



Smarten

Advanced Data Discovery

Powered by ElegantJ BI

Memory & Disk Space Sizing

Business Intelligence & Advanced Data Discovery

Document Information	
Document ID	Smarten-Memory-And-Disk-Space-Sizing
Document Version	8.0
Product Version	5.0 and above
Date	10-Nov-2018
Author	EMTPL

© Copyright Elegant MicroWeb Technologies Pvt. Ltd. 2018. All Rights Reserved.

Statement of Confidentiality, Disclaimer and Copyright

This document contains information that is proprietary and confidential to EMTPL, which shall not be disclosed, transmitted, or duplicated, used in whole or in part for any purpose other than its intended purpose. Any use or disclosure in whole or in part of this information without the express written permission of EMTPL is prohibited.

Any other company and product names mentioned are used for identification purpose only, may be trademarks of their respective owners and are duly acknowledged.

Disclaimer

This document is intended to support administrators, technology managers or developers using and implementing Smarten. The business needs of each organization will vary and this document is expected to provide guidelines and not rules for making any decisions related to Smarten. The overall performance of Smarten depends on many factors, including but not limited to hardware configuration and network throughput.

Contents

1	About this Document	4
2	Terminology	4
3	Estimation of Total Required Memory	4
3.1	Smarten Managed Memory	4
3.2	Smarten Object Memory	5
3.3	Example Calculation	6
4	Cube Disk Size Estimation.....	7
5	Notes.....	8
6	Product and Support Information	9

1 About this Document

This document will enable you to compute the memory requirement for real-time and cached cubes of Smarten. The document also explains calculation for disk space required for the cubes.

It is recommended that this document is read after reading “The Concept and Workings of Managed Memory Computing” document.

2 Terminology

Base Memory: The amount of memory used by the operating system and the application server. The Base Memory will vary depending on the application server and the Operating System being used.

Smarten Managed Memory: The memory used for Managed Memory Computing.

Smarten Object Memory: The memory used by the user objects, such as dashboards or analysis.

Number of unique aggregated records: The number of unique records in the cube derived from the result set retrieved from data source.

Symbol table: Table used to keep unique values of every column with its index value.

Aggregated Row Size: The sum of the size of each column in a particular UARS row.

Number Of Users: The maximum number of concurrent users.

UARSS: Unique Aggregated Record-Set Size. Data size of unique aggregated records for cubes that are enabled for Managed Memory Computing.

3 Estimation of Total Required Memory

In order to estimate Total Required Memory, Smarten Managed Memory and Smarten Object Memory must first be calculated. The equation to estimate Total Required Memory is:

Total Required Memory = Base Memory + Smarten Managed Memory + Smarten Object Memory

Note:

MDX and Real-time cubes or datasets do not use Managed Memory, and hence Managed memory sizing will be 0 (zero) for MDX and Real-time cubes or datasets.

Managed memory sizing is only applicable to Smarten Cached Cubes or datasets - with Aggregation and without Aggregation options.

3.1 Smarten Managed Memory

The calculation of Smarten Managed Memory first requires UARS to be calculated:

Smarten Managed Memory = Aggregated Cube 1 Managed Memory + ...+ Aggregated Cube n Managed Memory + Object 1 Managed Memory + ...+ Object n Managed Memory + Non-aggregated Cube 1 Managed Memory + ...+ Non-aggregated Cube n Managed Memory + Dataset 1 Managed Memory + ...+ Dataset n Managed Memory

*Managed Memory per Cube or Object = NoOfUARSSRecords * UARSRecordSize*

*UARSRecordSize = Dimension 1 Data type size +...+ Dimension n Data type size + (NoOfMeasures * 48) bytes*

*Managed Memory per Dataset or non-aggregated cube = NoOfRecords * RecordSize*

Variable	Description
<i>UARSRecordSize</i>	Record size in bytes for a single record in UARS. It will be calculated as per the formula shown above.
<i>NoOfUARSRecords</i>	Total number of records in the UARS
<i>NoOfMeasures</i>	Number of measures selected in UARS
<i>Dimension data type size</i>	Data type size of Dimension column in bytes. Please refer column data type size table shown below.
<i>NoOfRecords</i>	Total number of records in dataset or cube
<i>RecordSize</i>	Record size in bytes for a single record in dataset or cube. It will be a sum of each column size for a single record.

Column Data type size reference table

Column Data Type Size	
Data Type	Storage Requirement [In Bytes]
SmallInt	2
Integer	4
Double	8
Boolean	1
Bigint	8
Varchar	1 * No. of characters used in string
Timestamp	8

An example computation is noted below.

Cube / Object / Dataset	Record Size (Bytes)	No. of Records	Managed Memory Required (GB)
Aggregated Cube 1	600	5,000,000	2.8
Object 1	80	20,000	0.001
Dataset 1	400	2,000,000	0.75
Non-aggregated cube 1	300	1,000,000	0.28
Total			3.831

Smarten Managed Memory (GB) = Aggregated Cube 1 Managed Memory (GB) + Object 1 Managed Memory (GB) + Non-aggregated Cube 1 Managed Memory (GB) + Dataset 1 Managed Memory (GB)
= 3.831 GB

3.2 Smarten Object Memory

To calculate Smarten Object Memory, Total Cells must be calculated followed by Memory per Object. The memory used per object can be derived from the equation below. It is advisable to use the largest object size for calculating the Object Memory, which is usually the most complex dashboard.

Once the Memory per Object is known, calculate the Smarten Object Memory as follows:

*Smarten Object Memory = Memory per Object * number of concurrent users*

Memory per Object = TotalCells x MemoryVariant x ResultSetMultiplier

*Total Cells = NoOfRows * NoOfColumns*

MemoryVariant = 37 + Sum of used Dimension columns data type size

Variable	Description
NoOfRows	Total number of rows in the BI object
NoOfColumns	Total number of columns in the BI object
MemoryVariant	MemoryVariant is number of bytes required for one BI object cell. It is calculated based on above mentioned formula. It is total data type size of used dimension columns + 37 bytes for measure value.
ResultSetMultiplier	For cached cube objects This value will be 5. For MDX / Real-time cube objects This value will be 6.

An example computation is noted below, using the assumptions that the MemoryVariant has a value of 77 (37 + 40 bytes for 2 string type dimensions).

Total No. of Rows in BI object	Total No. of Columns in BI object	Total No. of Cells in BI object	Memory Required per BI Object per User (in MB)	
			Cached Cube Object	MDX / Real-time Cube Object
5,000	25	125,000	45.9	55.05
10,000	30	300,000	110.15	132.18
50,000	50	2,500,000	917.90	1101.49
200,000	20	4,000,000	1468.66	1762.39

3.3 Example Calculation

This section shows examples for total memory requirements calculations for below mentioned scenario.

Here is the example scenario:

- Base memory required by the operating system, application server and other running process on server is 1.5 GB.
- Cube1 is a cached aggregated cube with 10 Dimensions and 5 measures. UARS records are 100000. UARS record size is 490 bytes. Based on calculation shown above, cube 1 managed memory is:

Cube1 Managed memory	100000 * 490 = 49000000 bytes	46.73 MB
----------------------	-------------------------------	-----------------

- Cube2 is real-time cube with 10 dimensions and 5 measures. As real-time cubes do not use Managed memory, Cube2 Managed memory size is 0.
- Cube3 is a cached non-aggregated cube with 10 columns. Total records are 100000. Record size is 80 bytes. Based on calculation shown above, cube3 managed memory is:

Cube3 Managed memory	$100000 * 80 = 8000000$ bytes	7.63 MB
----------------------	-------------------------------	----------------

- Dataset1 is a cached dataset with 20 columns. Total records are 500000. Record size is 200 bytes. Based on calculation shown above, Dataset1 managed memory is:

Dataset1 Managed memory	$500000 * 200 = 100000000$ bytes	9.54 MB
-------------------------	----------------------------------	----------------

- Dashboard1 is created using Analysis1 and Analysis2. Analysis1 is created from Cube1 (cached cube) using 4 dimensions and 2 measures. Analysis2 is created from Cube2 (Real-time) cube using 3 dimensions and 2 measures. Memory per object is calculated as shown in section 3.2.

Object Name	No. Of Rows	No. Of Columns	Total no. Of cells	Memory Variant	Resultset Variant	Object Memory (MB)
Analysis1	750	100	75000	97	5	34.7
Analysis2	900	50	45000	82	6	21.1
Dashboard1						55.8

- There will be usage of 5 concurrent users.

Total Managed memory (MB) = 46.73 (Cube1) + 0 (Cube2) + 7.63 (Cube3) + 9.54 (Dataset1) = **63.9 MB**

Total Object Memory (MB) = 55.8 (Dashboard1 memory) * 5 (concurrent users) = **279 MB**

Total required memory

Base Memory (MB)	Total Managed Memory (MB)	Total Object Memory (MB)	Total Required Memory (MB)
1500	63.9	279	1842.9

4 Cube or Dataset Disk Size Estimation

To calculate the disk space required for Smarten Cubes and datasets, record size of the data, number of records, compression ratio and level of aggregation in the source data should be considered.

Cube or Dataset Size (GB) = $(\text{RecordSize} * \text{NoOfRecords} * \text{CompressionRatio}) * \text{AggregationVariant} / (1024 * 1024 * 1024)$

Total Disk space required is then calculated as follows:

Total Disk space required GB = Cube1 size + + Cube_n Size + Dataset1 size + + Dataset_n Size

Variable	Description
RecordSize	Record size in bytes for a single record in cube or dataset. It will be a sum of each column size for a single record.
NoOfRecords	Total number of records in the cube or dataset
CompressionRatio	Compression ratio factor. It will be 0.4 and 0.55.

AggregationVariant	<p>If drill through data is stored in the cube: This value is between 1.25 and 2 based on level of aggregation in the source data. If all records in the data source are unique, that results into UARSS size to be same as source data, and this multiplier would 2.</p> <p>If drill through data is not stored in the cube: This value is between 0.25 and 1 based on level of aggregation in the source data. If all records in the data source are unique, that results into UARSS size to be same as source data, and this multiplier would 1.</p> <p>If aggregated data is not stored in the cube: This value is 1.</p> <p>For datasets: This value is 2.</p>
--------------------	---

An example computation is noted below.

Cube	Record Size (Bytes)	No. of Records	Compression Ratio	Aggregated Variant	Disk space Required (GB)
Cube 1	600	5,000,000	0.4	1.3	1.45
Cube 2	900	20,000,000	0.5	1.5	12.57
Datase t1	500	1,000,000	0.5	2	0.47
Total					14.49

Note:

This Disk space size estimation is applicable only for Smarten Cached Cubes and datasets. For MDX and Real-time cubes or Datasets, only metadata information is stored in Smarten, so this calculation will not be applicable for MDX and Real-time cubes or Datasets.

5 Notes

- Sizing and other formulas mentioned in this document is based on measurement of performance and various activities using specific computer systems and/or components and reflects the approximate performance and requirements of Smarten as measured by those tests
- Any difference in system hardware, network or software design or configuration, may affect actual requirements
- Requirements may vary upon variation, non-performance or failure resulting out of third party software like operating systems, platforms, servers, tools, utilities and programs
- Requirements may vary upon data structure and level of usage
- Hypothetical or simulated tests have certain inherent limitations
- Under no circumstances will EMTPL be liable for any special, indirect, incidental, exemplary or consequential damages of any kind or nature whatsoever, whether based on contract, warranty, tort (including negligence), strict liability or otherwise, arising out of or in any way related to Smarten performance or sizing requirements.

6 Product and Support Information

Find more information about ElegantJ BI-Smarten and its features at www.smartent.com

Support: support@smartent.com

Sales: sales@smartent.com

Feedback & Suggestions: support@smartent.com

Support & Knowledgebase Portal: support.smartent.com