

Oracle GoldenGate: die Brücke zwischen Oracle und MySQL Datenbanken

Ileana Someşan
Oracle Deutschland B.V. & Co. KG
Potsdam

Schlüsselworte

Oracle GoldenGate, Oracle Datenbank, MySQL, Replikation, Synchronisation

Einleitung

In diesem Artikel wird Oracle GoldenGate im Umfeld von Oracle und MySQL Datenbanken vorgestellt. Mit GoldenGate kann man Oracle und MySQL Datenbanken nebeneinander betreiben und die zwei Datenbestände zu jedem Zeitpunkt auf dem gleichen Stand halten. Die Struktur der Daten darf sich unterscheiden; das ist immer der Fall, wenn verschiedene Clients auf Oracle und MySQL zugreifen. Nach einer kurzen Einführung in GoldenGate wird die unidirektionale Replikation von Oracle nach MySQL näher erläutert.

Problemstellung

Oracle und MySQL Datenbanken werden häufig im selben Unternehmen eingesetzt, wennauch für verschiedene Anwendungen. In diesem Fall können schnell die Anforderungen entstehen, dass ein Teil der Oracle Daten auch für MySQL-Clients zur Verfügung steht und dass eine Aktualisierung in der Oracle Datenbank für MySQL-Clients sofort sichtbar ist.

Dazu ein Beispiel: ein Lebensmittel-Großhändler pflegt seinen zentralen Produktkatalog, mitsamt den Preisen, in einer Oracle Datenbank. Die Produktpreise werden häufig angepasst, manchmal mehrfach täglich. Der Großhändler bietet einen kleinen Teil seines Sortiments auch online. Der Online Shop wird gehostet und basiert auf einer MySQL Datenbank. Die online Preise werden anhand der zentralen Preise berechnet und müssen bei jeder Änderung des Oracle Produktkatalogs sofort aktualisiert werden.

Es ist sicherlich möglich, seine eigene Synchronisationslösung zu implementieren. Der einfachste Ansatz basiert auf Export/Import-Jobs: die relevanten Daten werden regelmäßig aus der Oracle Datenbank exportiert und in MySQL geladen. Das Problem dabei ist, dass die zwei Datenbestände nicht zu jedem Zeitpunkt synchron sind. Export/Import-Jobs laufen zu festgelegten Zeitpunkten, oft einmal am Tag, während der Wartungsfenster. Eine häufigere Ausführung ist in der Regel nicht möglich, weil Export-Läufe einen Overhead verursachen und den Produktivbetrieb beeinträchtigen. Ein weiteres Problem ist, dass durch den Export große Datenmengen entstehen. Es werden immer alle Daten exportiert, unabhängig ob sie seit der letzten Ausführung verändert wurden oder nicht. Die Daten müssen auf das Zielsystem gebracht werden, was neue Fragen nach der Netzwerkbandbreite oder nach gemeinsamen Storage aufwirft.

Verteilte Transaktionen sind ein anderer Ansatz um Oracle und MySQL Datenbank auf demselben Stand zu halten. Bei dieser Methode werden sowohl die Oracle Datenbank als auch MySQL innerhalb einer Transaktion, aktualisiert. Somit sind die Daten zu jedem Zeitpunkt synchron. Oracle bietet den *Database Gateway for ODBC* (früher *Heterogenous*

Services for Oracle Database) zur Implementierung von verteilten Transaktionen zwischen einer Oracle und einer ODBC-basierten Datenbank, z.B. MySQL. Die Schwierigkeit bei diesem Ansatz besteht darin, dass die Oracle Endanwendung angepasst werden muss. Bei einer Anwendung von Fremdanbietern ist eine Anpassung ohne Weiteres nicht erlaubt und nicht möglich.

Die größte Herausforderung bei der Implementierung einer eigenen Lösung zur Datensynchronisation zwischen Oracle und MySQL ist der Umgang mit Unterschiedlichkeit (Heterogenität). Oracle und MySQL nutzen verschiedene Datentypen und „sprechen“ verschiedene SQL-Dialekte. Meistens haben die Daten unterschiedliche Datenstrukturen (z.B. Spaltenanzahl, -namen), aufgrund der darüber liegenden Anwendungen. Es ist aufwändig und fehleranfällig diese Unterschiede manuell aufzulösen. Und warum sollte man das tun?

Überblick Oracle GoldenGate

Oracle GoldenGate kam zu Oracle nach einer Aquisition im Jahr 2009. Seitdem wird es als selbständiges Produkt geführt und ist unabhängig von anderen Oracle Produkten, wie der Oracle Datenbank. GoldenGate spielt in Oracles Portfolio eine wichtige Rolle im Bereich Datenverteilung und -integration in Echtzeit. Eine Stärke von GoldenGate ist die Fähigkeit zwischen fast allen gängigen Datenbanksystemen zu replizieren. Das sind proprietäre Datenbanken, wie die Oracle Datenbank, IBM DB2, SAP Sybase, Microsoft SQL Server, HP NonStop, TimesTen oder Teradata, aber auch Open Source Datenbanken wie MySQL und Postgres. Die komplette Liste aller unterstützten Kombinationen aus Datenbanksystem/-version, und Betriebssystem/-version finden Sie unter [2].

GoldenGate kann für zwei Zwecke eingesetzt werden. Ein Einsatzgebiet ist das sogenannte Initial Load. Dabei wird eine leere Datenbank mit den Daten einer anderen Datenbank befüllt. Diese Operation kommt z.B. bei Datenbankmigrationen vor. GoldenGate kann alternativ zu anderen Migrationswerkzeugen verwendet werden.

Das Haupteinsatzgebiet ist die Synchronisation zwischen Datenbanken, nachdem das initiale Laden stattgefunden hat. GoldenGate überwacht eine Datenbank im Hinblick auf Änderungen und aktualisiert entsprechend eine andere Datenbank. Typische Änderungen sind Datenänderungen: neuen Datensatz einfügen, löschen oder aktualisieren. Für manche Datenbanken kann GoldenGate auch Änderung an der Datenstruktur erfassen, z.B. dass eine neue Spalte zu einer bestehenden Tabelle hinzugefügt wurde [3]. Man kann die Datenmenge auf bestimmte Datensätze, Tabellen, oder eine Teilmenge der Spalten einschränken. Weiterhin kann man zwischen Tabellen mit unterschiedlicher Struktur replizieren, z.B. Tabelle A mit den Spalten (m, n, p) und Tabelle B mit den Spalten (u,v).

GoldenGate besteht aus mehreren Komponenten, siehe Abb. 1. Einige davon laufen auf dem Quellsystem, andere auf dem Zielsystem. Der Extract Prozess nimmt Änderungen auf und schreibt sie in datenbankunabhängige Dateien, den sog. Trail Files. Der Data Pump Prozess (hat nichts zu tun mit der Data Pump Utility) verschickt die Änderungsdaten an das Zielsystem, optional verschlüsselt und komprimiert. Auf dem Zielsystem läuft eine andere Komponente von GoldenGate, der Replicat Prozess. Er erzeugt SQL-Aufrufe im Dialekt der Zieldatenbank und lässt sie ausführen.

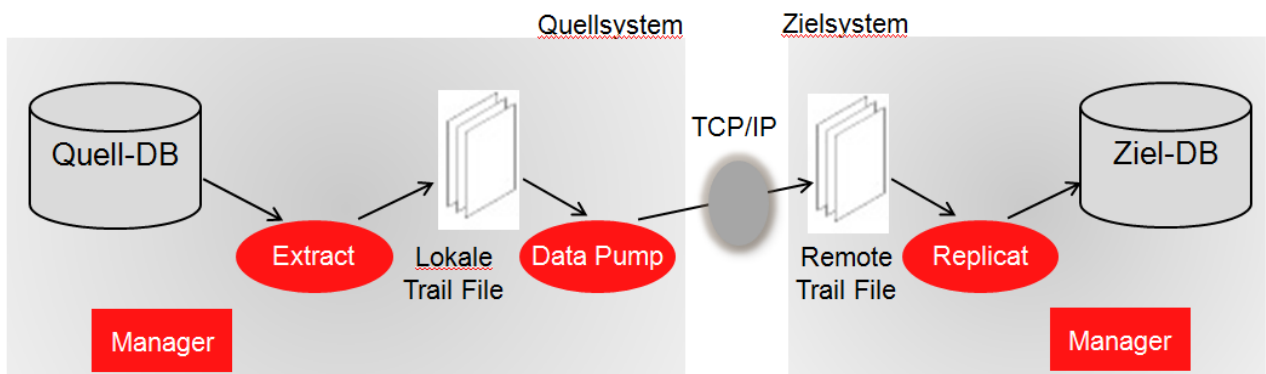


Abb. 1: Die wichtigsten Oracle GoldenGate Komponenten

Unidirektionale Replikation von Oracle nach MySQL

In diesem Abschnitt betrachten wir den Ablauf einer Replikation mit der Oracle Datenbank als Quelle und MySQL als Ziel.

Zunächst sollten Sie prüfen, ob GoldenGate Ihre Datenbankversionen unterstützt. Die aktuelle GoldenGate Version (11.2.1) unterstützt Oracle Datenbanken ab der Version 10.2.0.4 und MySQL ab Version 5.0 siehe [2]. Als nächstes sollte Sie sich vergewissern, dass all Datentypen, die in Ihrer Oracle und MySQL Datenbank verwendet werden, von GoldenGate repliziert werden können [4], [5]. Aktuell werden fast alle Datentypen bis auf wenigen Ausnahmen, wie z.B. ANYTYPE und ROWID in Oracle oder GEOMETRY und SET in MySQL, unterstützt. Sie können auf eine MySQL Datenbank alle Datenänderungen durch DML-Operationen (Insert, Update, Delete), jedoch keine DDL-Operationen, also Änderungen an der Struktur der Daten, replizieren.

Sie benötigen zwei Binaries von GoldenGate: *Oracle GoldenGate for Oracle* und *Oracle GoldenGate for MySQL* [1] und führen auch zwei Installationen durch. Eine Installation erfolgt typischerweise direkt auf dem Server, auf dem die Oracle Datenbank läuft, siehe Abb. 2. Wenn auf dem Produktivserver aus Compliance-Gründen keine weiteren Softwarekomponenten zugelassen sind, können Sie GoldenGate auch auf einem anderen Server installieren, auf dem Sie eine Standby Oracle Datenbank betreiben (z.B. mit DataGuard). Wichtig ist, dass GoldenGate auf die Redo Logs der Datenbank zugreifen kann. Die zweite Installation erfolgt auf dem MySQL Server.

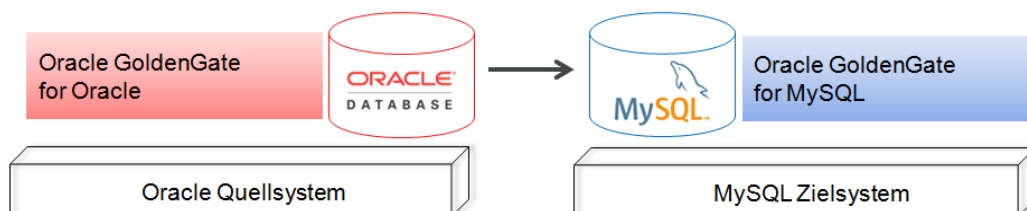


Abb. 2: Zwei Installationen von Oracle GoldenGate

Der Extract Prozess liest die Redo Logs der Oracle Datenbank entweder direkt, wenn Sie Extract im sog. *Classic Capture Mode* konfigurieren oder indirekt über einen Log Mining Server im *Integrated Capture Mode*. Letzteres ist erst ab GoldenGate 11.2.1 und Oracle Datenbank 11.2.0.3 mit entsprechendem Patch verfügbar. Der Integrated Capture Modus bietet bessere Integration mit der Oracle Datenbank, wenn es um erweiterte Funktionalität wie Real Application Clusters (RAC), Automatic Storage Management (ASM), Verschlüsselung oder OLTP-Komprimierung geht.

Welche Vorbereitungsschritte sind für die Oracle Datenbank notwendig?

Es ist empfohlen, die Oracle Datenbank im Archivelog Modus zu betreiben. So kann GoldenGate auf die Archived Redo Logs zugreifen, falls die Online Redo Logs nicht mehr verfügbar sind. Die Tabellen in der Oracle Datenbank müssen über Primärschlüssel oder wenigstens über Spalten mit eindeutigen Werten verfügen (UNIQUE Constraint). Ansonsten sollten Sie über GoldenGate einen Ersatzschlüssel definieren (KEYCOLS). Weiterhin müssen Sie in der Oracle Datenbank das Supplemental Logging aktivieren. Darüber werden zusätzliche Informationen in die Redo Logs hinein geschrieben, die GoldenGate in die Lage versetzen, Änderungen an Oracle Datensätzen den richtigen MySQL Datensätzen zuzuordnen. Bei der Replikation zwischen zwei verschiedenen Datenbanktypen muss auf dem Quellserver die GoldenGate DEFGEN Utility ausgeführt werden. DEFGEN beschreibt die Tabellen der Oracle Datenbank in einem generischen Format. Sie müssen die resultierende Definitionsdatei auf das Zielsystem übertragen, damit GoldenGate die Konvertierung zwischen Oracle und MySQL Tabellen vornehmen kann.

Welche Vorbereitungsschritte müssen Sie auf dem MySQL Server durchführen?

Sie sollten die Zieltabellen in der InnoDB Storage Engine speichern. Anderenfalls ist bei der Replikation die Transaktionsintegrität nicht gewährleistet. InnoDB ist die standard Speicher-Engine seit MySQL 5.5. Auch in MySQL sollten Sie für die Eindeutigkeit der Datensätze sorgen. Trigger und kaskadierende Update und Delete Constraints müssen auf dem Zielsystem deaktiviert werden. Es wird empfohlen, in MySQL eine Tabelle anzulegen, in der GoldenGate seine Checkpoints speichert.

Die Datenreplikation wird in zwei Schritten konfiguriert. Im ersten Schritt wird die MySQL Datenbank mit Oracle Daten befüllt, um einen initialen, gleichen Stand der Daten zu erzeugen. Das Initial Load kann mit GoldenGate oder mit anderen Ladewerkzeugen durchgeführt werden. Im zweiten Schritt kann die Synchronisation gestartet werden.

Die Initial Load Konfiguration ist in Abb. 3 gezeigt. Sie müssen lediglich einen Extract und einen Replicat Prozess konfigurieren und starten. Der Extract Prozess liest die Oracle Daten direkt aus den Tabellen aus. Es werden keine Redo Logs benötigt.

Für die Synchronisation konfigurieren Sie drei Prozesse Extract, Data Pump und Replicat und registrieren lokale und remote Trail Files, siehe Abb. 4. Die Konfiguration der Prozesse kann sehr einfach über die Kommandozeilenschnittstelle von GoldenGate, dem sog. GGSCI-Client, erfolgen.

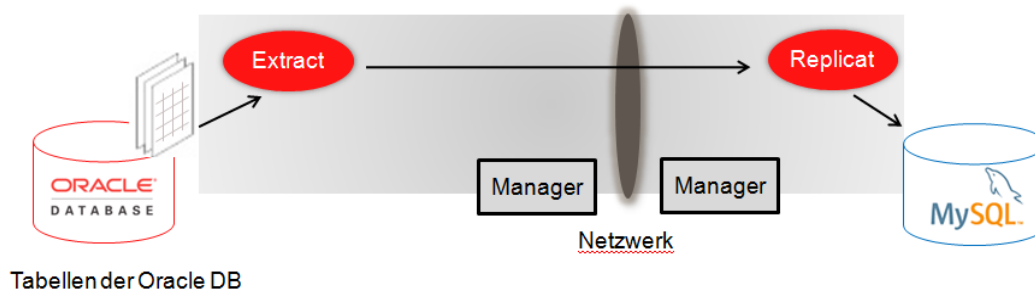


Abb. 3: GoldenGate Konfiguration für Initial Load

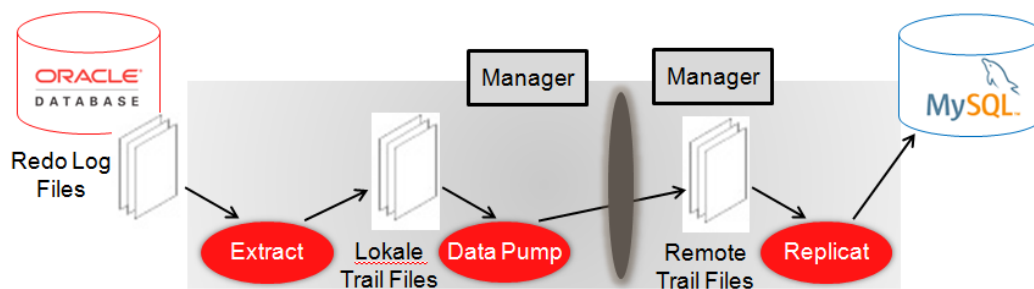


Abb. 4: GoldenGate Konfiguration für die Synchronisation

Schlussbemerkungen

Die Lizenzierung von GoldenGate ist unabhängig von Datenbanklizenzen. Das ist auch der Grund, warum GoldenGate mit jeder Edition der Oracle Datenbank (Standard Edition One, Standard Edition oder Enterprise Edition) eingesetzt werden kann. Im beschriebenen Fall der Replikation zwischen Oracle und MySQL müssen zwei Produkte lizenziert werden: *Oracle GoldenGate* und *Oracle GoldenGate for Non Oracle Database*. Es wird nach standard Oracle Metriken (Prozessor oder Named User Plus) lizenziert, wobei bei der Lizenzierung nach Benutzern die Enduser der beteiligten Datenbanken zählen.

GoldenGate ist ein mächtiges Produkt. In diesem Artikel wurde ein mögliches Replikationsszenario, von Oracle nach MySQL, gezeigt. GoldenGate ermöglicht auch die Replikation in die andere Richtung, von MySQL nach Oracle. Bidirektionale Konfigurationen lassen sich ebenfalls implementieren und Oracle GoldenGate bietet dafür eingebaute Methoden zur Konflikterkennung und -behebung, oder die Möglichkeit Konflikte über selbst implementierte Routinen (User Exits) aufzulösen.

Quellen und weitere Informationen

[1] Download Oracle GoldenGate von edelivery.oracle.com

[2] Von Oracle GoldenGate unterstützte Datenbank-/Plattformkombinationen

<http://www.oracle.com/technetwork/middleware/ias/downloads/fusion-certification-100350.html>

[3] Funktionsumfang abhängig vom Datenbanktyp, siehe Oracle GoldenGate *Installation and Setup Guide* für die jeweilige Datenbank

[4] Oracle GoldenGate Oracle Installation and Setup Guide

[5] Oracle GoldenGate MySQL Installation and Setup Guide

Kontaktadresse:

Ileana Someșan

Oracle B.V. & Co. KG

Schiffbauergasse 14

D-14467 Potsdam

Telefon: +49 (0) 331-2007 165

E-Mail ileana.somesan@oracle.com

Internet: www.oracle.com