

Case Study – I

Data Integration Engineer – Informatica

Prerequisites

- Informatica power center client
- Sql Server

Instructions

- Target table structures are already defined. DDL used to create the tables are given in appendix for reference.
- Two sample sources are given below for two different scenarios

Sample Input 1

 Sample_File - Notepad

File Edit Format View Help

Col1,Col2

A,A

A,A


A,B

B,A

B,B

B,B

Sample Input 2

 Dept - Notepad

File Edit Format View Help

Dept_id,Dept_name

10,Sales

20,Finance

1 Case Study

1.1 Problem Statement/Use Case

In first case Split record Source Sample file contains 2 columns Col1, Clo2 from that you want to fetch unique row and load into unique table and fetch duplicate record load it into Duplicate table. So we have two target table and achieve this scenario in two different approach in one approach consider source as Flat file in another approach consider source as SQL server in this approach try to implement performance tuning and optimize your mapping

Second case is Incremental loading ,record are incremented based upon Dept_id .So in initial load all records from source file Dept needs to be loaded to target table Dept and in upcoming loads only incremented data needs to be loaded

2 Split record

2.1.1 Staging –Approach 1

Create a mapping to load the staging file into Target tables Unique and Duplicate, Unique record from Source needs to be loaded to Unique table and Duplicate record needs to be loaded to Duplicate table. For source details please refer Sample Input 1

Field Name	Data Type	Source Mapping
COL1	VARCHAR(1)	From File
COL2	VARCHAR(1)	From File

2.1.2 Target table

In target table output needs to be with below data

Unique - Notepad	Duplicate - Notepad
File Edit Format View Help	File Edit Format View Help
COL1, CL02	COL1, CL02
A,B	A,A
B,A	A,A
	B,B
	B,B

2.1.3 Mapping

The mapping to achieve the output is

SQ->Aggregator->Router-Target

If they achieve the above scenario via different transformation apart from above transformations they can provide justification why it is better than above mapping

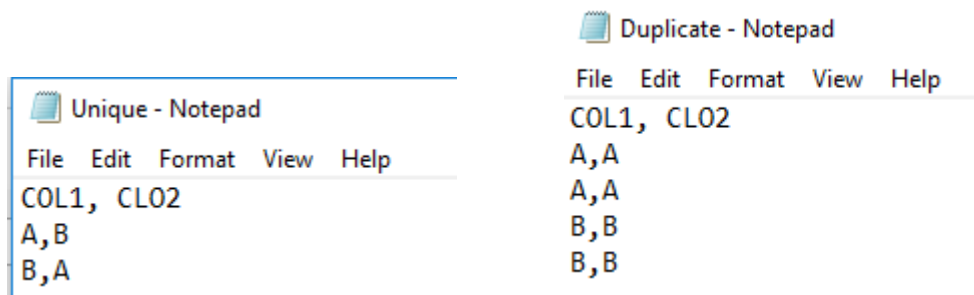
2.1.4 Staging –Approach 2

Create a mapping to load the staging table into Target tables Unique and Duplicate, Unique record from Source needs to be loaded to Unique table and Duplicate record needs to be loaded to Duplicate table. For source details please refer Sample Input 1

Field Name	Data Type	Source Mapping
COL1	VARCHAR(1)	From Table
COL2	VARCHAR(1)	From Table

2.1.5 Target table

In target table output needs to be with below data



2.1.6 Mapping

The mapping to achieve the output is , achieve this scenario in SQ override and mapped the output to router to split the record to load into target table

SQ ->Router-Target

If they achieve the above scenario via different transformation apart from above transformations they can provide justification why it is better than above mapping

2.1.7 Workflow

Create a individual session to invoke the each mapping in the below load order and trigger the session via Workflow sequentially

3 Incremental Loading

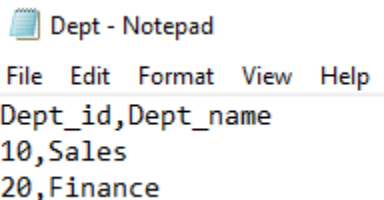
3.1.1 Staging

Create a mapping to load flat file Dept into target table Dept , in upcoming load only incremental records and updated records should be loaded. Incremental records and updated records are found using Dept_id field and also while inserting a record for the first please make sure to capture the system date time in CREATED_DATE field, for the updated records based of DEPT_ID updated records has to be replaced in destination table and UPDATED_DATE should be inserted so whenever if there is any updates in existing records then UPDATED_DATE has to updated as well

Field Name	Data Type	Source Mapping
DEPT_ID	Number(10)	From File
DEPT_NAME	VARCHAR(10)	From File
UPDATED_DATE	DATETIME	
CREATED_DATE	DATETIME	

3.1.2 Target table

In target table output needs to be in below format , in upcoming load only new Dept_id should be loaded



```
Dept - Notepad
File Edit Format View Help
Dept_id,Dept_name
10,Sales
20,Finance
```

3.1.3 Mapping

As it is incremental load we have to find the Max record for each session found that via using Mapping variable (Set Maxvariable) . We are using flat file so we have to filter old data in Filter transformation , in expression transformation find Max id in each load , mapping should be in below way

SQ->Filter->Expression-Target

3.1.4 Workflow

Create session to invoke mapping and trigger the session via Workflow

4 Expected Results

Sr. No.	Output Description	Output Snapshot
1	Data loaded into Target Unique Table	
2	Data loaded into Target DuplicateTable	
3	Initial data loaded into Target Dept Table	
4	Incremental data loaded into Target Dept table	