DB Partitioning & Compression

Applies to:
SAP BI 7, SQL Server 2005. For more information, visit the Business Intelligence homepage.

Summary
The purpose of this document is to outline a strategy for DB partitioning and compression of data of InfoCubes.

Author: Venkata Devaraj P
Company: Satyam Computer Services Limited
Created on: 14 January 2009

Author Bio
Venkata Devaraj P currently with Satyam Computer Services Ltd. is an astute & result oriented IT professional with overall 11 years of experience of which 7.10 Years in SAP BW Implementation, 1.5 Years in SAP FI Roll Outs and 3.7 Years of domain expertise in Finance (Internal audits).
# Table of Contents

1. **Strategy** .................................................................................................................. 3
   1.1. **Background** ........................................................................................................ 3
   1.2. **Options** ............................................................................................................. 4
   1.3. **Recommendations** ............................................................................................. 4

2. **Design Implementations** .......................................................................................... 5
   2.1. **Partitioning of New Cube** ................................................................................. 5
       2.1.1. **Setting Constant for Fiscal variant** ......................................................... 5
       2.1.2. **Definition for Partitioning** ....................................................................... 6
   2.2. **Repartitioning** ................................................................................................... 7
       2.2.1. **Setting Constant for Fiscal Variant** ......................................................... 8
       2.2.2. **Complete Partitioning** .............................................................................. 8
       2.2.3. **Repartitioning Monitor** ............................................................................ 12
   2.3. **How to see the Partitions** .................................................................................. 13
   2.4. **Steps and definition Compression** ...................................................................... 14
       2.4.1. **Option1** ...................................................................................................... 15
       2.4.2. **Option2** ...................................................................................................... 15
       2.4.3. **With Zero Elimination** .............................................................................. 16
       2.4.4. **Automating Compression using Process chain** ....................................... 16

3. **Data Recovery Options** ........................................................................................... 18
   3.1.1. **Before Compression of request** .................................................................. 18
       3.1.1.1. **Delete request from all the data targets** ............................................ 18
       3.1.1.2. **Delete the Data Mart Status in DSO** ................................................... 18
       3.1.1.3. **Delete the Data Mart Status in DSO** ................................................... 19
   3.1.2. **After Compression of request** ...................................................................... 19
       3.1.2.1. **Reverse posting** .................................................................................. 19
       3.1.2.2. **Selective Full Upload** ......................................................................... 20
       3.1.2.3. **Re-Initialization** ................................................................................... 20

4. **Archived and Deleted Partitions** ............................................................................ 21
   4.1.1. **Processing Option – Merging and Adding Partitions** ................................... 21

5. **Appendix A** ............................................................................................................ 22
   5.1. **Background information about Copying Data (refer oss note 1008833)** .......... 22
   5.2. **Background Information about Error Handling (refer oss note 1008833)** .... 22
   5.3. **Transport (refer oss note 1008833)** ................................................................. 22

6. **Appendix B** ............................................................................................................ 23
   6.1. **Example for Maximum Number of Partitions** ................................................. 23

**Related Content** ........................................................................................................ 24

**Disclaimer and Liability Notice** ................................................................................ 25
1. Strategy

1.1. Background

For Infocubes created after implementation of SP9 in BI, system creates default partitions for every request that is loaded with data in the cube.

In SQL Server 2005, the maximum number of partitions permissible per Infocube is 1000. Once this number is reached, the data load into that particular info cube would fail.

Most of the cubes are being loaded daily and if the source is only one, then before the end of 3 years the maximum number of partitions could be reached. And if the cube has multiple sources, then the maximum number would be reached even earlier.

Due to the above scenario there is a need of have a strategy in place.

So far in SQL Server 2005 table partitioning is possible for BW PSA tables, F-Fact tables, and E-Fact tables.

SAP supports SQL Server 2005 range partitioning on the following classes of tables in SAP BW versions 3.5 and higher:

- Persistent Staging Area (PSA) table. This is the main table for the staging data coming from the outside into SAP BW. Here the partition criteria is the ‘LoadID’. The LoadID is an artificial partitioning key for PSA tables that correlates to the number of rows per partition.

  You first have to execute the report SAP_RSADMIN_MAINTAIN with parameter object = DB4_PSA_PARTITIONING and Value = ‘X’ to set the RSADMIN parameter. SAP Note 815186 includes more comprehensive information on this.

- F-Fact (part of a SAP BW InfoCube). The F-Fact table in the cube allows duplicate rows. The F-Fact table contains the ‘Request ID’. This is the request that is associated with loading the data into the cube. The Request ID is the dimension to which the partitions are lined. Each new load gets stored in its own partition.

- E-Fact table (part of a SAP BW InfoCube). The partitioning of the E-Fact table is not mandatory, but it is recommended because the data from F-Fact table is compressed into the E-Fact table. Duplicates in the F-Fact table get rolled up into one row, making the E-Fact table smaller. Since the Request ID is not part of the E-Fact table, the implementation of SAP uses the time dimension for partitioning. Either the calendar month or the fiscal year/period can be used for E-Fact tables. The time range is defined at the creation time of the cube.
1.2. **Options**

The options that are available are given below:

1. Partition all the newly created Infocubes.
2. For existing cubes perform complete repartitioning based on Time characteristics.
3. Compress the request of all cubes on a timely basis. This will transfer the data from the F-Fact table to the E-Fact table.

1.3. **Recommendations**

1. Create Time DB partitions for all new Infocubes using time characteristics Calendar Month (0CALMONTH) or Fiscal Year / Period (0FISCPER).

2. Repartition the infocubes using time characteristics Calendar Month (0CALMONTH) or Fiscal Year / Period (0FISCPER). (refer example in Appendix B for no of partitions)

3. Create Maximum of XXX partitions \((XX \text{ Years} \times YY \text{ Month} \text{ or } Z \text{ QTR}) +2\)

4. Compress the requests that are older than XXX days.

5. Keep the PSA and Change Log retention period in sink with the compression strategy.
2. Design Implementations

2.1. Partitioning of New Cube

You use partitioning to split the total dataset for an InfoProvider into several, smaller, physically independent and redundancy-free units. This separation improves system performance when you analyse data delete data from the InfoProvider.

You can only partition a dataset using one of the two partitioning criteria ‘calendar month’ (0CALMONTH) or ‘fiscal year/period (0FISCPER). At least one of the two InfoObjects must be contained in the InfoProvider.

If you want to partition an InfoCube using the fiscal year/period (0FISCPER) characteristic, you have to set the fiscal year variant characteristic to constant.

2.1.1. Setting Constant for Fiscal variant

Got to the Edit mode of the Cube → Expand the Time Dimension → Right Click on Fiscal Year Variant → Provider-Specific Properties

Enter a Constant “01” and click on Continue

Panel: Provider-Specific Properties of the InfoObject
2.1.2. Definition for Partitioning

RSA1 → Modelling → Infoprocessor → Right Click on Cube → Change

From Menu choose Extras → DB Performance → Partitioning

Choose the Time characteristic based on which you would like to partition and click on continue.

Enter the Time range and also enter the maximum number of partitions and click on continue.
2.2. Repartitioning

Repartitioning can be useful if you have already loaded data to your InfoCube, and:

- You did not partition the InfoCube when you created it.
- You loaded more data into your InfoCube than you had planned when you partitioned it.
- You did not choose a long enough period of time for partitioning.
- Some partitions contain no data or little data due to data archiving over a period of time.

*SAP recommends a complete back up of the database before you execute this function. This ensures that if an error occurs (for example, during a DB catalogue operation), you can restore the system to its previous status.*

In most of the cases there would be a daily back up that would be taken as a normal practice. Therefore it is always easier to take a differential back-up of BI before starting the repartitioning.

RSA1 → Modeling → Infoprovider → Right Click on Cube → Additional Functions → Repartitioning

OR - RSA1 → Administration → Repartitioning
2.2.1. Setting Constant for Fiscal Variant

To partition InfoCubes, Fiscal Variants if used in infocubes, must be set to constant value. For empty infocubes the same can be done from the change mode of the infocube. If data has already been loaded to the InfoCube, the fiscal variant can no longer be changed in the change mode.

Using transaction SE38, the program “RSDU_SET_FV_TO_FIX_VALUE” should be used to set the fiscal variant to a fixed value, provided all the data records loaded to date are assigned to just one fiscal variant. If this is not the case, subsequent partitioning is not possible.

The execution of the above program the update rules / transformation of the infocube would be deactivated since the rule for updating the fiscal variant would be changed from direct move to constant.

The update rules/ transformation will have be manually activated and collected in a transport request and transported to higher regions after the execution of the above mentioned program.

<table>
<thead>
<tr>
<th>Update rule for fiscal variant Before setting the same as constant.</th>
<th>Update rule for fiscal variant after setting the same as constant.</th>
</tr>
</thead>
</table>

2.2.2. Complete Partitioning

Complete Partitioning fully converts the fact tables of the InfoCube. The system creates shadow tables with the new partitioning schema and copies all of the data from the original tables into the shadow tables. As soon as the data is copied, the system creates indexes and the original table replaces the shadow table. After the system has successfully completed the partitioning request, both fact tables exist in the original state (shadow table), as well as in the modified state with the new partitioning schema (original table). You can manually delete the shadow tables after repartitioning has been successfully completed to free up the memory. Shadow tables have the namespace /BIC/4F<Name of InfoCube> or /BIC/4E<Name of InfoCube>.

You can only use complete repartitioning for InfoCubes. A heterogeneous state is possible. For example, it is possible to have a partitioned InfoCube with non partitioned aggregates. This does not have an adverse effect on functionality. You can automatically modify all of the active aggregates by reactivating them.

After entering the technical name of the cube you want to repartition. Click on Initialize.
Choose the Time characteristic (Fiscal Year / Period or Calendar Year / Month) based on which you want to partition the cube. Click on continue.

Enter the Range of Time that you want to partition and the maximum number of partitions. Click on continue.
Refer Appendix B for example on determining the number of partitions.

You will get a message for rebuilding of aggregates after repartitioning. Click Yes if aggregates exist on the infocube else choose No.

On clicking Yes, you will get the below message. Click on continue.

You can now schedule the repartition immediately or at a particular date and time.

You will get the below message that the repartitioning has been scheduled.
Click the Monitor to check the status of repartitioning

**Repartitioning of InfoProviders**

Read the documentation and SAP Note 100883

Repartitioning of InfoProviders

**InfoCube**

**ZC00M_C02**

**Processing Options**

- Adding Partitions
- Merging Partitions
- Complete Repartitioning

**Repartitioning Request**

- Initialize
- Monitor
- Unlock

Check the status of the repartitioning

**Monitor Requests**

<table>
<thead>
<tr>
<th>Object Name</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>CCREA_SHD_EFACT</td>
<td>Create shadow table /BICO/ZC00M_C02 for F fact table /BICO/ZC00M_C02</td>
</tr>
<tr>
<td>CCREA_SHD_FFACT</td>
<td>Create shadow table /BICO/ZC00M_C02 for F fact table /BICO/ZC00M_C02</td>
</tr>
<tr>
<td>SPACE_CHECK</td>
<td>Check free space on DB</td>
</tr>
<tr>
<td>COPY_TO_SHD_EFACT</td>
<td>Copy all data records from /BICO/ZC00M_C02 to /BICO/ZC00M_C02</td>
</tr>
<tr>
<td>COPY_TO_SHD_FFACT</td>
<td>Copy all data records from /BICO/ZC00M_C02 to /BICO/ZC00M_C02</td>
</tr>
<tr>
<td>CREATE_IDX</td>
<td>Create indexes on both shadow tables</td>
</tr>
<tr>
<td>SET_READ_LOCK</td>
<td>Set read lock for InfoCube ZC00M_C02</td>
</tr>
<tr>
<td>VIEW_AGG</td>
<td>Deactivate all active aggregates of InfoCube ZC00M_C02</td>
</tr>
<tr>
<td>DELETE_FACTVIEW</td>
<td>Delete view of fact tables</td>
</tr>
<tr>
<td>CHECK_EFACT</td>
<td>Check data consistency of table /BICO/ZC00M_C02</td>
</tr>
<tr>
<td>CHECK_FFACT</td>
<td>Check data consistency of table /BICO/ZC00M_C02</td>
</tr>
<tr>
<td>SWITCH_EFACT</td>
<td>Swap E fact table /BICO/ZC00M_C02 with /BICO/ZC00M_C02</td>
</tr>
<tr>
<td>SWITCH_FFACT</td>
<td>Swap F fact table /BICO/ZC00M_C02 with /BICO/ZC00M_C02</td>
</tr>
<tr>
<td>CREATE_FACTVIEW</td>
<td>Recreate view of fact tables</td>
</tr>
<tr>
<td>POST_ACT</td>
<td>Adapt BW metadata for InfoCube ZC00M_C02</td>
</tr>
<tr>
<td>REPAIR_IDX</td>
<td>Repair indexes for both fact tables</td>
</tr>
<tr>
<td>ANALYZE</td>
<td>Calculate DB statistics for both fact tables</td>
</tr>
<tr>
<td>RELEASE_READ_LOCK</td>
<td>Release read lock for InfoCube ZC00M_C02</td>
</tr>
<tr>
<td>ACTIVATE</td>
<td>Reactivate all aggregates of InfoCube ZC00M_C02</td>
</tr>
<tr>
<td>CLEANUP</td>
<td>Various cleanup jobs</td>
</tr>
</tbody>
</table>
2.2.3. Repartitioning Monitor

You can monitor the repartitioning requests using a monitor. The monitor shows you the current status of the processing steps. When you double-click, the relevant logs appear. The following functions are available in the context menu of the request or editing step:
Delete: You delete the repartitioning request; it no longer appears in the monitor, and you cannot restart. All tables remain in their current state. The InfoCube may be inconsistent.

Reset Request: You reset the repartitioning request. This deletes all the locks for the InfoCube and all its shadow tables.

Reset Step: You reset the canceled editing steps so that they are reset to their original state.

Restart: You restart the repartitioning request in the background. You cannot restart a repartitioning request if it still has status Active (yellow) in the monitor. Check whether the request is still active (transaction SM37) and, if necessary, reset the current editing step before you restart.

Refer Appendix A for background information about copying data, error handling and transport.

2.3. How to see the Partitions

You can use transaction Se38 and program RSMSSPARTMON to see the no of partitions created and the number of records saved in each of the partition.

### Status of a Partitioned Table

<table>
<thead>
<tr>
<th>Partition ID</th>
<th>Boundary low</th>
<th>Boundary High</th>
<th>Number of rows</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>MINVAL</td>
<td>0000000029</td>
<td>655</td>
</tr>
<tr>
<td>2</td>
<td>0000000029</td>
<td>0000000029</td>
<td>1,073</td>
</tr>
<tr>
<td>3</td>
<td>0000000029</td>
<td>0000000030</td>
<td>88</td>
</tr>
<tr>
<td>4</td>
<td>0000000030</td>
<td>MAXVAL</td>
<td>0</td>
</tr>
</tbody>
</table>

Table info:
- Table Exists: YES
- Table Partitioned: YES
- Number of Partitions: 4
- Created At: 11/44:55/2908:12:31
- Part.Type: RANGE
- Part.Col.Name: KEY_ZCOM_C99P
- Part.Col.Type: int
- Part.Col.Length: 10
- LEFT/RIGHT: RIGHT
- Number of Rows: 1,772
- Next FileGroup
- Decimal Compressed: NO
Similarly you can see the partitions created for the PSA table and the Change log table.

### 2.4. Steps and definition Compression

It is always advisable to first create the partition and then compress the request in the infocube. This is because the E-fact able would not have any data if none of the request are compressed and the creation of the partition would be a lot quicker. Steps for Manually Compressing the request:

In RSA1 → Modeling → Infoprovider → Right Click on Cube → choose manage

Go to Collapse tab
2.4.1. Option1

Request ID. - In this option you can compress the request up to the request id that you enter.

2.4.2. Option2

Calculated Request ID. - In this you can specify a threshold value. During the runtime it is then determined if and to which calculated request ID the data of the request is compressed.

With calculated requests you can
- either specify the number of requests that are not compressed, or
- Specify that only requests older than XXX days are compressed.
2.4.3. With Zero Elimination

In both the options you can opt for "With Zero Elimination". In this case, entries where all the key figures are blank are deleted.

If you want to avoid entries that only contain zero values as key figures (for example reverse posting) are contained in the InfoCube after compressing, you can execute a zero elimination at the same time as compressing.

Zero elimination is only permitted for InfoCubes where key figures with the aggregation behaviour 'SUM' exclusively occur. With non-cumulative values in particular you may not execute a zero elimination.

2.4.4. Automating Compression using Process chain

The first time compression can be done manually and then the same can be added as the last process in the existing process chains as shown below.

Right Click on the Collapse Process Variant → to Display the variant
### Process Display: Compression of the InfoCube

**Variant:** ZCCA_ACTL_CUBE

<table>
<thead>
<tr>
<th>Last Changed By</th>
<th>Changed On</th>
<th>Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>DDIC</td>
<td>09.01.2002</td>
<td>08:54:55</td>
</tr>
</tbody>
</table>

To Select, Press F4 On The Object Type, Then F4 on the Name.

**Selections**

<table>
<thead>
<tr>
<th>Obj.</th>
<th>Object Type</th>
<th>Object Name</th>
<th>Object Name</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>InfoCube</td>
<td>ZCCA_ACTL</td>
<td>Cost Center Actual Values</td>
</tr>
</tbody>
</table>

- Collapse only those requests that were loaded X days ago
  - 189
- Number of requests that you do not want to collapse
  - 0

- Switch on Zero Elimination
- No Marker For Inventory Cube
3. Data Recovery Options

3.1.1. Before Compression of request

There may be a scenario when the data in the cube is not correct or a particular request has not been loaded from the DSO/PSA to the Cube/DSO as the case may be. The steps to recover the data when the request has not been compressed and the data in the PSA is not deleted are given below:

3.1.1.1. Delete request from all the data targets

First selectively delete the request in the Cube, till the day the data was not loaded in the Cube.

3.1.1.2. Delete the Data Mart Status in DSO

Delete the data mart status of the request in the DSO

Click on the data mart status of the request. You will get a screen as shown below. Click on to reset the delta to the cube in the DSO. Below is the 3.X DSO screen shot.

Click Yes to reset the delta to the data target.
In BI7 DSO,

Click on the data mart status of the request. You will get a screen as shown below.

You can delete the request from all the data targets that the request has been loaded. And this will reset the data mart status to the cube in the DSO.

3.1.1.3. Delete the Data Mart Status in DSO

After deleting the data mart status, you can reload the request to the cube using the further update in 3.X or DTP in BI7 scenario or from the PSA to DSO and then the Cube as the case may be.

3.1.2. After Compression of request

The recovery options after the request in the cube has been compressed are given below:

3.1.2.1. Reverse posting

When the data is loaded to BW Infocube and then it is compressed then you can't delete the data pertaining to this request in the cube. For this what you can do is, if the request is loaded via PSA then you can click on the Request Reverse Posting option on the Monitor screen of the particular Request. This will reverse the sign of the key figures loaded into the InfoCube for that particular request only, so that it will make overall key figure value in the cube for this particular request to 0.
It appears that the reverse posting can be done only up to version 3.X (Update rules). The same does not appear active in BI7 (Transformation/DTP).

### 3.1.2.2. Selective Full Upload

Selective full Upload can be used when a data mismatch or inconsistency is identified and the same has to be loaded to BI to correct the mismatch or inconsistency. The selective Full upload can be done only under the below mentioned scenarios:

1. You can narrow down to the exact data mismatch and decide on the selections of the full load.
2. The data load is routed through a DSO.
3. If the load is directly to a cube. After ensuring that the relevant data is not present in the cube or after doing a selective selection.
4. The datasource supports for the relevant selections.

### 3.1.2.3. Re-Initialization

In some of the cases the only and last option available to do a re-initialization of the delta.
4. Archived and Deleted Partitions

4.1.1. Processing Option – Merging and Adding Partitions

When the data of a particular period in the cube has been archived or deleted, the partitions that have been defined for that period would become empty. You can use the merge partitions options of repartitioning to merge the partitions.

InfoCube partitions are either merged at the bottom end of the partitioning schema (merge), or added at the top (split).

Ideally, this operation is only executed for the database catalog. This is the case if all the partitions that you want to merge are empty and no data has been loaded outside of the time period you initially defined. The runtime of the action is only a few minutes.

If there is still data in the partitions you want to merge, or if data has been loaded beyond the time period you initially defined, the system saves the data in a shadow table and then copies it back to the original table. The runtime depends on the amount of data to be copied.

With InfoCubes for non-cumulative, all markers are either in the bottom partition or top partition of the E fact table. Whether mass data also has to be copied depends on the editing options. For this reason, the partitions of non-cumulative InfoCubes cannot be merged if all of the markers are in the bottom partition. If all of the markers are in the top partition, adding partitions is not permitted. If this is the case, use the Complete Repartitioning editing option.

You can merge and add partitions for aggregates as well as for InfoCubes. Alternatively, you can reactivate all of the aggregates after you have changed the InfoCube. Since this function only changes the DB memory parameters of fact tables, you can continue to use the available aggregates without having to modify them.
5. Appendix A

5.1. Background information about Copying Data (refer oss note 1008833)

By default, the system copies a maximum of six processes in parallel. The main process splits dialog processes in the background. These dialog processes each copy small data packages and finish with a COMMIT. If a timeout causes one of these dialog processes to terminate, you can restart the affected copy operations, after you have altered the timeout time. To do this, choose Restart Repartitioning Request.

5.2. Background Information about Error Handling (refer oss note 1008833)

Even if you can restart the individual editing steps, you should not reset the repartitioning request or the individual editing steps without first performing an error analysis.

During repartitioning, the relevant InfoCube and its aggregates are locked against modifying operations (loading data, aggregation, rollup and so on) to avoid inconsistent data. In the initial dialog, you can manually unlock objects. This option is only intended for cases where errors have occurred and should only be used after the logs and datasets have been analysed.

5.3. Transport (refer oss note 1008833)

Since the metadata in the target system is adjusted without the DB tables being converted when you transport InfoCubes, repartitioned InfoCubes may only be transported when the repartitioning has already taken place in the target system. Otherwise inconsistencies that can only be corrected manually occur in the target system.

The only transport would be the activation of the update rule / transformation after setting the fiscal variant as a constant.
6. Appendix B

6.1. Example for Maximum Number of Partitions

If you choose the partitioning criterion 0FISCPER and determine the value range
From 001.2004 To 012.2023

If you want to create one partition for each Fiscal Year / Period then specify the maximum
number of partitions as 242.

20 years x 12 months + 2 = 242 partitions (2 partitions for values that lay outside of the
range, meaning < 001.2004 or >012.2023).

If you want to create one partition for each QTR then specify the maximum number of
partitions as 82.

(20 Years X 4 QTR in a Year) = 80 + 2 = 82 partitions (2 partitions for values that lay outside
of the range, meaning < 001.2004 or >012.2023).
Related Content

Repartitioning Help
Compressing Infocubes
How to Contribute

For more information, visit the Business Intelligence homepage.
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.