

Sizing

ASAP FOR BW ACCELERATOR SAP BUSINESS INFORMATION WAREHOUSE



How to size of a BW system

Version 3.1
December 2004

Table of Contents

ASAP for BW Accelerator.....	1
1 INTRODUCTION	1
2 SIZING	1
2.1 R/3 System.....	1
2.2 BW-Systems	1
2.2.1 Overview	1
2.2.2 First sizing approach (T-shirt sizing).....	3
2.2.3 Detailed sizing approach	4
2.2.3.1 Disk Space Requirements in General	4
2.2.3.1.1 Overview	4
2.2.3.1.2 Estimating an InfoCube	4
2.2.3.1.3 Dimension Tables	4
2.2.3.1.4 Fact Tables.....	5
2.2.3.1.5 PSA tables.....	5
2.2.3.1.6 ODS tables.....	5
2.2.3.1.7 Limitations.....	5
2.2.3.1.8 Re-Estimation	5
2.2.3.2 Simplified estimation of disk space (used by the QuickSizer)	7
2.2.3.3 CPU requirements in general	8
2.2.3.3.1 Simplified CPU sizing for data staging	9
2.2.3.3.2 Simplified CPU sizing for queries	9
2.2.3.4 Memory requirements in general.....	11
2.2.3.4.1 Simplified memory sizing for data staging	12
2.2.3.4.2 Simplified memory sizing for queries.....	13
2.2.3.5 Sizing result for 2-Tier systems	14
2.2.3.6 Sizing result for 3-Tier systems.....	14
2.2.3.7 Frontend PC requirements	14

1 Introduction

This document describes how to estimate hardware resources that are necessary to run a particular workload on a BW System, providing guidelines for the definition of hardware requirements like main memory, disk space and CPU. Generally, it is much more difficult to estimate the sizing requirements of a BW system than an R/3 system. The reason is that user behavior is more difficult to predict than in an R/3 system, as BW provides great flexibility in order to meet the varied reporting needs of an organization. With this in mind, the goal of this document is to present general guidelines that can help in determining hardware requirements for a BW system. **It is always encouraged to work directly with hardware vendors to receive further assistance in sizing a BW system.**

Software Version Supported

This document applies to BW 3.x versions.

The recommendations provided in this document, regarding estimation of hardware requirements of BW systems, are assumptions based on empirical data gathered from existing installations and based on the output of certified BW benchmarks.

2 Sizing

Resources requirements for achieving a particular throughput and response times depend on several factors such as e.g. database, operating system, number of users, amount of online/background operations, load profile.

The values always refer to a minimum configuration and should be increased in cases of doubt.

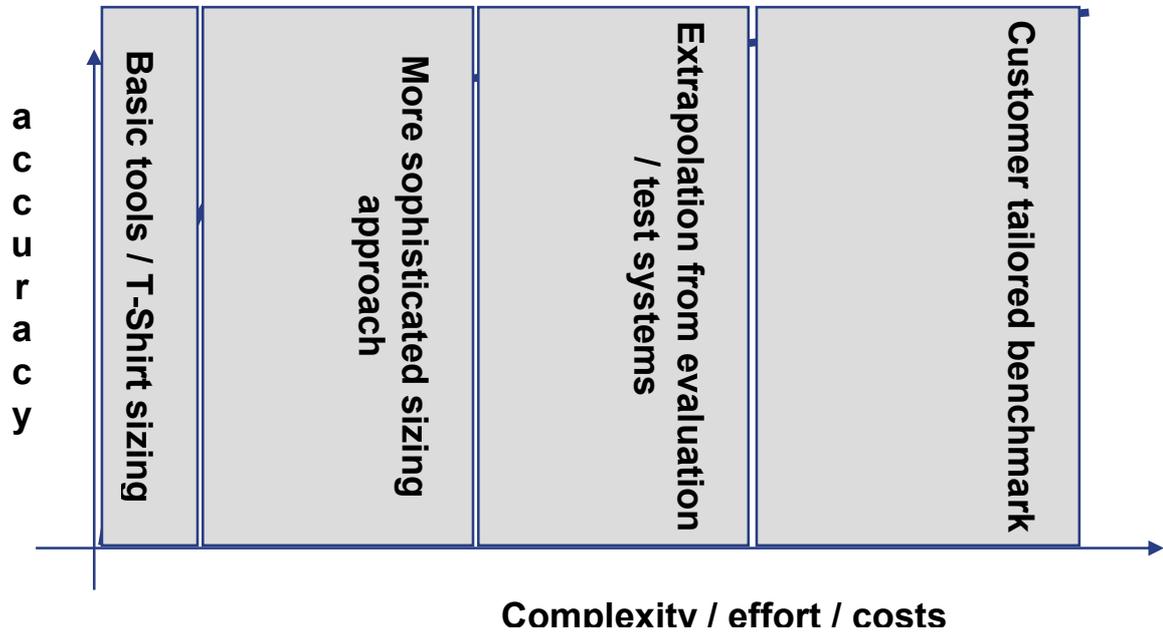
2.1 R/3 System

When connecting a BW system to an R/3 system it has only a minor net impact on the hardware sizing of the R/3 system. The reason for this is that the presence of a BW system alleviates much of the load that reporting typically places on the R/3 system. On the other hand data extraction can require significant R/3 system resources.

2.2 BW-Systems

2.2.1 Overview

The reliability of a BW sizing is based on the given information. Based on the information you can choose one out of the 2 different sizing methods. The first approach (T-shirt sizing) is used for starting projects gathering first experiences with BW. That means neither the customer nor a consultant is able to provide detailed information about the future implementation. If there is more detailed information available you should use the second approach including Disk, CPU and Memory sizing.



In the figure above you can easily see the dependencies between accuracy and the costs of the sizing procedure. The more accurate the sizing, the more complex and costly is the sizing procedure. The detailed sizing approach introduced in this document is based on assumptions described in the following chapters. Depending on how well these assumptions are met, the result of the sizing procedure can differ from the actual resource requirements. A reliable sizing can only be done running a customer tailored benchmark.

2.2.2 First sizing approach (T-shirt sizing)

Categories of sizes have been developed for the hardware sizing for a BW System. They correspond to T-shirt sizes.

BW Category (T-Shirt size)	Configuration	Hardware	No. of users
XS	DB + App-server on one server	>= 2 GB of RAM, >=2CPUs, DB 30 – 50GB	1 – 10
S	DB + App-server on one Server	>= 2 GB of RAM, 2 – 4 CPUs DB 50 – 100GB	10 – 20
M	DB -, App-server on one or several servers dependant on the hardware configuration	>= 4 GB of RAM, 4 CPUs for DB- and App-server DB 100 – 200 GB	20 – 50
L	DB- + App-server on several servers dependent on the hardware configuration	>=4 GB of RAM, DB instance >=4 CPUs, App-server >=2CPUs, >= 2 GB of RAM DB 200 – 1000 GB	50 – 150
XL	One DB instance, multiple App- servers	>=8 GB of RAM, DB instance >=8 CPUs, App-server >=4CPUs, >= 4 GB of RAM >1000 GB	> 150

2.2.3 Detailed sizing approach

2.2.3.1 Disk Space Requirements in General

2.2.3.1.1 Overview

Disk space requirements depend heavily on the design of InfoCubes and data distribution. The following hints on estimating disk space do not account for possibly sparse tables or compression the database might use. There is a QuickSizer available on <http://service.sap.com/QuickSizer> that helps you perform what is described below.

The Business Information Warehouse software including the development workbench requires approximately 9 GB of disk space.

The required disk space for data depends on:

- Number of InfoCubes
- Number of key figures and dimensions per InfoCube
- Volume of data loaded
- Number of aggregates
- Amount of master data
- Number and size of flat files used for data load, residing on application servers
- The number and size of PSA and ODS tables
- Indexes

2.2.3.1.2 Estimating an InfoCube

When estimating the size of an InfoCube, tables like fact and dimension tables are considered. However, the size of the fact table is the most important, since in most cases it will be 80-90% of the total storage requirement for the InfoCube.

When estimating the fact table size consider the effect of compression depending on how many records with identical dimension keys will be loaded.

The amount of data stored in the PSA and ODS has a significant impact on the disk space required. If data is stored in the PSA beyond a simply temporary basis, it could be possible that more than 50% of total disk space will be allocated for this purpose.

2.2.3.1.3 Dimension Tables

- Identify all dimension tables for this InfoCube.
- The size and number of records need to be estimated for a dimension table record. The size of one record can be calculated by summing the number of characteristics in the dimension table at 10 bytes each. Also, add 10 bytes for the key of the dimension table.
- Estimate the number of records in the dimension table.
- Adjust the expected number of records in the dimension table by expected growth.
- Multiply the adjusted record count by the expected size of the dimension table record to obtain the estimated size of the dimension table.

2.2.3.1.4 Fact Tables

- Count the number of key figures the table will contain, assuming a key figure requires 17 bytes.
- Every dimension table requires a foreign key in the fact table, so add 6 bytes for each key. Don't forget the three standard dimensions.
- Estimate the number of records in the fact table.
- Adjust the expected number of records in the fact table by expected growth.
- Multiply the adjusted record count by the expected size of the fact table record to obtain the estimated size of the fact table.

2.2.3.1.5 PSA tables

The record length of a PSA table can be estimated by adding the field lengths of all the fields of the respective InfoSource and 48 bytes for request-id and other fields being added by the system.

Multiplying the record length by the number of records being held in the PSA table will give you an estimate on disk space needed per PSA table.

Data in PSA might be deleted regularly depending on your scenario (e.g. use ODS for permanent storage). Estimate the number of records accordingly.

2.2.3.1.6 ODS tables

The record length of a ODS table can be estimated by adding the field lengths of all the fields of the respective InfoSource.

After activation of data the activation queue is empty, so the space needed for new data can in most cases be neglected.

For the change log table add 48 bytes for request-id and other fields being added by the system.

Multiplying the record length by the number of records being held in the ODS table will give you an estimate on disk space needed per ODS table.

The number of records of the change log table depends on the average number of updates and the delta type used for the ODS object.

The minimum size of the change log table has the same size like the active data table (assuming that all records are distinct).

2.2.3.1.7 Limitations

There are no limits to the number of InfoObjects, number of InfoCubes, or the number of reports per InfoCube.

The system administrator can create 13 dimension tables in addition to the dimensions for time, Packet ID, and measure of unit for each InfoCube.

Number of characteristics in all dimension tables:

Each of the dimension tables can have 248 InfoObjects, computed by $255 - 6$ (no access to these fields) $- 1$ (Dimension ID) to obtain 248. For performance reasons we recommend a maximum of 16 InfoObjects per dimension.

A fact table can have up to 233 key figures, computed by $255 - 6$ (no access to these fields) $- 16$ (dimensions) to obtain 233.

2.2.3.1.8 Re-Estimation

SIZING AND PERFORMANCE

After making an estimate, you should verify it by loading a small amount of data, such as a few thousand records and then analyzing the space used to confirm the original estimate. You should also analyze the input sources to determine the number of records from each source system.

If RAID (Redundant Array of Inexpensive Disks) is utilized, adjust the overall estimate accordingly. For example, when using RAID 1 (disk mirroring), you should double your estimate.

Estimates on disk space should be checked frequently. Data Warehouse requirements can change over time and can be dramatically changed by apparently minor modifications. Therefore, the impact on data storage requirements should be considered for every change. Also, reviews need to be performed periodically, such as every three to six months, in order to ensure that current growth estimates are correct.

2.2.3.2 Simplified estimation of disk space (used by the QuickSizer)

The following assumptions are made in order to simplify the disk sizing:

- compression factor resulting out of identical records is 0% (all records are distinct)
- index space is added with 100%
- space for aggregates is additional 100%
- compared to the size of the fact table we ignore the size of the dimension tables
- PSA data is deleted after upload or after a certain time.
- No. of master data records is less than 10% of the fact table (could be wrong for Retail customers)
- We ignore compression algorithm of the different database vendors.

A simplified estimation of disk space for the BW can be obtained by using the following formula:

For each cube:

Size in bytes =

$(n + 3) \times 10 \text{ bytes} + (m \times 17 \text{ bytes}) *$

[rows of initial load + rows of periodic load * no. of periods]

n = number of dimensions

m = number of key figures

+

For each ODS (incl. change log):

Size in bytes =

$(n \times 10 \text{ bytes}) + (m \times 17 \text{ bytes}) *$

[rows of initial load + rows of periodic load * no. of periods]

n = number of character fields

m = number of numeric fields

Remarks:

- The change log has the same size as the active ODS table if you load only distinct data. This is reflected in an uplift (see below).

- For ease of use we have assumed an average length of 10 bytes per character field. If you use character fields which are significantly longer you should increase the number of fields in the QuickSizer accordingly.

To the obtained results you have to add

- + 20% for PSA
- + 100% of total size of the InfoProviders for aggregates / change log (ODS)
- + 100% for indexes
- + 10% for master data
- + twice the size of the biggest fact table - at least 10 GB

2.2.3.3 CPU requirements in general

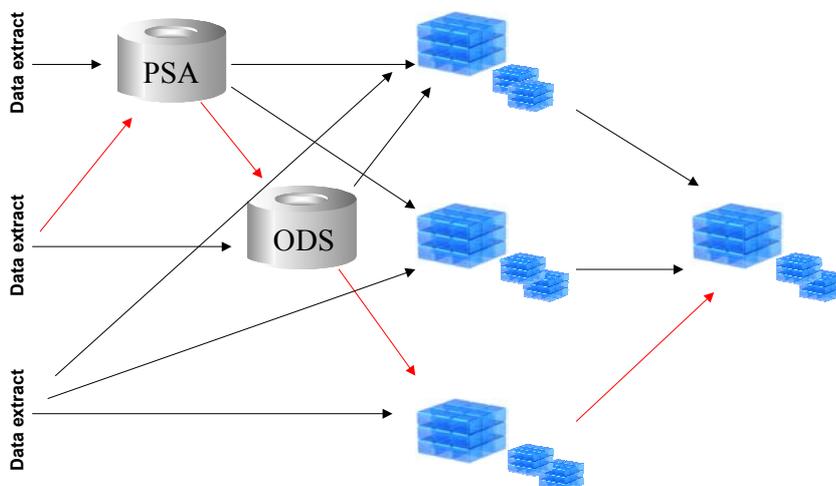
The CPU requirements of BW depend heavily on the design of the data model and the queries/user behavior.

The CPU resource requirements for data staging depend on:

- number of InfoProviders
- volume of data loaded
- number of aggregates
- amount of rollups/change-runs
- number of distinct requests

In order to find the CPU requirements you have to determine the critical path for the data staging. This path is different from customer to customer and could be different from upload to upload.

As an example, the red path in the picture below describes the critical path including load, rollup, change-run, index rebuild, gathering statistics. This path has to be measured in order to get the number of records per hour, which could be applied to the system.



The CPU resource requirements for queries depend on:

- number of InfoProviders
- number of connected users
- number of aggregates
- types of queries

In order to find the CPU requirements, you have to classify your queries in different query classes like “easy”, “medium” and “heavy” and analyze the different response times in order to get relational factors. Additionally you have to classify your user behavior in different groups like “Information Consumers”, “Executive Users” and “Power Users” (a definition of the properties for these user groups can be found below) and to examine which user type is using which kind of query.

2.2.3.3.1 Simplified CPU sizing for data staging

The following assumptions are made in order to simplify the CPU sizing for data staging:

- A system is able to apply 750000 records per hour per job
- No special Transformation rule and update rule are used
- The critical path contains an ODS –object , an InfoCube and 10 aggregates
- The data applied to the system could be split into several requests (one job per request)
- 350 SAPS are used for one job

The CPU requirements are calculated as follow:

$$\text{SAPS} = \#\text{jobs} * 350 \text{ SAPS} / 0.65 \text{ (at 65\% CPU consumption)}$$

While the number of jobs is calculated as following:

$$\#\text{jobs} = \{\text{sum (delta upload)} / \text{maintenance window}\} / 750.000$$

2.2.3.3.2 Simplified CPU sizing for queries

The number of users working in the system forms the basis for estimating CPU resources of the database and application servers. Different definitions of “user” are available:

- Information Consumers
- Executive Users

- Power Users
 - **Information Consumers:** accesses the BW System from time to time, typically are seeking information occasionally using predefined reports and generate in average **1 navigation step** per hour (averaged over 1 week).
 - **Executive Users:** access the BW System regularly and continuously, typically navigating within reports supported by aggregates, generating **11 navigation steps** per hour (averaged over 1 week).
 - **Power Users:** work intensively with the BW System, running ad hoc queries which in most cases are not supported by aggregates, generating **33 navigation steps** per hour (averaged over 1 week).

Remarks:

- The classification of the users corresponds to the user classification within the early watch report.
- A default user distribution based on a customer survey is 71% : 27% : 2%

Additionally we have to define different types of reports being used within BW:

- Easy queries
 - Medium queries
 - High queries
-
- **„Easy“ queries:** predefined reports using optimized aggregates. These queries are weighted by a factor 1.
 - **„Medium“ queries:** slicing and dicing, navigating in reports using various aggregates These queries are weighted by a factor 1,5 (i.e. these queries require 50% more resources than easy queries)
 - **„Heavy“ queries:** ad-hoc reports with unpredictable navigation paths, access of detail data, full table scans. These queries are weighted by a factor 5 (i.e. these queries require five times as many resources as easy queries)

Remark:

These factors are based on the BW benchmark scenario. In different scenarios using other reports the factors could be different.

SIZING AND PERFORMANCE

Given the user and query definitions described above, a matrix describes the usage of different query types by different user groups. The values in the matrix are based on best practice and can be changed in SAP QuickSizer if required:

		Easy query 1	Medium query 1,5	Heavy query 5
# normal users	X	80%	20%	0%
# advanced users	Y	50%	50%	0%
# power users	Z	0%	0%	100%

Using the numbers of users in the three different user categories (in the formula denoted by 'x', 'y', 'z') we calculate the total number of navigation steps to be handled by the system:

$$\begin{aligned}
 \text{Total \# of Nav/h} = & X * (0,8 * 1 + 0,2 * 1,5 + 0 * 5) * 1 + \\
 & Y * (0,5 * 1 + 0,5 * 1,5 + 0 * 5) * 11 + \\
 & Z * (0 * 1 + 0 * 1,5 + 1 * 5) * 33
 \end{aligned}$$

Remark:

The SAP BW benchmark derives a total number of navigation steps as one of the key performance indicators, allowing for a close correlation of benchmark results and QuickSizer results.

Finally, the number of navigation steps needs to be converted into SAPS. We use the following formula for this:

$$\text{SAPS} = (\text{Nav/h} * 1.33 * 2.28 * 9 / 60) / 0.65 \text{ (at 65\% CPU consumption)}$$

The various factors in this formula were derived from several tests. In a comparison of SAP BW benchmark results with SD benchmark results we have obtained a factor of 2.28 (The definition of SAPS is based on the SD benchmark and the notion of dialog steps). In average one dialog step produces the same load as 9 navigation steps. Finally, we have introduced an adjustment factor of 1.33 by comparing the obtained results with productive installation figures.

2.2.3.4 Memory requirements in general

The memory requirements of BW depend heavily on the design of the data volume and the number of connected users

The memory resource requirements for data staging depend on:

- number of parallel jobs
- number of records loaded within one data package
- Complexity of update and transformation rule
- degree of parallel processes within the RDBMS
- BW buffers configured for the system (table buffers, program buffers, export/import,...)
- RDBMS buffers (data buffer, process buffer)

In order to obtain precise information about the memory requirements, you would have to check the memory consumption for each distinct load step in your system and determine the maximum memory consumption for all parallel jobs. Note that the critical path described above is not always the step which consumes the most memory.

The memory resource requirements for queries depend on

- number of connected user
- keep-alive time for WAS
- data volume of the result set

During the daily load, you would have to monitor how many users are connected to the system. The memory used by these users can be measured with the transaction sm04 and could be summarized. Another parameter which impacts memory consumption is the keep-alive time of the Web Application Server. The WAS keeps a connection open for this time even if the user has closed his browser.

2.2.3.4.1 Simplified memory sizing for data staging

The following assumptions are made in order to simplify the memory sizing for data staging:

- Number of records per data package is 50000
- No special Transformation rule and update rule are used
- The critical path contains an ODS object, an InfoCube and 10 aggregates
- 300 MB are used for one job
- minimum of BW buffer is 500 MB
- minimum RDBMS buffer is 700 MB
- Max. RDBMS buffer is 10 GB
- 10 MB buffer is used for each million of records of the largest InfoCube
- 120 MB is used for each process connected to the RDBMS
- for parallel processing 2 times * number of jobs is used within the RDBMS

The BW memory requirements are calculated as follow:

$$\text{Memory BW} = \text{\#jobs} * 300 \text{ MB} + 500 \text{ MB}$$

The number of parallel load jobs is calculated as follows:

$$\text{\#jobs} = \{ \text{sum (delta upload)} / \text{maintenance window} \} / 750.000$$

This number can also be used as a minimum number of batch processes to be configured.

The RDBMS memory requirements are calculated as follow:

$$\text{Memory} = 700 \text{ MB} + \max(\text{rows of InfoProvider}/1000000) * 10 \text{ MB} + \text{\#Jobs} * 120\text{MB} + 2 * \text{\#Jobs} * 120\text{MB}$$

The number of parallel jobs is calculated as described above.

2.2.3.4.2 Simplified memory sizing for queries

The following assumptions are made in order to simplify the memory sizing for queries:

- Users are connected via WAS
- Keep alive time is 60 seconds
- 30 MB are used for one job
- minimum of BW buffer is 500 MB
- minimum RDBMS buffer is 700 MB
- Max. RDBMS buffer is 10 GB
- 120 MB is used for each process connected to the RDBMS
- no parallel processing for queries
- One BW process can handle 1 high or 2 medium or 5 low users concurrently

The BW memory requirements are calculated as follows:

$$\text{Memory} = 500 \text{ MB} + \text{number of connected users} * 30 \text{ MB}$$

The number of connected users is calculated as following:

$$\text{Number of connected users} = (\text{\#Information Consumers} + \text{\#Executive Users} + \text{\#Power Users}) / 3600 \text{ s} * \text{keep-alive time}$$

The RDBMS memory requirements are calculated as follow:

$$\text{Memory} = 700 \text{ MB} + \max(\text{rows of InfoProvider}/1000000) * 10 \text{ MB} + \text{number of concurrent users} * 120\text{MB}$$

The number of connected users is calculated as described above.

2.2.3.5 Sizing result for 2 –Tier systems

If the system should be installed on central server the following calculation should be made in order to obtain the resource consumption figures:

SAPS = max (SAPS data staging ; SAPS query)

**Memory = max (memory data staging (RDBMS + BW) ;
memory query (RDBMS + BW))**

2.2.3.6 Sizing result for 3-Tier systems

If the system is separated into a dedicated RDBMS server and a BW server the distribution of resources is:

BW:

SAPS = max (50% SAPS data staging ; 80 %SAPS query)

Memory = max (memory data staging (BW) ; memory query (BW))

RDBMS:

SAPS = max (50% SAPS data staging ; 20 %SAPS query)

Memory = max (memory data staging (RDBMS) ; memory query (RDBMS))

Note that the total amount of SAPS for a 3-tier system could be higher than for a 2-tier system since CPU resources of a central system can be shared dynamically.

2.2.3.7 Frontend PC requirements

Please refer to note 321973