Dell Boomi: Efficient Database Integration using Dell Boomi

A Dell Technical White Paper

Dell | ITDS
Ganesh Kumar
Efficient Database Integration using Dell Boomi™

Contents

Introduction .......................................................................................................................... 2
Assumption .......................................................................................................................... 2
Cloud based integration ....................................................................................................... 2
Data Integration between SqlServer and Oracle ................................................................. 2
Data Integration using Boomi Atomsphere ......................................................................... 3
Scheduling Data Integration Job ........................................................................................ 7
Summary ............................................................................................................................. 7
Addendum ........................................................................................................................... 8

Figures

Figure 1. Integration between SqlServer and Oracle ............................................................. 2
Figure 2. Data Integration Process ...................................................................................... 3
Figure 3. Source Database Connection ............................................................................... 4
Figure 4. Select SqlServer Profile using Stored procedure ................................................. 5
Figure 5. SqlServer Map (Data Transformation) ................................................................. 5
Figure 6. Destination Database Connection ....................................................................... 6
Figure 7. Insert Oracle Profile ............................................................................................ 6
Figure 8. Scheduling Data integration Job.......................................................................... 7
Introduction
It is known that Boomi Atomsphere is the leading SaaS Cloud integration platform of choice. Due to its simplicity and versatility, it is also emerging as a leading platform to integrate data between heterogeneous applications or databases (ex: MySql or Sql Server to Oracle).

Traditional ETL tools are too cumbersome to exchange or transfer nominal or limited set of data between heterogeneous databases. It is not efficient to deploy a replication based solution like GoldenGate or Shareplex to transfer nominal data. Relying on vendor tools involves installing and configuring Heterogeneous Gateway software on Oracle Database Servers and managing them. As the number of applications which require such heterogeneous database integration increases, it will be cumbersome to manage all such database servers with custom installations for different gateway products. Do we have a better way to integrate our databases?

Assumptions
We assume that the readers have knowledge about various database technologies like

- Oracle, PL/sql programming
- Sql Server, T-Sql, Stored procedure
- Basic knowledge using Boomi Atompshere or other integration or ETL tools.

Cloud based integration
One of the key factors for efficiency is to standardize the database practices, and this includes the Oracle installation amongst other best practices. It will be very efficient if DBAs don’t have to worry about such non-standard deployments like Gateway installs etc. It will be best to entirely avoid software installation, configuration and instead focus on designing and developing the integrations. The key is to take advantage of Boomi Atomsphere, which is a Cloud based integration Service, which offers many advantages and productivity gains.

Integration between Sql Server and Oracle
Its so simple to design an integration using Boomi Atompshere due to the revolutionary visual interface and the availability of a wide variety of connectors and components. This Whitepaper focuses on how to design a repeatable integration to periodically transfer data from a Sql Server database to an Oracle Database.
To design a repeatable data integration, we need to store certain information in a table which includes columns like

- owner,
- table_name,
- last_extracted_date (Date/Time when the data was extracted),
- incremental_hours
- next_date which can be computed by adding the incremental hours to Last_extracted_date. For example, the last_extracted_date is: '01-jan-2012 00:00:00' and Incremental_hours is 6, then, the next_date will be '01-jan-2012 06:00:00'. So, effectively, this pulls any data for those 6 hours on 1st Jan 2012. After successfully extracting the data, we will have to update last_extracted_date as '01-jan-2012 06:00:00'.

The below database objects are needed to support the integration

- Control table which stores information like Owner, TableName, Last_extracted_date, Incremental_hours etc
- History table which stores all executions like Owner, TableName, Last_extracted_date, NumRowsSource, Incremental_Hours etc
- A stored procedure which extracts (selects) data from source database between "Last_extracted_date" to "Last_extracted_date + Incremental_hours".

Data Integration using Boomi Atomsphere

Data Integration in 3 easy steps.

1. Define source database connection and stored procedure to extract data
2. Define Data transformation map between the source and destination table
3. Define destination database connection and Insert statement to load data

The below is a picture of the actual data integration process. Let's look at these in detail.

1. Define source database connection and stored procedure to extract data
   Drag and drop the connector icon into canvas and enter the source (sqlServer) database connection detials as shown below.
For “Connection”, Enter the Database Type as MSSQL, and input information like host, port, database name etc.

For “Operation”, we would select the “Type” as “Stored Procedure Read” and input the name of the stored procedure “get_data” as in this example. The Stored procedure returns 5 columns as shown below. This data will have to be loaded into the destination (Oracle) database table.

- S_NUM
- CURR_DATE
- STRING1
- DATE_ARCHIVED
- SEQNO
2. SqlServer Map (Data Transformation)
In this step, we would map the columns from source (Sql server) and the destination (Oracle) database. This component has a variety of string or date functions to manipulate the input data and map them to destination columns. Since the column names between the source and destination tables are the same, this will be a straight column to column mapping.

Left side (Source) SqlServer Database ----------------- Right side (Destination) Oracle Database
3. Define Destination database connection and the Insert statement

Drag and drop the connector icon into the canvas and enter the destination (Oracle) database connection.

**Connector Action**

<table>
<thead>
<tr>
<th>General</th>
<th>Parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td>Connector:</td>
<td>Database</td>
</tr>
<tr>
<td>Action:</td>
<td>Send</td>
</tr>
<tr>
<td>Connection:</td>
<td>dest labad</td>
</tr>
<tr>
<td>Operation:</td>
<td>Insert Oracle labad</td>
</tr>
</tbody>
</table>

For “Connection”, Enter the Database Type as Oracle, and input all host, port, database name etc.

For “Operation”, we would select the “Type” as “Standard Insert/Update/Delete” and type in the Insert statement as shown below, passing ‘?’ for each column which Boomi will substitute for the real data at run time.
Scheduling Data Integration job

Such an integration can be designed and scheduled using the “Scheduling” option as shown below. This runs the integration every 6 hours on all days of the year.

Summary

This whitepaper has demonstrated with practical examples how to integrate SqlServer with Oracle using Dell Boomi Atompshere technology. This has also covered the many benefits of using Cloud based integration and the advantages we gain over traditional ETL or Replication tools. Dell Boomi supports all leading Database technologies and has in-built connectors to them. This technique can be used to integrate any two different heterogeneous databases.
Addendum

Stored Procedure “get_data”

cREATE PROCEDURE get_data
AS
DECLARE @v_date DATETIME, @v1 INTEGER
SELECT @v_date = MIN(curr_date) FROM dbo.log_table WHERE curr_date > (SELECT last_extracted_date FROM dbo.archive_control)
IF @v_date IS NOT NULL
BEGIN
TRUNCATE TABLE stg_log_table
INSERT INTO dbo.stg_log_table SELECT * FROM dbo.log_table WHERE curr_date >= @v_date AND curr_date < DATEADD(DAY, 1, @v_date)
SET @v1 = @@ROWCOUNT
UPDATE dbo.archive_control SET last_extracted_date = @v_date, num_rows = @v1
INSERT INTO archive_history VALUES ('dbo', @a_table_name, @v_date, NULL, @v1, NULL, NULL);
SELECT s_num, CURR_DATE, STRING1, DATE_ARCHIVED, SEQNO FROM STG_log_table
END
GO