code generation has been added alongside the existing "classic" code generation style. Oracle Warehouse Builder refers to its version of ODI knowledge modules as code templates.

A new type of mapping, code template mappings, uses the new ODI-based code generation framework. Code template mappings are created in the same mapping editor as conventional mappings. The OWB mapping editor has been extended to allow creation of code template mappings as well as classic mappings.

Note that classic mappings continue to be available, exactly as before, and can still be used in existing and new designs. No forced migration to the code template mapping framework is required. Existing mappings work without change using existing connectivity technology such as gateways and database links, or can be converted to use the new data movement options.

A companion whitepaper, "Knowledge Module Technology in Oracle Warehouse Builder," will describe the OWB-EE implementation of knowledge modules and code template mappings in more detail.

Change Data Capture Mappings

As already noted, OWB now supports heterogeneous CDC, and code templates are used to generate code specific to CDC source and target platforms. OWB also provides for other aspects of CDC processing management directly in the UI:

- Specifying the CDC change set
- Managing CDC subscribers
- Starting and stopping execution of the change capture process

A future code template will support integration with Oracle GoldenGate for change capture in Warehouse Builder.

Advanced Queue Support in Mappings

OWB supports the use of advanced queues as sources and targets in ETL mappings, in batch mode or trickle-feed mode. This capability, built upon Oracle Streams, provides new event-based data integration capabilities.

SOA Integration: Publishing and Consuming Web Services

OWB now supports publishing mappings and process flows as SOAP-based web services. This permits use of OWB in scenarios where Oracle BPEL or similar products are used to coordinate activities across the enterprise.

OWB can also call other SOAP-based web services from within OWB process flows, using a new Web Service process flow activity, and pass the returned XML document into an ETL mapping for further processing.

OWB-EE 11gR2 Data Warehousing and BI Enhancements

As noted, building Oracle data warehouses continues to be the central focus for Warehouse Builder. OWB 11gR2 has added significant enhancements in the areas of data warehousing and business intelligence to improve functionality, better exploit improvements in the underlying database, and keep in step with other changes in Oracle's product offerings. The major enhancements in this area include:

- Automated orphan management policies for loading dimensional objects
- Oracle OLAP Cube-Organized Materialized Views support
- Improved Key Lookup Operator
- Integration with Oracle Business Intelligence Enterprise Edition (OBI-EE)
- Chunking for Parallelizing Large Table Updates

Automated Orphan Management for Loading Dimensions

Setting an orphan management policy for dimensions or cubes lets developers declare how to handle source rows in ETL that do not meet the requirements necessary to form a valid dimension or cube record. The following orphan management policies are available:

- Assign a default parent
- Reject orphan rows
- Perform no maintenance

Automated orphan management policies improve ETL developer and administrator productivity by addressing an important cause of cube and dimension load failures, without requiring developers to explicitly build logic to handle these orphan rows.

OLAP Cube-Organized Materialized Views Support

OWB now supports OLAP cube storage in cube-organized materialized views in Oracle Database 11g. This brings the performance advantages of such storage to users of OWB data warehouse design. Query performance is greatly improved, without the need to make any modification to your queries.

Improved Key Lookup Operator Functionality

OWB 11gR2 supports a significantly enhanced key lookup operator. Apart from more efficient use of screen real estate, the new operator includes support for non-equality lookups and dynamic lookups, where the lookup table may be modified during mapping execution. These changes make the lookup operator more powerful in many situations, including improving Type 2 slowly changing dimension support.

OBI-EE Integration

OWB now supports integration with Oracle Business Intelligence Enterprise Edition (OBI-EE). This integration includes:

- Derivation of ready-to-use physical, business model and presentation layer metadata for OBI EE from a data warehouse design, or direct design of those objects
- Visualization and maintenance of OBI-EE business definitions from within OWB
- Deployment of the derived objects in the form of an RPD file that can be loaded into OBI-EE

Like other objects maintained in the OWB repository, the OBI-EE metadata objects participate in OWB data lineage and impact analysis. Data lineage of objects in OBI EE reports can be traced back to source tables, down to the individual column level. This support complements the OBI-SE (Discoverer) support in previous versions of OWB, to bring OWB 11gR2 in line with Oracle's current strategic business intelligence tool.

Chunking for Parallelizing Large Table Updates

Chunking in Warehouse Builder automates the use of a "divide and conquer" approach to parallelize the processing of large updates. Users enable chunking for a mapping and define chunking criteria to partition the updates. OWB generates PL/SQL code for the mapping, and at execution time, updates are divided according to chunking criteria, a pool of threads is allocated, and the chunks are processed in parallel.

The benefits of applying chunking include:

- Chunking provides the only method of automatically parallelizing PL/SQL code in Warehouse Builder.

- Chunking avoids the need for large rollback segments required by set-based processing of large updates. Set-based SQL statements for large updates require large rollback segments, because a single set-based statement does not perform intermediate commits.
- Large updates can be performed incrementally, and if interrupted, chunks already processed do not have to be processed again. Without chunking, if a large update terminates for some reason, all processing must be repeated.

Chunking therefore allows fuller exploitation of hardware resources even in cases where use of set-based code is not an option or not the best option.

OWB-EE 11gR2 Administration Enhancements

Significant enhancements for OWB 11gR2 administrators include:

- Multiple Configuration Management usability enhancements
- Optional Feature Manager
- Check for Updates

Multiple Configuration Management Enhancements

For customers using multiple configuration management to handle promotion of code from development to test to production environments and associated configuration changes, there are several enhancements. The OWB user interface for viewing and managing multiple configurations has been redesigned to simplify and clarify previously complex tasks, including:

- Side-by-side editing of attribute values for multiple configurations for an object
- Editing the configuration values for an object across all configurations
- Configuring groups of objects at the same time
- Using configuration templates to set default configuration values for different object types
- Copying and pasting of configuration attribute values

These improvements add business value by enabling users to take full advantage of the flexibility provided by multiple configurations.

There are also extensions to multiple configuration management to support associating multiple execution unit and code template selections with a single logical mapping.

Optional Feature Manager

The Optional Feature Manager, available in OWB 11.2.0.1 for Win32 and OWB 11.2.0.2 on other platforms, enables repository owners to selectively enable or disable groups of OWB features that are covered by licensed OWB options. Developers working on projects in that repository are prevented from using the associated features at design time, at run time, or when importing MDL that references related features. This prevents inadvertent usage of licensed features.

OWB 11gR2 Productivity and UI Enhancements

There are extensive new and improved features in OWB 11gR2 that are focused on improving ETL developer and administrator productivity. These include:

- Redesigned OWB Design Center UI Based on JDeveloper
- Advanced search capabilities across the product
- User folders in projects for managing large numbers of objects
- Grouping and spotlighting operators in mappings
- Check for Updates

Redesigned OWB Design Center UI Based on JDeveloper

The Oracle Warehouse Builder Design Center user interface has been updated to use the Fusion Client Platform, the same core Integrated Development Environment (IDE) platform as Oracle JDeveloper and Oracle SQL Developer.

The advantages of this user interface framework include:

- More consistent behavior across different parts of the OWB user interface, and more consistency between the OWB Design Center user interface and other development tools from Oracle such as JDeveloper and SQL Developer.
- More efficient and flexible use of screen real estate.
- Support for opening multiple editors of the same type, for example, editing several ETL mappings at once in different windows.

This change brings the Design Center in Oracle Warehouse Builder in line with other development tools from Oracle.

Advanced Search Capabilities in Design Center

Advanced search capabilities have been added to OWB 11gR2 Design Center, to simplify development of large projects. For example:

- The Mapping Editor has been enhanced with advanced find capabilities to make it easier to locate and make updates to operators, groups, and attributes in a mapping diagram, in the Available Objects tab, and in the Selected Objects tab. This boosts developer productivity when working with large and complex mappings and complex data sources with large numbers of tables, views, or columns.
- The Dependency Manager, which is used to browse data lineage and impact analysis information, now includes advanced metadata searching capabilities. Users can now more easily locate specific objects in large and complex dependency graphs. This improves productivity by making it easier to find specific objects and their lineage, and discover impacts from design changes.

Mapping Debugger Enhancements

There are numerous enhancements to the OWB 11gR2 Mapping Debugger, including:

- Improved support for watch points
- Enabling and disabling of individual break points
- Support for user-defined type columns
- Enhanced support for numerous operators, such as VARRAY, EXPAND, and CONSTRUCT
- Support for key lookup and table function operators
- Support for correlated joins
- Improved cleanup of debugger-specific objects

These enhancements will improve productivity for ETL mapping developers, especially when working with complex mappings where the mapping debugger adds the most value.

User Folders in Projects

Users can create hierarchically nested folders to logically group related objects. Folders can be created within Oracle and non-Oracle database modules, non-Oracle database modules, and application modules. User folders can also be created to contain pluggable mappings. User folders can be nested as necessary to organize objects further.

Folders can be used to group related objects. Any object in one of the supported module types, such as a table or a mapping, can be moved into a folder.

For example, if a single database module contained tables, views, and ETL mappings for product and customer data, folders "Product" and "Customer" could be created, and the objects related to each category moved into the separate folders.

User-created folders improve ETL developer productivity on complex projects, by making it easier to logically group and manage large numbers of objects.

Grouping and Spotlighting of Operators in Mappings

You can now temporarily or permanently group objects in the Mapping Editor so that they are collapsed to a single icon, thus hiding complexity in mappings. Users can also spotlight a single operator, which temporarily hides all objects in the mapping except for those objects that connect directly to the operator.

Check for Updates

Like other JDeveloper-based tools, OWB 11gR2 now supports a Check for Updates mechanism that lets customers find new code templates, information resources, utilities, and other extensions for OWB 11gR2. This facility will enable Oracle to make new capabilities available to customers on an ongoing basis and in a systematic way, outside of normal release cycles.

OWB 11gR2 Resources

Future collateral will be posted on the Oracle Warehouse Builder product page on OTN, here:

http://www.oracle.com/technology/products/warehouse

Advanced techniques are frequently discussed on the Warehouse Builder Blog, at:

http://blogs.oracle.com/warehousebuilder

The Oracle Data Integration LinkedIn group, numbering over a thousand members, is home to an active community of OWB and ODI users. For details, join the group here:

http://www.linkedin.com/groups?gid=140609

The Warehouse Builder Forum on OTN continues to be the best place to ask and answer detailed questions about using Warehouse Builder. Access the forum here:

http://forums.oracle.com/forums/forum.jspa?forumID=57

Oracle Warehouse Builder 11gR2 New Features: Data Integration and Data Warehousing

May 2010
Author: Antonio Romero

Oracle Corporation
World Headquarters
500 Oracle Parkway
Redwood Shores, CA 94065
U.S.A.

Worldwide Inquiries:
Phone: +1.650.506.7000
Fax: +1.650.506.7200
oracle.com

Oracle is committed to developing practices and products that help protect the environment

0109