

Features des Oracle Warehouse Builder 11.2

Franz von Sales Hohenberg
ORDIX AG
Wiesbaden

Schlüsselworte:

OWB, Repository, Location, Module, Mappings, Workspace, Rechte, Design Center, Control Center, Konfigurationen, Quellsysteme, Targets, Operatoren, owbsys, Workflow, Experts, Speicherorte, Projekte, Code Template Mappings, CDC

Einleitung

Der Oracle Warehouse Builder hat sich durch die Verbesserungen der letzten Versionen zu einem voll qualifizierten ETL-Tool entwickelt. Durch die Umstellung auf Workspaces hat Oracle die Möglichkeit geschaffen, die Rechtestruktur zu vereinfachen. Die einfache Aufteilung und seine komplexen Verarbeitungsmöglichkeiten, die von einfachen Flatfiles bis zu komplexen Verbindungen fremder Systeme reichen, lassen den Oracle Warehouse Builder in die High Performance Ebene aufsteigen.

Installation

Die Installation des Oracle Warehouse Builder gestaltet sich, ob unter Windows oder Unix, relativ einfach. Der schwierigste Teil ist das Sammeln von Informationen, die man vor der Installation bereithalten soll:

- Welche Benutzer sollen Zugriff auf das Respository haben?
- Wie viele Projekte möchten Sie anlegen?
- Sollten alle Projekte eigene Workspaces erhalten?
- Benutzen Sie Entwicklungs-Repositorys?
- Haben Sie ein Deployment-Verfahren?
- Wie werden Änderungen in andere Repositorys übertragen?
- OMBPLUS

Location / Speicherort

Speicherorte stellen die erste wichtige Grundfunktion im OWB da. Sie müssen sorgfältig gewählt und eingerichtet werden. In diesen Locations werden die Verbindungsinformationen hinterlegt, in denen generell ein Location-Name angegeben werden muss, z.B.:

ORA_UIDMAP01_SDBC

Je ausführlicher der Location-Name beschrieben wird, desto einfacher gestaltet sich später die Verwendung. Innerhalb der Location werden unter anderem benötigt:

- Benutzername für den Zugriff auf ein Schema
- Kennwort des Schemas
- Physikalischer Rechnername
- Datenbankname

- Schemaname
- Version

Um das Error Handling zu vereinfachen, werden sogenannte Präfixe automatisch erstellt oder bei der Installation oder später angegeben. Je nach Betriebssystem können die Masken zum Erstellen der Location variieren. Die Grundeinstellungen bleiben aber gleich.

File Locations (Speicherorte)

Um auf Flatfiles zugreifen zu können, werden unter Dateien / Files sogenannte File Locations angelegt. Diese zeigen nur auf das lokale System, da hier keine Verbindungsinformationen benötigt werden. Beachtet werden nur die Dateisystemrechte auf dem lokalen System.

Projekte

Projekte sind das Herzstück des Oracle Warehouse Builder und die Grundlage zum Trennen von Firmeninternen Prozessen. Prozesse auf Projektebene zu trennen hat keine Auswirkungen auf die Location. Sie können eine Location in mehreren Projekten benutzen. Der Zugriff aber aus mehreren Projekten auf eine Location zur gleichen Zeit, sollte getestet werden.

Datenbanken

Das Festlegen vom Quell - oder Targetsystem erfolgt unter den Datenbanken des jeweiligen Herstellers. Der Namen des Moduls sollte ebenfalls an dieser Stelle im Voraus festgelegt werden. Zudem wird hier definiert, um welche Art von Zugriff es sich handelt:

- Entwicklung
- Qualitätssicherung
- Produktion

Das einfache System des OWB kommt an dieser Stelle zur Geltung. Es wird hier eine Location eingebunden, die im ersten Schritt definiert wurde. Dass die richtige Location eingebunden wird, liegt bei dem Entwickler. Änderungen der Location werden nur an zentraler Stelle durchgeführt. Das Definieren eines Moduls ist gleichzusetzen mit dem Entwickeln einer eigenständigen Umgebung. Sie laden dann die Oracle Objekte aus der Datenbank in das Repository, die benötigt werden:

- Views
- Tabellen
- Externe Tables
- Materialized Views
- Sequences
- Benutzerdefinierte Typen

Jederzeit lassen sich auch eigenständige Objekte im OWB entwickeln und später in die Datenbank per Deployment einspielen. Es können pro Projekt mehrere Quell- und Target-Module angelegt werden. Die Übersichtlichkeit der Namensgebung und der einfachen Bedienung liegt im Auge des Betrachters.

Control Center

Kontrolle ist alles, Vertrauen ist nichts. Damit Sie nicht vor einer Entscheidung stehen, wurde unter OWB das Control Center eingeführt. Das Betreiben unter einem anderen User als bei dem Design

Center ist unter rechtlichen Gesichtspunkten und der Vorgabe des Kunden zu betrachten. Wenn Sie das Design Center öffnen können wird unter:

Extras / Control Center Manager

eine Umgebung geöffnet, die es erlaubt Objekte aus dem OWB auf die HDD des lokalen Rechners oder je nach Control Center auf jedes beliebige Repository im Netzwerk zu deployen.

Beim Deployment sollte grundsätzlich die Reihenfolge der Objekte beachtet werden. Damit unnötige Fehlermeldungen unterbleiben. Das Betrachten aller oder nur geänderter Objekte ist als Auswahlmenü vorhanden.

Objekte werden hier zur Übertragung auf andere Systeme deployed. In der Praxis hat sich herausgestellt, dass bestimmte Deployment-Verfahren für das Übertragen von Tabellen, Views usw. vorgefertigte Scripte verwenden und das Deployment des OWB auf das Mapping beziehen.

Jedes Projekt, das über ein Control Center deployed werden soll, muss auch im Control Center registriert werden. Zudem werden die Locations registriert. Dadurch wird verhindert, dass unbedachte Handlungen und Änderungen an den Locations zu Fehlern führen. Damit Änderungen in den Locations nach einer Registrierung durchgeführt werden können, muss die Registrierung aufgehoben, die Änderung durchgeführt und eine neue Registrierung abgeschlossen werden.

Konfigurationen

Konfigurationen dienen dazu, den OWB in kürzester Zeit auf ein anderes Repository umzustellen. Wie bereits beschrieben, werden im Control Center die Locations registriert. Eine Konfiguration wird einem Control Center zugewiesen. Damit kann eine Konfiguration als Default ausgewiesen werden, das automatisch ein Control Center angehängt hat.

Damit lassen sich Einstellungen (Experts-Module) anwenden, um zum Beispiel mehrere Dimension Mappings auf Bulk Insert umzustellen. Sie können auch das Verhalten der Fehler im Mapping-Lauf definieren und zwar für alle Mappings gleichzeitig.

Die Möglichkeiten die vom OWB hier geboten werden sind immer mit anderen Features kombiniert.

Datenprofile

Das Erstellen von Datenprofilen für Tabellen wird meistens unterschätzt. Das grafische Handling und der saubere Überblick täuschen etwas vor, was eine sehr komplexe und umfangreiche Anwendung ist. Das 100 %-Wissen über die Datenstruktur, dessen Lieferumfang und Qualität sind Voraussetzung für das korrekte Einrichten und Verarbeitung beim Übertragen.

Die bearbeitenden Reiter

- Datenprofil
- Profilojekt
- Aggregation
- Datentyp
- Datenregel
- Referenziell
- Funktionale Abhängigkeiten
- Eindeutiger Schlüssel
- Muster

bedürfen einer Vorplanung und lassen sich nicht einfach umsetzen.

Code Template Mappings

Neu im Oracle Warehouse Builder 11R2 sind die sogenannten Code Templates und Template Mappings. Die Mappings basieren auf der ODI-Technologie (Knowledge-Module). Die Template Mappings können analog zu den klassischen Mappings erstellt und eingesetzt werden. Der wichtigste Hinweis ist, dass diese in einer Java-Laufzeitumgebung ausgeführt werden. Zurzeit stehen vier Arten von Code Templates zur Verfügung:

- Load (Extrahieren von Daten)
- Integration (Schreiben von Daten)
- Control (Errorlogging)
- Change Data Capture

Template Mappings werden unabhängig von den vorhandenen Datenbankmodulen in einem eigenen Ordner definiert. Der Editor zum Erstellen der Code Template Mappings kommt dem Expert Editor sehr nahe. In dem Editor werden verschiedene Tasks angeboten:

- JAACL (Java Command Language, Java TCL Interpreter)
- JDBC (Führt SQL-Befehle aus)
- Jython (fr. JPython, Java-Implementierung von Python)
- BS/OS (Aufruf von Betriebssystembefehlen)
- Runtime API

Die OWB Code Templates können für folgende Datenbanken verwendet werden:

- Oracle
- DB2
- Microsoft SQL Server

Die Vorteile der neuen Code Template Mappings liegen klar auf der Hand. CDC kann komplett aus dem OWB verwaltet werden. Das Laden von heterogenen Daten aus unterschiedlichen Datenquellen in die Staging Area ist eine Aufgabe der Template Mappings. Die klassischen Mappings sind zur Weiterverarbeitung in das Data Warehouse gedacht.

Kontaktadresse:

Franz von Sales Hohenberg
ORDIX AG
Westernmauer 12-16
D-33098 Paderborn

Telefon: +49 (0) 5251 / 1063-0
Fax: +49 (0) 1801 / 67439 - 0
E-Mail: info@ordix.de
Internet: www.ordix.de