

Oracle Warehouse Builder 11g Release 1

An overview

March 2007

Note:

This document is for informational purposes. It is not a commitment to deliver any material, code, or functionality, and should not be relied upon in making purchasing decisions. The development, release, and timing of any features or functionality described in this document remains at the sole discretion of Oracle.

This document in any form, software or printed matter, contains proprietary information that is the exclusive property of Oracle. This document and information contained herein may not be disclosed, copied, reproduced, or distributed to anyone outside Oracle without prior written consent of Oracle. This document is not part of your license agreement nor can it be incorporated into any contractual agreement with Oracle or its subsidiaries or affiliates.

Oracle Warehouse Builder 11g Release 1

An overview

INTRODUCTION

As your organization's demands and requirements for more knowledge grow, it is important to understand the direction of the tools that help you create that knowledge. By understanding the newest capabilities you can deliver quality knowledge for your organization to act upon.

Today's information architecture is much more dynamic than it was just a few years ago. Executives now demand more information, want more and more analysis capabilities, want to predict behavior. Most importantly, executives want all of this from throughout the organization to make a difference competitively.

Oracle Warehouse Builder 11g Release 1 enables you to deliver on the information requirements of your organization. In this paper we will show you the new capabilities of the tool to create your knowledge delivery infrastructure. After reading this article you will understand how this new Warehouse Builder release will drastically improve your organization's ability to act.

IDENTIFYING THE MAIN TRENDS IN INFORMATION INTEGRATION

Many trends are being identified on a daily basis. We should, however, be careful when following or implementing all of these “latest and greatest” trends. We have to deliver information in the format in which the consumer wants it and to the location where he or she want it. This is the realm in which we need to evaluate the trends.

Keeping this prerequisite in mind, we can identify the following three trends:

- Increased demands for information quality
- The requirement to derive more knowledge from existing information
- Better time-to-market for your solutions information solutions

These trends are as important for your information consumers as for anyone developing infrastructure to deal with these trends. So let's take a closer look at what these trends mean for an organization.

Increased Information Quality

While it has been talked about for quite some time now, the days are finally coming that information assets are properly managed. The biggest impact on information asset management will come from quality problems and quality assessments. These two factors will identify the need for action in the information quality area.

A crucial point to understand is that we are not talking about raw data quality but about information quality. Because information consists of raw data and metadata (or business definitions) this broadening of the horizon will lead to higher data *and* metadata quality requirements. That change impacts not just your architecture but also processes and system designs.

Derive More Knowledge from Existing Information

As information becomes a true asset, your consumers want to get more and more results from these assets. This result is knowledge. To get more knowledge out of existing information assets people will go on unpaved roads to discover new information or to use existing information in new ways. As the technology to drill, pivot, rotate, and mine becomes more accessible, users will want more of that on their information. In short, you must be prepared to supply more and more high-end analysis capabilities to more and more people in order to produce knowledge for your organization to act upon.

Delivery from more Systems

The requirement to deliver information at the exact right time is further complicated by the demand to have more and more data consolidated at that right moment. As the information needs grow, so do the number and types of the sources that require tapping into to deliver the information

The combination of more systems to consolidate and the right-time information delivery requirements has significant ramifications for the information infrastructure you are building.

Better Time to Market

With business accelerating and information requirements exploding, you must find a way to keep up with ever shortening delivery cycles on those new information requirements.

For that, you will need a toolset that allows you to automate repetitive tasks, encapsulate business knowledge in tasks and deliver solutions quicker. Warehouse Builder embeds extensibility, automation of tasks and capturing of specialist information within its capabilities, allowing you to deliver complex systems at an ever faster pace.

PACKAGING

Oracle Warehouse Builder is an integral part of Oracle 11g Database. The product runs on all versions (Standard Edition, Standard Edition One, Enterprise Edition) and typically all platforms the Oracle 11g database is certified on and ported to. The core features of the product areas represented as a no-cost option to the database license. The options are separately priced and available with database enterprise edition only.

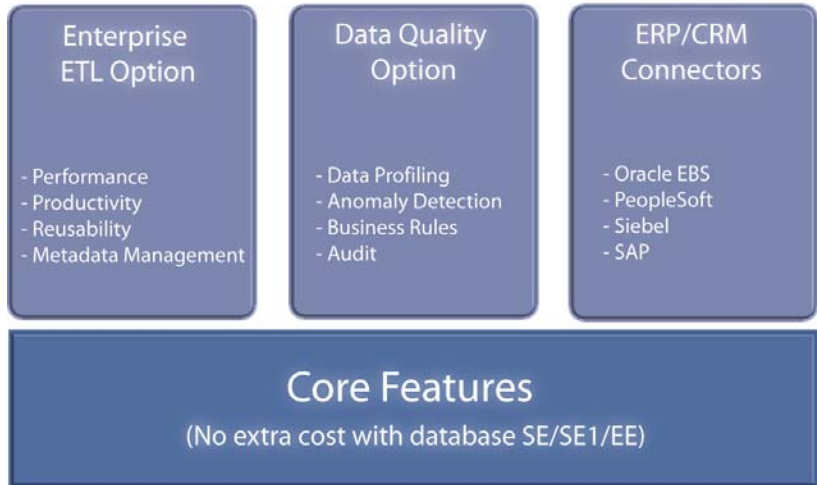


Figure 1 Packaging and Options

For all details on the options, please visit the Oracle Database Options description in the documentation or on the internet. The following sections show a high level summary of the options.

Core Features

The core features of the product are intended to enable data integration and modeling for most environments. It is an ideal stepping-stone into the Warehouse Builder world as you can start using it with no extra cost. It is by no means intended to just satisfy small implementations.

Enterprise ETL Option

This option, is specifically geared to increase performance (enabling high performance extractions) and productivity allowing for extensive reuse capabilities. Some advanced metadata management features are also incorporated into this option.

Data Quality Option

As a fully integrated part of the product, a very distinct advantage over all other tools in the market, the data quality option provides support for data profiling, data rules (in essence business rules) and information compliance features. Because of

the integration, the information gathered during data profiling can be used to automatically generate data corrections.

Connectors

Connectors allow for better access to and tighter integration with ERP and CRM packaged applications. For Warehouse Builder 11g the Siebel connector is a new addition to the Warehouse Builder connector family.

Integration is Key

To get the maximum benefit from your tool, ensure that the ETL component is integrated with the other pillars of information management.

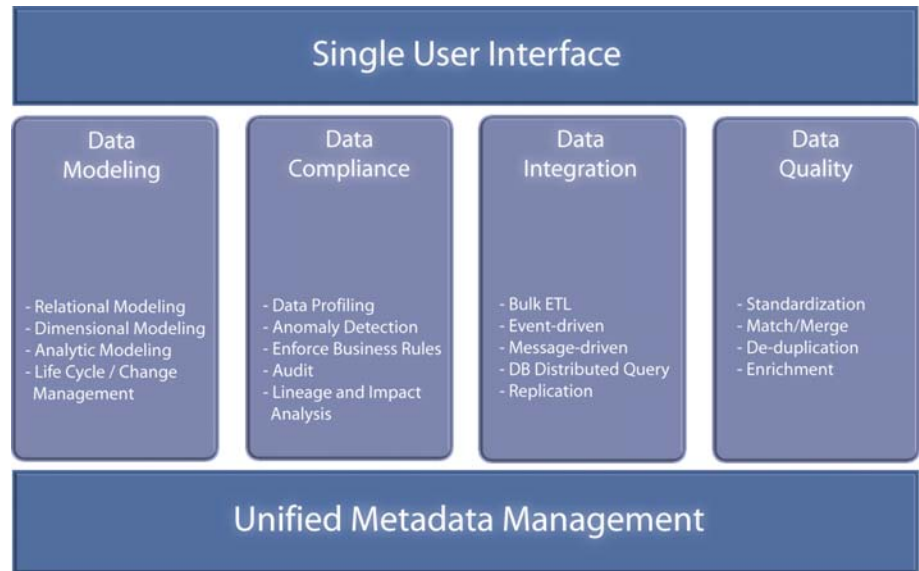


Figure 2 An integrated product

Data modeling, data compliance, and data quality are core features that your information management tool must possess.

Beyond the claims of vendors, try to understand from day one how the tools interact, how they are integrated, whether they are sold as a set, or whether they are built as a single product. These questions will show how well the product is integrated.

Now, why is integration key? For years we addressed (and some vendors still address) the topics in Figure 2 as individual tools. Then as a customer you try to figure out how to get the data model into the ETL tool and you probably end up buying some third party integration software to achieve that goal.

This case is the more common case today. When looking at the very near future, results from data profiling tools need to be communicated to the ETL developers. How is that going to work? How is that data profiling information going to flow into the direct data quality transformations required to address bad data? All of these questions should be answered with a single word, integration.

So, to really get benefits out of the tools and techniques out there, you must find an integrated tool that allows you to do the tasks you need to do faster and better.

WAREHOUSE BUILDER INCREASES INFORMATION QUALITY

As we discussed in the first trend, delivering just data is not good enough in the current and near future landscape. As an information provider or integrator you will need a toolset that will help you create quality information. Warehouse Builder forms a crucial part of that toolset.

Metadata Quality

By now you must all have heard the gospel of data quality being sung, and we all realize that data quality is important. Interestingly enough, metadata seems to evade any quality discussion, let alone any product delivering metadata quality features.

The new Warehouse Builder dramatically changes that. For the first time¹ a tool will assist you in actually resolving some of the issues created by changes in your environment.

The new Warehouse Builder Dependency Management services allow you to uncover the ripple effect of changes in metadata anywhere in your system *before* these changes break your system.

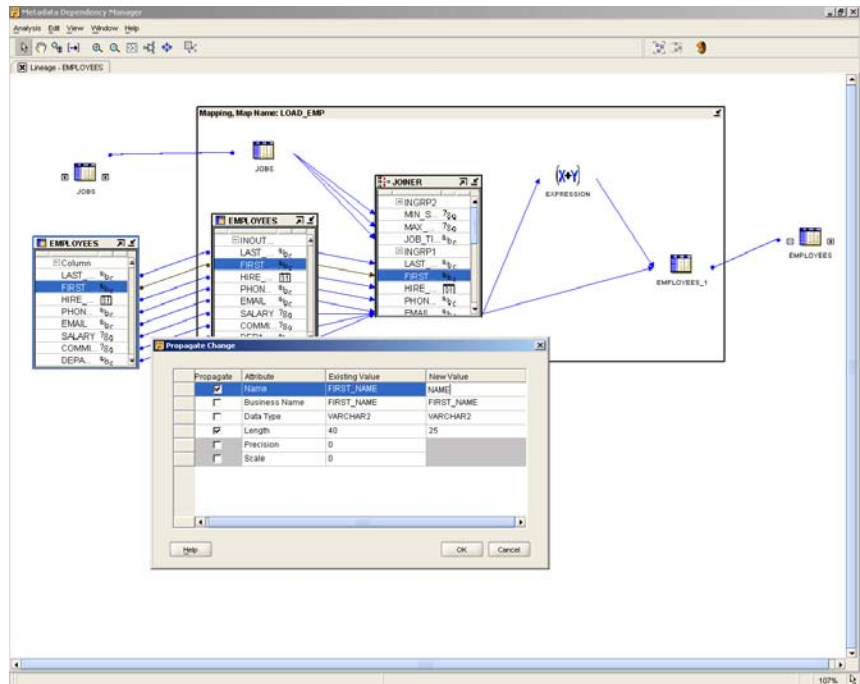


Figure 3. Propagating Change in your system

For the first time you will be proactively managing change in your system. As an example (see Figure 3), you can propagate changes through your system within the mapping editor.

¹ Patents Pending

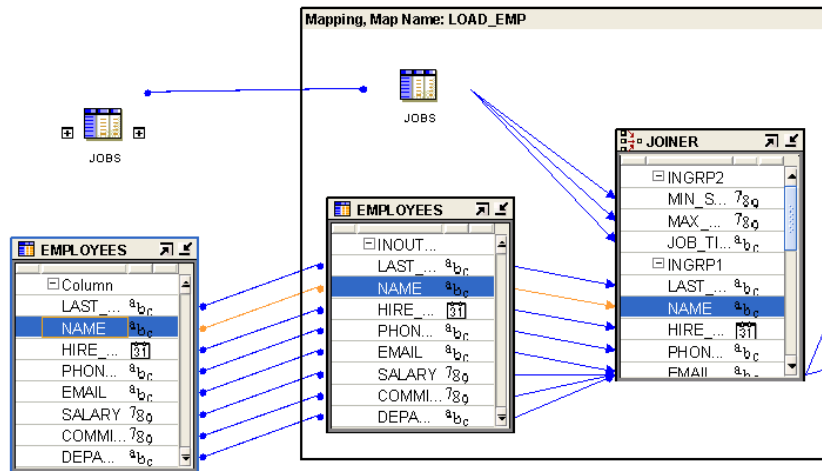


Figure 4 Propagation impact shown for attribute Name

By doing this Warehouse Builder not only saves you a lot of time and effort but also improves the quality of your metadata, allowing you to create valid information for your end users.

Data Profiling

The other crucial component in making information valuable and accurate is of course ensuring data quality levels. The new Warehouse Builder adds more punch to the already extensive data quality set now available.

The big-ticket item regarding information quality in Warehouse Builder is the capability to assess the quality of your data within Warehouse Builder, commonly known as Data Profiling. The high-level data profiling results are shown in Figure 5.

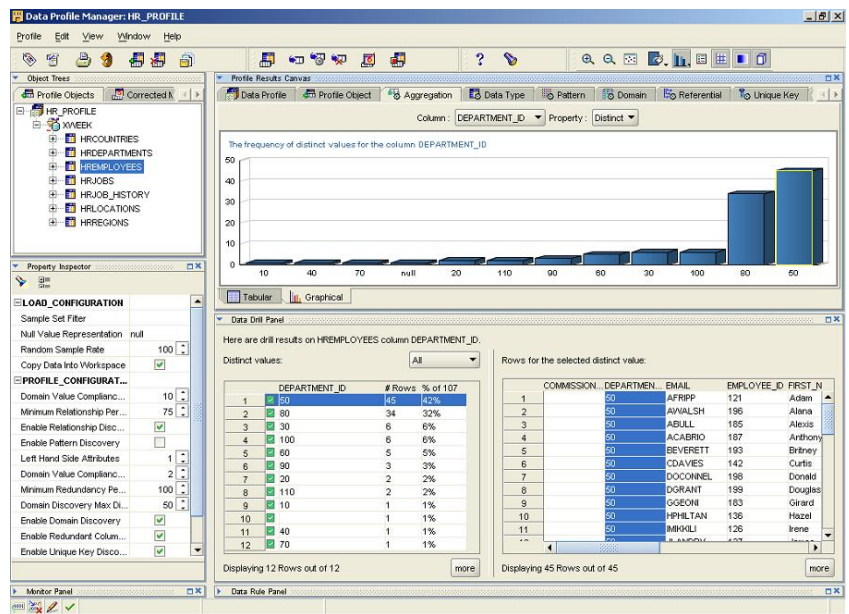


Figure 5. High-level data profiling summary

Using these profiling results Warehouse Builder allows you to derive data cleansing rules, which you can apply to cleanse the data via Extract, Transform and Load routines (ETL) called correction mappings. This integration within the creation of the ETL process is a major differentiator between Warehouse Builder and the classic profiling solutions. The integration makes data profiling with Warehouse Builder much more effective, delivering improved data quality immediately to the business.

To ensure your data quality in your target does not fall below your corporate thresholds, you can add a Data Auditor to your process that continuously monitors the quality of your data. If the quality goes below a specified threshold (this can be measured in 6-sigma or %defects), the data auditor will warn you and you can undertake appropriate action.

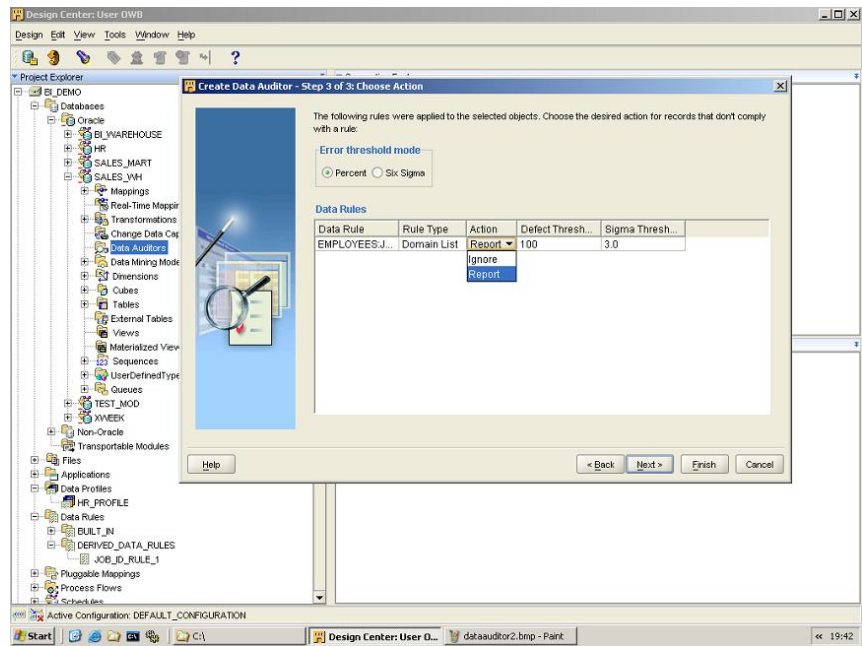


Figure 6. Creating a Data Auditor

Data Quality

The current feature set is improved by utilizing core database technology. The data quality algorithms can now be pipelined to achieve better throughput. New and enhanced matching and merging algorithms ensure you produce better results faster. All of this comes with Warehouse Builder and is available to you within the core ETL components.

Also, non-software improvements have been made by opening up the Warehouse Builder application-programming interfaces (APIs) to specialized data quality vendors, Warehouse Builder allows you to work with your preferred vendor.

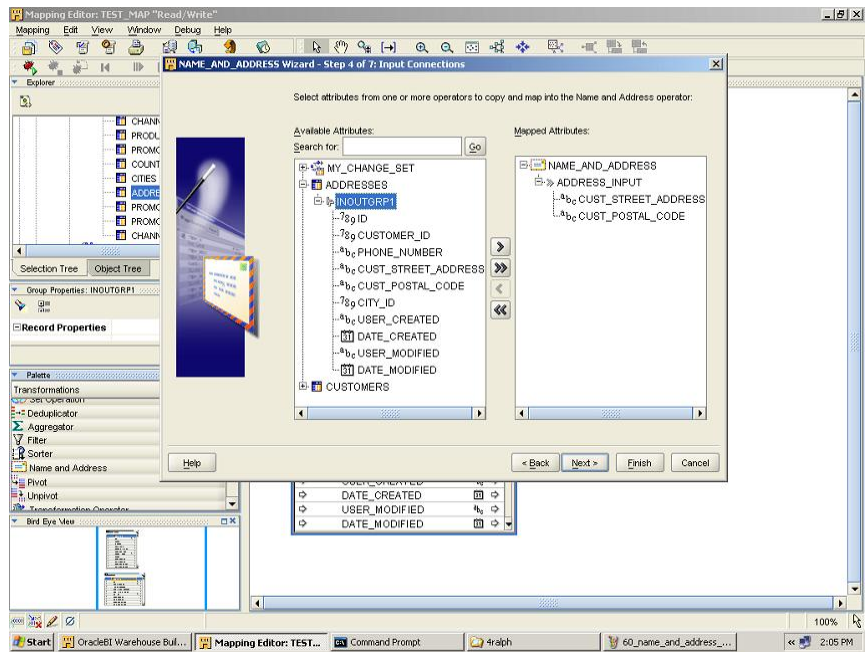


Figure 7. Name and Address Cleansing

While this functionality was introduced in the 9.2 release of the product, the new Warehouse Builder promises even more data quality partners (see [OTN](#)) to work with making data cleansing affordable and easy to do.

WAREHOUSE BUILDER INCREASES INFORMATION VALUE

The second trend is all about enabling advanced analysis on more of your data in a central place. With all the sources that Warehouse Builder can reach into and the advanced data quality features, you can now give more information to your end users.

However, end users now want simple-to-use, high-end analytics to do what-if analysis, to mine their data for previously hidden patterns and apply this knowledge to their business problems.

The new Warehouse Builder will enable some of the hottest technology for you to satisfy this high-end analysis requirement. Direct OLAP target design, advanced calculations (share, ranking and time series) and ETL straight into an Oracle OLAP database are major new components in the capabilities of Warehouse Builder.

Apart from OLAP, Warehouse Builder enables some more advanced capabilities that come with the Oracle platform, like Oracle Spatial and Oracle Data Mining. These technologies allow you to answer questions like where do I make money and why would customers leave for a competitor. In other words, show information based on location and investigate information much deeper to really understand why things are, or might happen.

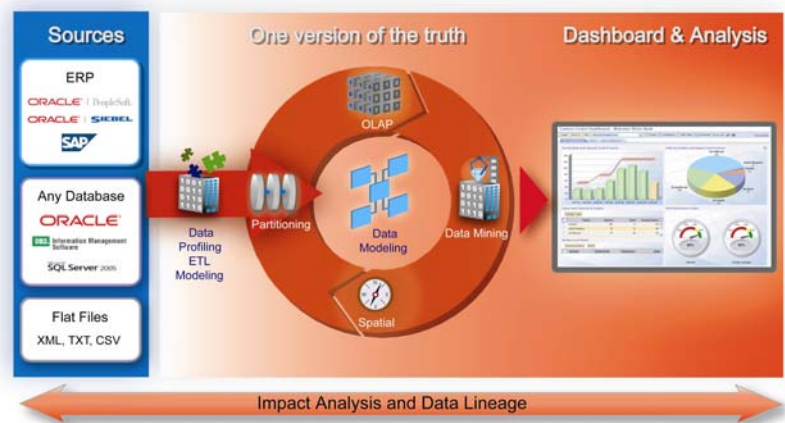


Figure 8 Enabling the Oracle platform for advanced analytics

Once you are utilizing the Oracle platform, you can move into the display world and show maps, gauges and other visualization techniques with Oracle's Business Intelligence products.

Traditional ETL Functionality

The first step in all of this is still traditional ETL technology. We should take a look at the new features for ETL in this new Warehouse Builder release.

The new mapping editor is a key addition to Warehouse Builder (Figure 9).

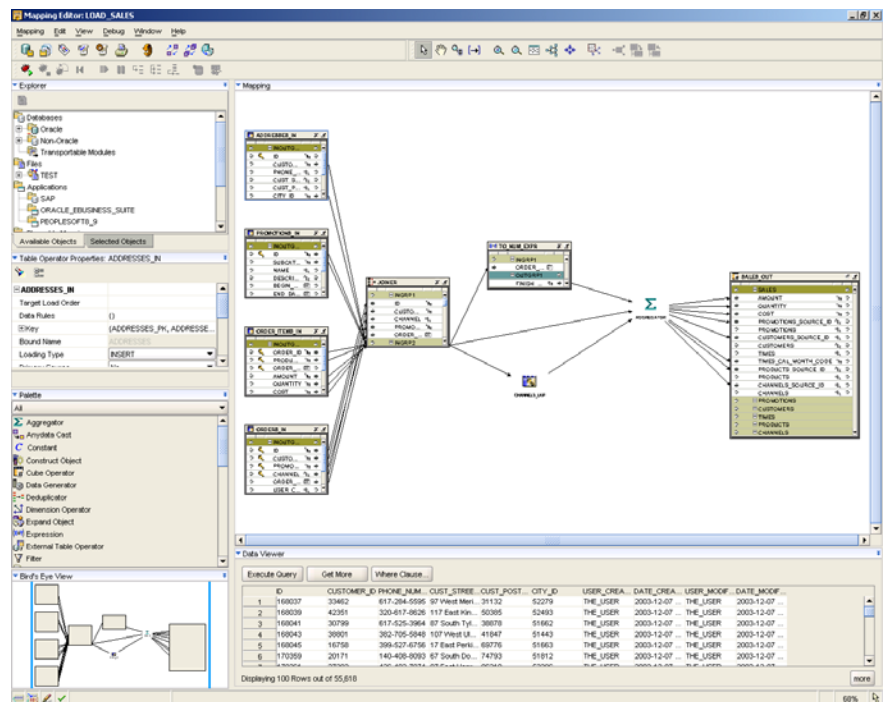


Figure 9. The Mapping Editor in Warehouse Builder

By adding numerous features to the user interface, the new Warehouse Builder makes for a much better user experience in creating large mappings, increasing your productivity, and reducing the number of errors made.

Apart from helping you with visual and productivity enhancements, various new components are added to allow you to create ETL processes faster. You will be able to create reusable mapping components from an existing mapping. Advanced ETL capabilities such as loading slowly changing dimensions or loading data into an on-line analytical processing (OLAP) database are literally just a few clicks away. All in all, the new Warehouse Builder will deliver even more ETL functionality to make you deliver information faster and with higher quality.

Enabling more Data Sources

As organizations implement more and more standard applications, connectivity to these packaged applications becomes more important. The new release of Warehouse Builder adds functionality to deal with this situation by enhancing the already present SAP connector and adding a number of packaged applications connectors to the product.

Capturing Changes

The key to delivering information at the right time is being able to capture relevant information on the source and propagate it to the target for delivery. With the new Warehouse Builder you can utilize Oracle Change Data Capture (CDC) using the latest database techniques.

Through partner solutions² Warehouse Builder allows you to integrate mainframe data, including leveraging CDC applications from those partners.

eBusiness Suite Integration

As part of the push to enable more and more packaged applications, Oracle's eBusiness Suite is obviously the first addition that comes to mind. With the new Warehouse Builder release that integration (outside of the integration Oracle has with the Embedded Data Warehouse) is now added.

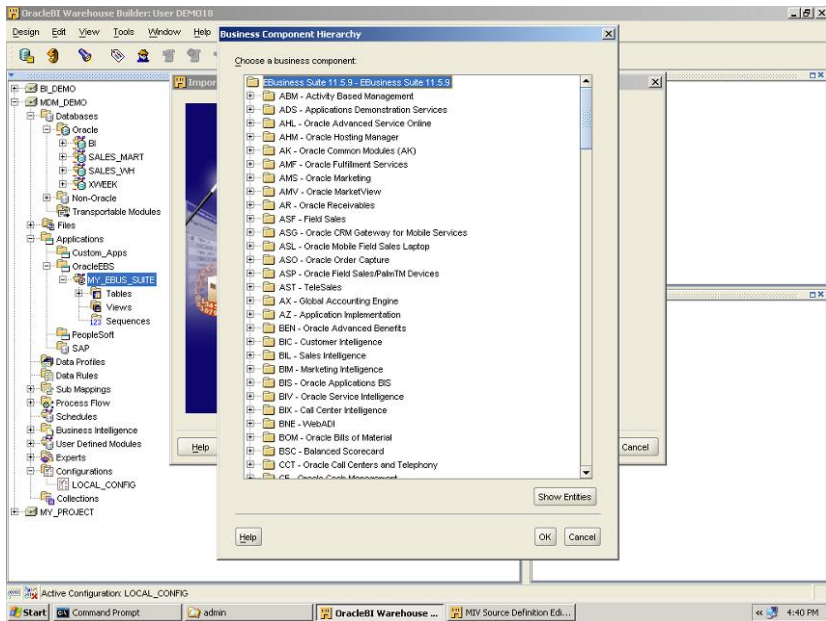


Figure 10 eBusiness Suite Connector

Warehouse Builder now completely understands the metadata in the eBusiness Suite and makes extracting from it much simpler and more efficient.

Another potent feature in the EBS connector is the seamless integration with the Concurrent Manager in the EBS product suite. This functionality uniquely enables tight integration of Warehouse Builder processes in the generic EBS scheduling and job management framework.

PeopleSoft Integration

Next to the eBusiness Suite integration, PeopleSoft's packaged applications are now also added into Warehouse Builder as a standard connector. The PeopleSoft connector allows Warehouse Builder to natively understand the metadata structures of the PeopleSoft applications. Once you have imported the PeopleSoft application metadata you can use all the facilities in Warehouse Builder to integrate this data within the rest of the environment.

² Visit OTN for more information
http://www.oracle.com/technology/products/warehouse/htdocs/otn_partners.html

Siebel Integration

The Siebel connector is very similar to both the EBS and the PeopleSoft connector. For Siebel, the connector understands the structures in the Siebel metadata and will give you the business view of these structures. As with the previously discussed connectors, the Siebel connector is a SQL based connector.

SAP Integration

The SAP Connector is improved for this release. Contrary to most other ETL in Warehouse Builder, which uses SQL and PL/SQL, extraction from SAP is done using ABAP, a native SAP language. This difference ensures that the extraction programs run on the SAP server in a native language.

The new version of Warehouse Builder supports direct deployment and execution of SAP programs on the SAP server. While your SAP staff probably continues to manage programs in production servers, this enhancement greatly improves development and testing efforts by simplifying the deployment process for SAP extraction. Next to this extension a myriad of performance and productivity enhancements have been made to the SAP connector.

The combination of complete functionality, native ABAP code generated and its price point, make the SAP Connector a valuable component in extracting data from your SAP system.

Enabling Advanced Analytics

The biggest benefit of an OLAP solution is the capability to deliver highly advanced calculation capabilities with amazing response times to end users. Now that the Oracle OLAP option is available inside the Oracle database, Warehouse Builder enables this technology for you in an easy-to-use way.

In Warehouse Builder you will create a logical design describing your OLAP cubes in dimensions, hierarchies, measures, calculated measures, and all the components you will need. Then Warehouse Builder uses the new XML API with the OLAP option to create the analytic workspace and the metadata required in the database catalogs. Warehouse Builder also allows you to choose your implementation. Do you want a pure multidimensional OLAP solution (MOLAP) or a relational OLAP solution? With the flip of a switch you can decide or change your mind if testing points you to a different direction.

Once you have created your schema or analytic workspace, you can use the Warehouse Builder ETL modelers to create the load programs. Warehouse Builder is capable of directly loading any data into the analytic workspace, allowing you for the first time to use the wealth of transformation power on OLAP data loads.

Allowing you to load data from any source directly into the arena of high-end analytics of the OLAP option brings high-end analytics to the masses, allowing many people to access and work with OLAP data.

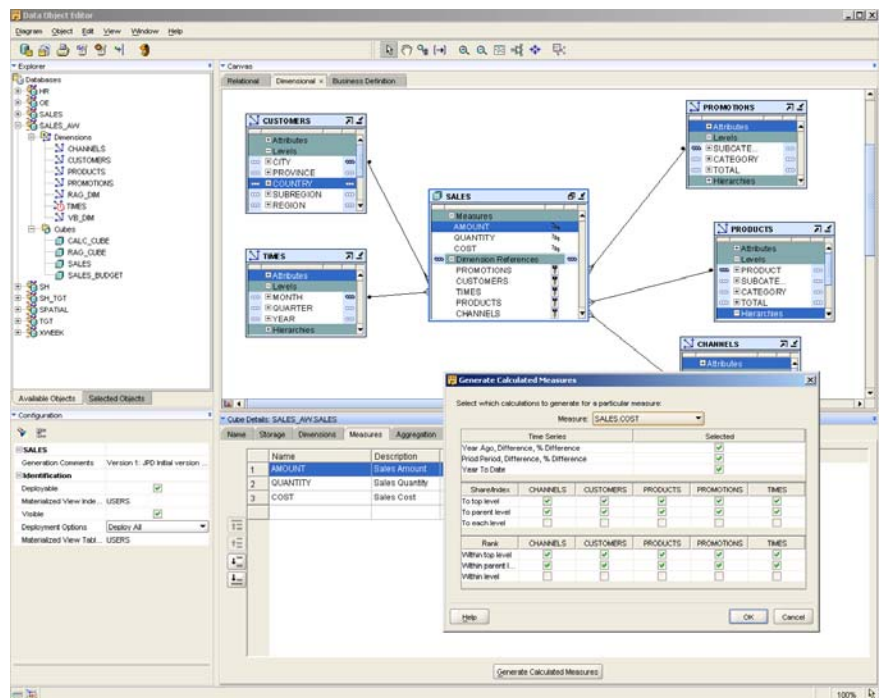


Figure 11 Adding analytic content to a cube

By adding a generation option for advanced analytics directly into the OLAP cube editor, you can now directly prepare the cube for analysis. As these calculated measures are now stored in the database as opposed to your end users creating them, you gain performance and consistency.

Deriving BI Metadata

While you expect Warehouse Builder to deliver a robust infrastructure, you probably are surprised to see that the tool delivers a first cut (template) report for you to show to your end users. By doing so Warehouse Builder saves you a lot of time in your development cycle.

As you define the database (relational or OLAP) and load data into it, you create and collect a lot of metadata. Traditionally, you then build an application on top of this database design essentially duplicating the work of creating metadata and designs all over again.

Warehouse Builder puts a stop to this duplication of effort and metadata because it derives the end user metadata from your design. This not only reduces the work required, it also reduces the risk of errors and the efforts to keep metadata in sync, thus increasing your metadata quality.

Another crucial aspect for metadata quality is that, by having these definitions in one place, you can deliver end-to-end impact analysis. Not only can you determine up front the impact of changes on your database, you will now be able to determine the impact on your BI environment.

So in Warehouse Builder you can select your cube metadata and, for example, derive an Oracle BI Standard Edition (OBI SE) End User Layer, then store this information first in the Warehouse Builder repository. You can then edit the OBI SE objects within Warehouse Builder. Once you are satisfied, you simply deploy this to OBI SE.

WAREHOUSE BUILDER INCREASES PRODUCTIVITY

We have looked at two trends and how Warehouse Builder can assist you in dealing with improving quality and increasing knowledge in your organization. But now we need to deal with shorter cycles to deliver more of this to the organization. Can Warehouse Builder assist with that?

Encapsulating Advanced Functionality

Many tasks in data integration can be encapsulated into standard activities. We already discussed generating advanced analytic calculations (see Figure 11). Rather than having developers or end-users create these calculations in the tedious one-by-one process, we enable you to generate them in a single step.

Slowly Changing Dimensions

The same is done with Slowly Changing Dimensions. The concepts are not very hard to understand – capture changes in crucial information and version it within the data – but since everyone implements it with a slight nuance it makes for hard work in data integration.

In Warehouse Builder slowly changing dimension logic is designed in the actual dimension metadata. The dimension captures all logic that will be applied to the data coming into the dimension.

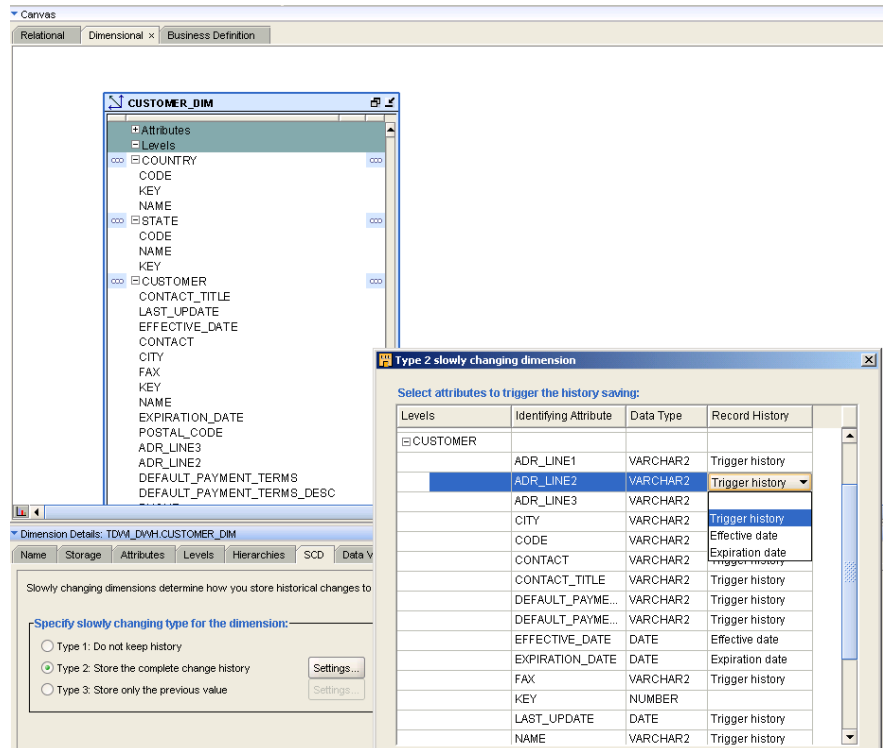


Figure 12 A single implementation for Slowly Changing Dimensions

Once the business user has decided which attributes are important, the data modeler designs the dimension. Within the ETL steps the developer now sees the dimension as any other dimension. This developer does not need to worry about how to handle changes and updates, Warehouse Builder automates that based on the dimension definition. The combination of these steps, design and standardization make the process of effectively dealing with slowly changing dimensions much faster.

Other features include such things as smart storage by using partitioning to improve load and query times, the appropriate definition of indexes on fact tables to enable star query transformation, the ability to generate time dimensions and more.

These examples are just two examples of how advanced functionality can be embedded in simple design constructs. The next productivity booster makes examples like that look rather pale though.

Data Corrections

Imagine you can generate an entire schema based on what your data looks like, have that schema and the mappings generated, and then correct the data errors while it flows into the target schema? Sounds like science fiction? Well think again, because this is now reality with the correction modules in Warehouse Builder.

In a nutshell this is how it works. Based on the results from data profiling (your data), you can derive or design data rules that describe what the data should look like. Then you choose to generate correction mappings and a schema that will act as a target for these correction mappings. When generating the correction mappings, you can choose cleansing methods to ensure the data fits the data rules while you move it into its target schema.

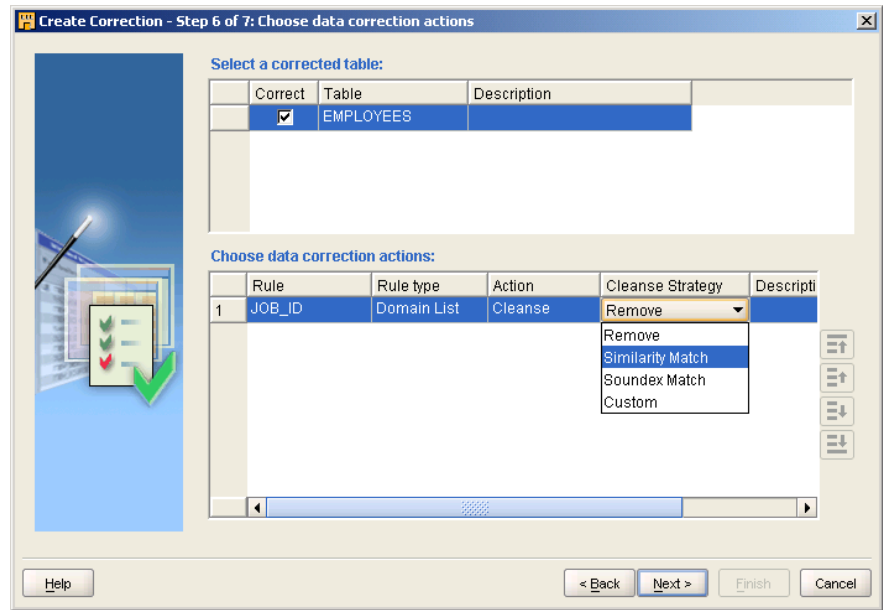


Figure 13 Choosing how to cleanse data

By guiding you through this process, Warehouse Builder speeds up the data integration process tremendously. But since the process is now generated from the data rules, recreating the process when rules change is going to save you even more time. No need to redevelop, simply change the rule’s metadata and regenerate the process and schema and you are up and running again.

This is not merely a productivity gain, this is a productivity turbo charger.

Encapsulating Your Own Business Practices

Where a number of features in Warehouse Builder give you automation or encapsulation of common practices or even allow you to generate data cleansing routines, sometimes you need to find a way to encapsulate your own business rules and practices.

Pluggable Mappings

When you integrating data from various sources, you want to make sure that certain transformations happen in a specific way according to your organization’s specifications. In Oracle you can always write PL/SQL to encapsulate these rules, and then use these programs in Warehouse Builder. Maintaining and changing them however is a nightmare.

With pluggable mappings you can now model transformations or ETL constructs in the same visual and open way as you can regular mappings. Pluggable mappings allow you to define an interface (both input and output) to make the component generic to consuming ETL processes.

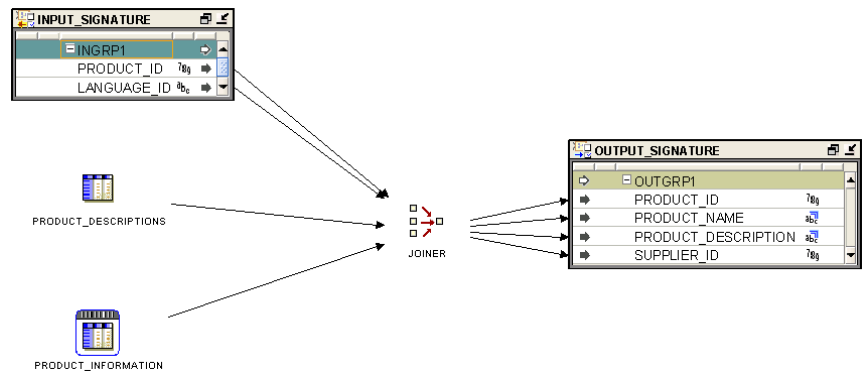


Figure 14 A pluggable mapping with input and output signature

Because these are now components that have a graphical design, they are much easier to understand and maintain. Even if the developer has left the organization, interpreting a logical data flow is much easier than weeding through lines and lines of code, thus saving you time and effort.

Experts

On top of reusing transformation logic, wouldn't you like to repeat certain more complex tasks as you can in for example Microsoft Excel using its macros?

That can be done in Warehouse Builder now with a feature called experts. Experts allow you to program tasks into task flows using both a graphical editor and a scripting language (TCL with OMB³ and OMU extensions). Because Warehouse Builder is a highly graphical tool, experts also allow you to reuse Warehouse Builder user interface components creating your own "wizards".

³ Oracle MetaBase (OMB) and Oracle Metabase User interface (OMU) commands are extensions to the standard TCL language

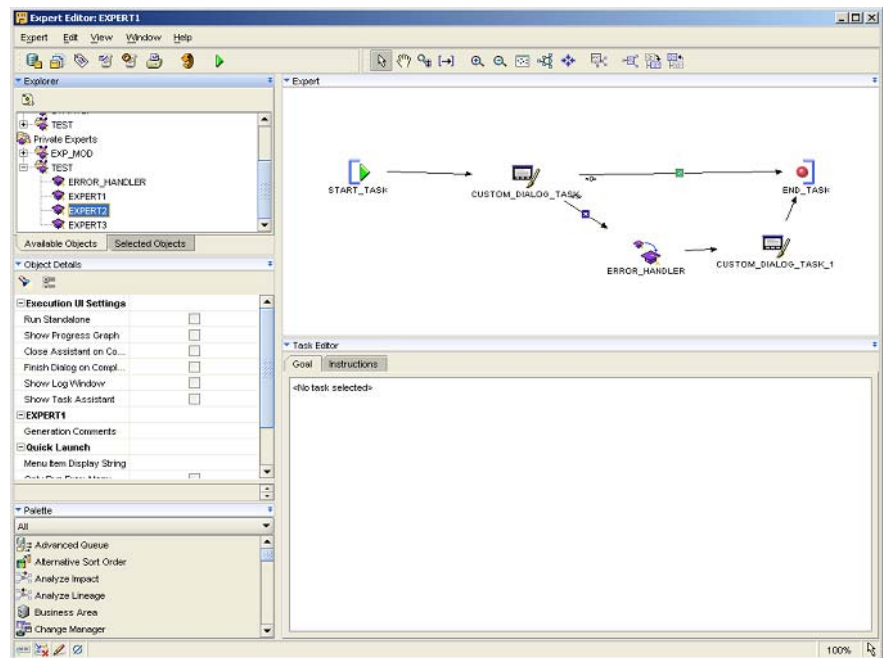


Figure 15 Designing experts in a graphical editor

With this capability you can now achieve a number of objectives. You can allow less experienced developer do jobs that have been defined, via an expert, by more experienced developers, making both more productive.

You can also encapsulate business practices, ranging from simple naming conventions to specifying how certain objects must be designed. Rather than documenting guidelines, you can now embed these guidelines into the product and enforce them. This in turn will reduce costly reviews, audits and more importantly rework and non-compliance.

Should you find the need to create your own user interface components, like simple dialog boxes or more complex shuttles or other custom components, Warehouse Builder allows you to either design custom dialogs in the expert editor or to use java components. In both cases you can interact with the repository metadata using the built in language components.

You can also expose experts to non-Warehouse Builder users because you can run an expert without being logged into the Warehouse Builder user interface. This way you can have more business-oriented users interact with the repository. One example would be a user that must investigate a data anomaly in his OLAP based spreadsheet.

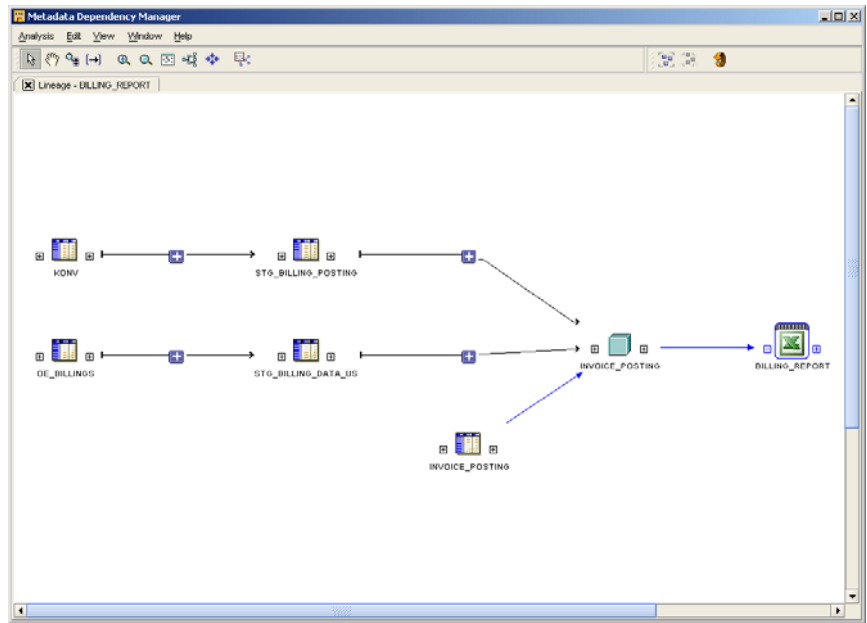


Figure 16 Linage from Excel through to the sources

By integrating the lineage user interface in an expert, and then exposing this via a batch file into Excel, you can have any business user understand the forensics of your data.

SUMMARY AND CONCLUSION

If you look at the trends in information integration you must face the fact that you will have to:

- Deliver information of better quality
- More knowledge derived from existing information
- At shorter intervals and with a shorter time-to-market

To be able to do all of this within the constraint spending limits, you need to look into spending your money wisely. Warehouse Builder 10 Release 2 allows you to deliver on these requirements out of the box for a low price.

Therefore, we can say that Warehouse Builder enables your organization to turn its data into tangible results quicker and with better quality.



Oracle Warehouse Builder 11gR1 - An Overview

March 2007

Author: Jean-Pierre Dijcks

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
www.oracle.com

Copyright © 2007, Oracle. 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 is a registered trademark of Oracle Corporation and/or its affiliates. Other names may be trademarks of their respective owners.