



# Lessons Learnt : KITS Datastage Migration using BODS



1/27/16

## KITS ETL Migration using BODS

Domain/ Service/ Technology /Geo: **Retail & CPG**

Name of the Project/Support Fn: **KITS ETL Migration**

Name of the Author: **SOUMYAJEET DHAR**

Date Created: **27<sup>th</sup> Jan 2016**

# Description

## Project/Context

- This document explains the lessons learnt during the migration of ETL components from Datastage v7.5 to SAP Business Objects Data Services (BODS) v4.2 in Multi-User Environment

# How it happened

## Process Adopted

Below are the general practices adopted during the ETL code development of the project using SAP BODS

1. Use of 'Check Out Without Replacement' functionality in BODS while committing code into the Central Repository.
2. Turning off Push Down utility to process complex SQL within BODS engine rather using direct push down to database.
3. Use of XSLT scripts to convert Target XML into Canonical form.
4. Use of looping mechanism to create separate output XML for each record in Source Table.
5. Use of VB scripts to deliver any email having multiple attachment from BODS
6. Set File Format property "Maximum errors to stop job" as "1" to handle any error faced during File writing in BODS.

# Issues & Analysis

## Issues Faced

- ❑ Frequent designer crashes during development and loss of unsaved data.
- ❑ Corruption of development workspace.

- ❑ BODS jobs often fail to push down any complex SQL directly into Database.

## Causal

- ❑ BODS designer does not support development on the same local repository from multiple sessions at the same time.

- ❑ BODS v4.2 engine converted complex SQL in case of Push down into internal command which often fail to execute.

# Issues & Analysis(Contd..)

## Issues Faced

- ❑ XML in canonical form can not be produced as target in BODS

- ❑ Separate output XMLs can not be produced for each source records in BODS.

## Causal

- ❑ BODS designer does not support XSD which is being derived from any canonical XML file format.

- ❑ Like “Trigger Column” functionality present in datastage, BODS designer does not have any functionality to create separate XML for each source record.

# Issues & Analysis(Contd..)

## Issues Faced

- Unable to send email with multiple attachments using BODS.

- BODS jobs succeed even if there exists any error while writing to a file as output.

## Causal

- BODS in-built email function “mail\_to()” does not support multiple attachments to be send along with email body.

- BODS is unable to catch the error while writing data to a File. Hence,no failure is being observed.

# Results

## Justification

Using Check Out Without Replacement:

- ❑ SAP Data Services marks all appropriate objects as checked out—in both the object library and in the workspace—but does not copy any objects from the central repository to the local repository.

This is helpful only if the code being checked in to the central repository is the latest working copy which has all the latest updates included.

## Benefits

Various benefits achieved by avoiding 'Check Out without replacement' functionality are:

- ❑ Less development effort spent on need for retrofitting changes.
- ❑ Reduced defects that might occur from retrofitting changes.
- ❑ Synchronized version history among developers during development leveraging increased ease of use in tracking changes.



# Results(Contd..)

## Justification

Turning off Push Down:

- ❑ For complex SQL we can skip direct push down and process this via BODS engine to avoid the unpredictable failure.

## Benefits

- ❑ Though this technique is less recommended by SAP ,use of this may helpful when Push down is failing for Complex SQL.

# Results(Contd..)

## Justification

Use of XSLT Scripts:

- ❑ BODS designer does not support XSD which is being derived from any canonical XML file format.

To convert any output XML format into canonical form we need any external scripting i.e XSLT

## Benefits

Various benefits achieved by using XSLT Scripts are:

- ❑ XML in canonical form can be produced as output.
- ❑ XSLT being well structured scripting language maintenance of code is easy.

# Results(Contd..)

## Justification

Use of looping mechanism for XML output:

- Like “Trigger Column” functionality present in datastage, BODS designer does not have any functionality to create separate XML for each source record.

In order to achieve above looping is done using multiple query transformations .

## Benefits

Various benefits achieved by using looping mechanism are:

- Separate XML for each source record is created in single Dataflow.

- Datastage equivalent functionality is achieved which in turn fulfill the project purpose.

# Results(Contd..)

## Justification

Use of VB Script for email attachment:

- ❑ BODS in-built email function “mail\_to()” does not support multiple attachments to be send along with email body.

Hence external VB script is used to send email to multiple users with compressed file attachment(max upto 5 files)

## Benefits

- ❑ Email with multiple attachments can be send to users.

# Results(Contd..)

## Justification

Change in File format property:

❑ BODS is unable to catch the error while writing data to a File. Hence, no failure is happening.

File Format property “Maximum errors to stop job” as “1” is set to handle any error faced during File writing in BODS.

## Benefits

Various benefits achieved by changing File Format are:

- ❑ BODS job can handle any error occurred during writing of output file.
- ❑ Debugging of failed job is easier with the help of error log.

# Solutions & lessons

## Solutions

---

- The proposed solutions can be applied to any ETL project that is running using SAP Business Objects Data Services tool.

## Lessons Learnt

---

- 'Check in without replacement' functionality should be used very less frequently when the project is running in a multi developer environment.
- Pushdown can be turned off for complex SQL if it is taking too much time to execute.
- To generate XML in canonical form XSLT can be used to modify the file structure.
- In BODS separate output XML for each source record can be generated using loop mechanism.
- Email sending with multiple attachments can be achieved using VB script in BODS.
- Error handling can be done in BODS while writing to a file by modifying File format property.

# How this may be prevented elsewhere

## Prevention

The proposed lessons learnt can be applied to any ETL project that is running under Multi developer environment using SAP Business Objects Data Services tool.

**Soumyajeet Dhar**

**Soumyajeet.dhar@tcs.com**



# Thank You

IT Services  
Business Solutions  
Consulting