Explore to the Extraction Tab of Data Transfer Process in SAP BI 7.0

Applies to:

SAP BI 2004s or SAP BI 7.x. For more information visit the Enterprise Data Warehousing.

Summary

This article will explain about what are all the options available in the Extraction tab of Data Transfer Process and which options has to select in different scenarios.

Author: Venkateswarlu Nandimandalam

Company: Mahindra Satyam

Created on: 20 September 2011

Author Bio

Venkateswarlu Nandimandalam, working in Mahindra Satyam has one and half year experience in SAP BI 7.0 and he has been working in Production Support and Custom Development

Table of Contents

Data Transfer Process:

Data transfer process (DTP) loads data within BI from one object to another object with respect to transformations and filters. In short, DTP determines how data is transferred between two persistent objects. It is used to load the data from PSA to a data target (Cube or DSO or Info Object) thus, it replaced the data mart interface and the Info Package.

Now in this article I am going to explain what are all the options available in the Extraction tab of DTP, and how to use those based on our requirements.

Data Source:

This is the source from which DTP has to extract the data. The below mentioned are different types of sources from which we can extract data using DTP.

1) Data Source

2) Info Cube

3) Data Store Object

4) Info Set

5) Info Object: Attributes

6) Info Object: Texts

7) Info Object: Hierarchies

After we select the type of source, we have to give the technical name of that object in the next field. You can see in below screens respectively.

Extraction:

We have two options in this to load data from the source object to Target object, i.e. Full and Delta

Full: If we want to load entire data from source to target, then we can choose this option.

Delta: Unlike delta transfer using an Info Package, no need to do initialization separately for delta transfer with DTP. When we execute DTP in delta mode for the first time, it retrieves all the requests from the source and the delta status is initialized.

Note: Data mart and delta time stamp will be maintained in table RSMDATASTATE.

In table RSMDATASTATE, field DMALL keep track of delta and it increases when delta completes.

Field DMEXIST is used to prevent deletion of Source request

Deleting data request deletes the entry in table RSSTATMANREQMAP which keep track of the data mart.

If we select extraction mode as Delta, then we will have two more options. They are
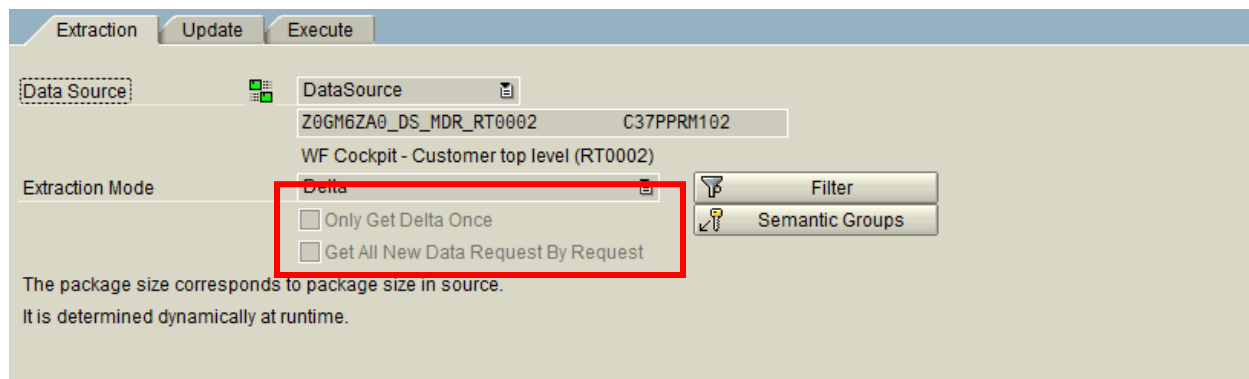
Only Get Delta Once: If we want to transfer the source request only once to the target, then we can select this option. Setting this option ensures that content of the Info Provider is an exact representation of the source data.

A scenario for this is, sometimes we always want an Info Provider to contain the most recent data for reporting, but data source which severs data to the Info Provider is not capable of delta.

If we select this option, before loading the new request, we have to delete the previous request (Normally this will be done using process type in process chain). When there is a new request in data source then only DTP determines it as delta request.

Get All New Data Request By Request: If we run delta DTP first time, it bundles all the requests in source and load it as a single request to the target. If we have huge data in the source then it will cause slow data loading and the sometimes failure of load. To avoid this kind of situations, if we select this option then DTP will load data from Source to target as request by request. For every request in source it will create a separate request in target.

Note: If you set the indicator for a DTP that was created prior to Net Weaver 7.0 Support Package Stack 13, the DTP request only retrieves the first source request. This restricts the way in which the DTPs can be used because requests accumulate in the source, and the target might not contain the current data. To avoid this, you need to execute the DTP manually until all the source requests have been retrieved. The system therefore also displays the following indicator for such DTPs: Retrieve Until No More New Data. If you also set this indicator, the DTP behaves as described above and creates DTP requests until all the new data has been retrieved from the source.

Special Scenarios in Data Extraction

A) If we select data source as Data Source Object (DSO) and extraction mode as full we will get a few more options under extraction. We can see in the below screen.

We can load data from three tables of the DSO to CUBE.

Active Table (with Archive): DTP reads data from Active table of DSO and from the archive or from near line storage if any exists. We can select this option even if DSO doesn't have archived data.

Active Table (Without Archive): DTP reads data from Active table, it doesn't read data in the archive or from near line storage even if it's exists.

Archive (Only Full Extraction): DTP reads data from the archive or from near line storage and it doesn't read data from active table.

If we select extraction mode as Delta, then we will get one more additional table i.e. Change Log table.

Generally delta records will be stored in the change log table and for further loading we will load the data from the change log table.

In some special cases it is possible to load delta from other tables also as shown below.

B) If we select the source as Info Cube then we will have below option to load data from Info Cube.

Info Cube Tables: Data is extracted only from a database (E table, F table and aggregates)

If we select "Use Aggregates" option, the outgoing objects in the transformation for the source Info Cube is compared with the aggregates. If all Info Objects in the target are used in aggregates, the data is read from the aggregates during extraction instead of from the Info Cube's fact table (F table) and table of compressed data (E table). This setting enables us to improve the performance of the data transfer process, as reading aggregates is more efficient than reading the Info Cube's F table and E table.

If we do not select this option, then the data is read from the Info Cube's F table and E table during extraction.

Archive (Only Full Extraction): The data is only read from the archive or from near line storage.

We can see the details in the below screen.

Package Size:

This determines the number of records for one package. By default this is 50000. If you are working in support and DTP load fails with time out error, you can try the load again by reducing the DTP Package size from 50000 to 30000.

Filter ( ):

Instead of transferring large amount data at a time, we can divide the data by defining filters. The filter thus restricts the amount of data to be copied and works like the selections in the Info Package. We can specify single values, multiple selections, intervals, selections based on variables, or routines. To define filters we have to click on the Filter symbol ( ). Choose Change Selection to change the list of Info Objects that can be selected.

The ✔ icon next to pushbutton Filter indicates that predefined selections exist for the data transfer process.

Semantic Groups ( ):

We chose Semantic keys to specify how we want to build the data packages that are read from the source. Based on the key fields we define in semantic keys, records with the same key are combined in a single data package.

This setting is only relevant for Data Store objects with data fields that are overwritten. This setting also defines the key fields for the error stack. By defining the key for the error stack, you ensure that the data can be updated in the target in the correct order once the incorrect data records have been corrected.

Currency Conversion:

This determines whether we want to convert the currency key figures or not.

Key Date master data:

If we have time dependent master which has to be analyzed while loading, then we can specify here.

Related Content

[Creating Data Transfer Processes](#)

http://www.sdn.sap.com/irj/scn/go/portal/prtroot/docs/library/uuid/80dd67b9-caa0-2e10-bc95-c644cd119f46?QuickLink=index&overridelayout=true

For more information visit the [Enterprise Data Warehousing](#)

Disclaimer and Liability Notice

This document may discuss sample coding or other information that does not include SAP official interfaces and therefore is not supported by SAP. Changes made based on this information are not supported and can be overwritten during an upgrade.

SAP will not be held liable for any damages caused by using or misusing the information, code or methods suggested in this document, and anyone using these methods does so at his/her own risk.

SAP offers no guarantees and assumes no responsibility or liability of any type with respect to the content of this technical article or code sample, including any liability resulting from incompatibility between the content within this document and the materials and services offered by SAP. You agree that you will not hold, or seek to hold, SAP responsible or liable with respect to the content of this document.