Introduction to and Overview of Performance and Sizing

Dr. Ulrich Marquard
SAP AG
Agenda

Performance and scalability

The Quick Sizer

- User-based sizing
  - CPU
  - Disk space
  - Memory
- Quantity structure-based sizing
  - CPU
  - Disk space

Considering special processes
Forecasting Specific Business Processes: Scalability

Small test system, one user

Forecast

Thousands of users
High throughput
High concurrency

© SAP AG 2001, WR16D3W1, TechED 2001

THE BEST-RUN E-BUSINESSES RUN SAP
Analyzing Possible Metrics to Define "Large Systems"

- Number of users
- Number of servers
- Size of network, geographical distribution of users
- Data throughput, number of transactions
- Size of database and database tables
Analyzing the "Number of Users"

Possible definitions
- Concurrent users
- Logged-on users
- Active users
- Named users

Design criteria
- Ensure scalability
- Minimize the consumption of critical resources during the thinktime of the users
Analyzing the "Number of Servers"

The multi-tier c/s architecture requires
- Presentation servers
- Internet servers
- Application servers
- Database servers

Design criteria
- Scalability
- Local data and local buffers
- System administration
- Workload distribution and balancing
- Redundancy and fault tolerance
Analyzing "Size of Network" and "Geographical Distribution of Users"

Possible network settings
- WAN, LAN
- Requirements to network bandwidth, costs, latency

Design criteria
- Minimizing network traffic (Roundtrips and amount of data)
- Synchronous and asynchronous communication
- Thin client
- Software distribution and installation
Analyzing "Data Throughput" and "Number of Transactions"

Technical terms
- Screen changes
- Database transactions
- Database calls

Business application terms
- Number of
  - Business objects
  - Business transactions
  - Business processes

Design criteria
- Scalability
- Load balancing
- Parallelization
Analyzing the "Size of Database and Database Tables"

Database settings
- Number and size of tables, views, indexes, ...
- Number and size of rows, fields
- Number of blocks
- Size of buffers

Design criteria
- Data and index design
- Monitoring and administration tools
- Parallelization (query, index creation, backup, ...)
- Table, index partitioning
- Data archiving
- I/O subsystems
Agenda

Performance and scalability

The Quick Sizer

- User-based sizing
  - CPU
  - Disk space
  - Memory

- Quantity structure-based sizing
  - CPU
  - Disk space

Considering special processes
Basic Design Guidelines of the Quick Sizer

Availability 24 x 7
Customers, partners, and SAP have access to the same project info
Be a cost-effective tool
Guarantee up-to-date information for everyone at the same time
Consider standard business applications implemented in mySAP.com
Make a reliable forecast for 80% of all sizings
The Sizing Approach - Three Steps

Step 1 – User-Based Sizing

- Check of basic feasibility
- Useful for initial budget planning
- Limit set to 800 SD (Sales & Distribution) benchmark users
- 33% CPU utilization
CPU Sizing

CPU utilization is calculated against 33%

- User actions can only be estimated
- Ensure good response times even at peak load times

Basics of the formula

- Normalize the number of users per application to high users
- Each application has a respective load
  - The load factors stem from SAP Standard Application Benchmarks and feedback from customer systems
- Normalize to SD Benchmark users (including reference factor)

Result of CPU sizing is the number of SD Benchmark users and SAPS
"Load Factors" based on the SAP Standard Application Benchmarks Version 4.6 - Example

Platform & Release dependent

Rel. CPU-Usage / Dialog step

- Dialog
- Update
- Database

SAP Business Component

SAP AG 2001, WR16D3W1, TechED 2001
Disk Sizing

Average disk space consumed per user

Considers the number of workdays per year

Offset for system itself

Basics of the formula

- All normalized users over all applications are summed up
- Number of normalized users of type medium (not high) is determined
- 1.65 MB per medium norm user
- The disk size in GB is then

\[
\text{(all medium norm users} \times \text{workdays} \times 1.65 + 16,500) / 1,024
\]
Memory Sizing

For application server and DB server (optimal and minimal)

Basics of the formula

- Different memory consumptions by applications
  - 5 different memory classes (e.g. FI, MM-WM + BWP in class one)
  - The classes are each subdivided into application server and DB server
  - The memory class is determined by the most expensive reference application
  - Class 5 is extra (SD + PP, if entered in the same project)

- Optimal memory sizing (all users)
- Minimal memory sizing (only medium and high users)
Example: Scope of User-Based Sizing

Call up Quick Sizer, customer number 32432, project name User-based_00

Available user inputs (510 users altogether)
- FI: 390 low, 90 medium, and 30 high users

Check the results
- CPU sizing
- Disk sizing
- Memory sizing
Example: Scope of User-Based Sizing

Call up Quick Sizer, customer number 32432, project name User-based_00B

Available user inputs (510 users altogether)
- FI: 200 low, 90 medium, and 30 high users
- SD: 190 medium

Check the results
- CPU sizing
  - Explain the jump to category 7+
- Disk sizing
  - Explain the jump to category 7+
- Memory sizing
Conclusions

Cannot consider business processes and their implementation

CPU consumption and disk space used cannot be approximated
The Sizing Approach - Three Steps

Step 2 – Quantity Structure-Based / User-Based Sizing

For more than 800 SD benchmark users
More detailed input
Necessary for batch oriented load
65% CPU utilization
CPU Consumption - Assumptions

Initial Question: Number and size of objects processed within a certain time frame

There is no distinction between

- Processing documents in batch or in dialog
- "Create with reference" or create without reference

Optimizations due to mass processing in batch are neglected

- Invoicing, goods movement, ...

Dependent documents which are created automatically are contained

- MM document + FI, SD-invoice + FI
  - Load from aggregation is accounted for
- Double counting is possible
Disk Space I

Initial Question: How large are they and how long do they reside in the system

Some impacts are not asked for

- **Basis system Source, load, ...**
  - Included in the installation requirements

- **Master data**
  - Customer, addresses, material, accounts, cost centers, BOMs, knowledge base, ...
  - Can be neglected when compared to document type data

- **Objects that only reside a very short time in the System**
  - "Intermediate" data
    - IDOC, WORKFLOW, SPOOL, batch input, job log
  - Data that are deleted automatically
    - Purchase requisition, planned order created by MRP run
    - Requirements, incompletion protocol and due list created by order
Disk Space II

Only consider "directly" created objects

- Automatically created objects will only be taken into account, if they cannot be avoided

- Examples
  - Order + Pricing document, but no CO or LIS data
  - Post goods issue + MM-Document, but no FI document
  - Invoice, but no FI or CO-PA document

Attention: Possible disk impacts that cannot be anticipated

- FI documents created by interfaces to FI
  - Aggregation may cause many FI documents
  - If material valuation is not used there are no FI documents

- Number of MM documents for backflush

- Analysis tools (for example LIS, CO-PA)
  - Depend on data constellation
  - Very little influence when highly aggregated
Example: Many Objects vs. Large Objects

Call up Quick Sizer, customer number 32432, project name QS-based_00

Inputs in FI

- FI-Documents per year: 1,000,000
- Line items: 10
- Retention period: 12 months

Check the results

Create a new project and change the following:

- Objects per year: 10,000
- Line items: 1,000
- Retention period: 12

Draw conclusions from the differing results
Example: Influence of the Retention Period

Call up Quick Sizer, customer number 32432, project name QS-based_00A

Inputs in SD

- Objects per year: 10 000 000
- Line items: 10
- Retention period: 12

Check the results

Create a new project and change the retention period to 3

Draw conclusions from the differing results
The Sizing Approach - Three Steps

Step 3 – Considering Special Processes and Factors

Step 2 – Quantitative Structure-Based / User-Based Sizing

For more than 800 SD benchmark users

More detailed input necessary for batch oriented load

65% CPU utilization

Step 1 – User-Based Sizing

Check of basic feasibility

Useful for initial budget planning

Limit set to 800 (Sales & Distribution) benchmark users

33% CPU utilization
Calculating Specific Business Processes: Database

Determine the number of database calls
- SAP SQL-trace (summary)

Compare results with standard benchmark transaction
- RSAMON80

Size benchmark transactions

Scale sizing results for database server with scaling factor

\[ P_{DB} = \frac{\text{No. of database calls of customized process}}{\text{No. of database calls of standard benchmark transaction}} \]
Calculating Specific Business Processes: SQL Trace

Tune summary shows number of database accesses and table rows

<table>
<thead>
<tr>
<th>Trade/prog</th>
<th>Table</th>
<th>SQL op</th>
<th>Accesses</th>
<th>Recs.</th>
<th>Time Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>VA01</td>
<td>TVZBT</td>
<td>SEL</td>
<td>1</td>
<td>56</td>
<td>60.114 0.4</td>
</tr>
<tr>
<td>VA01</td>
<td>TWP4</td>
<td>SEL</td>
<td>1</td>
<td>32</td>
<td>537 0.2</td>
</tr>
<tr>
<td>VA01</td>
<td>TWTCTRL</td>
<td>SEL</td>
<td>1</td>
<td>29</td>
<td>14.573 0.1</td>
</tr>
<tr>
<td>VA01</td>
<td>TNT_ACTION</td>
<td>SEL</td>
<td>1</td>
<td>12</td>
<td>12.106 0.1</td>
</tr>
<tr>
<td>VA01</td>
<td>USR02</td>
<td>SEL</td>
<td>1</td>
<td>1</td>
<td>31.225 0.2</td>
</tr>
<tr>
<td>VA01</td>
<td>U134H</td>
<td>SEL</td>
<td>1</td>
<td>1</td>
<td>11.358 0.1</td>
</tr>
<tr>
<td>VA01</td>
<td>VAMPA</td>
<td>SEL</td>
<td>1</td>
<td>0</td>
<td>22.724 0.1</td>
</tr>
<tr>
<td>VA01</td>
<td>VAMUE</td>
<td>SEL</td>
<td>1</td>
<td>1</td>
<td>6.168 0.0</td>
</tr>
<tr>
<td>VA01</td>
<td>V3DPPFLFM</td>
<td>SEL</td>
<td>2</td>
<td>1</td>
<td>12.477 0.1</td>
</tr>
<tr>
<td>VA01</td>
<td>V3DINTRF</td>
<td>SEL</td>
<td>1</td>
<td>2</td>
<td>11.394 0.1</td>
</tr>
<tr>
<td>VA01</td>
<td>V Ext Act</td>
<td>SEL</td>
<td>1</td>
<td>0</td>
<td>12.258 0.1</td>
</tr>
<tr>
<td>VA01</td>
<td>V PROJ1</td>
<td>SEL</td>
<td>1</td>
<td>1</td>
<td>25.148 0.2</td>
</tr>
<tr>
<td>VA01</td>
<td>V PRSS1</td>
<td>SEL</td>
<td>9</td>
<td>9</td>
<td>226.798 1.4</td>
</tr>
<tr>
<td>VA01</td>
<td>V PROJ2</td>
<td>SEL</td>
<td>1</td>
<td>1</td>
<td>13.054 0.1</td>
</tr>
<tr>
<td>VA01</td>
<td>VPROJ1</td>
<td>SEL</td>
<td>1</td>
<td>0</td>
<td>12.705 0.1</td>
</tr>
<tr>
<td>VA01</td>
<td>VNR35</td>
<td>SEL</td>
<td>1</td>
<td>0</td>
<td>9.067 0.1</td>
</tr>
<tr>
<td>VA01</td>
<td>VNTADD1</td>
<td>SEL</td>
<td>1</td>
<td>0</td>
<td>4.352 0.0</td>
</tr>
</tbody>
</table>

Total: 736 10.477 16161 465 100.0
Calculating Specific Business Processes: SQL Trace II

<table>
<thead>
<tr>
<th>Table</th>
<th>SQL type</th>
<th>Accesses</th>
<th>Records</th>
</tr>
</thead>
<tbody>
<tr>
<td>KNK</td>
<td>SEL</td>
<td>5</td>
<td>1</td>
</tr>
<tr>
<td>KNMP</td>
<td>SEL</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>KNV</td>
<td>SEL</td>
<td>8</td>
<td>1</td>
</tr>
<tr>
<td>KDOU</td>
<td>SEL</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>KOMP</td>
<td>SEL</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>KOTO003</td>
<td>SEL</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>LT08</td>
<td>SEL</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>MHK</td>
<td>SEL</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>NHMA</td>
<td>SEL</td>
<td>6</td>
<td>16</td>
</tr>
<tr>
<td>NHAC</td>
<td>SEL</td>
<td>2</td>
<td>26</td>
</tr>
<tr>
<td>NM0D</td>
<td>SEL</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>NH017</td>
<td>SEL</td>
<td>19</td>
<td>5</td>
</tr>
<tr>
<td>NM03M</td>
<td>SEL</td>
<td>1</td>
<td>7</td>
</tr>
<tr>
<td>NMAN</td>
<td>SEL</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>NLS01</td>
<td>SEL</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>NLS02</td>
<td>SEL</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>NLR01</td>
<td>SEL</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>NLR02</td>
<td>SEL</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>NLR01</td>
<td>SEL</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>NLR02</td>
<td>SEL</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>NLR01</td>
<td>SEL</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>HR0001</td>
<td>SEL</td>
<td>4</td>
<td>4</td>
</tr>
</tbody>
</table>
Calculating Specific Business Processes: Application Server

Use SAP single statistic records (ST03) to determine CPU-time
Use memory display of SM04 to determine memory consumption
Size benchmark transactions
Scale memory consumption with scaling factor

\[ P_{\text{appl. Mem}} = \frac{\text{Memory consumption}}{\text{Memory consumption of user-based sizing}} \]

Scale CPU requirements with scaling factor

\[ P_{\text{appl. Mem}} = \frac{\text{Measured CPU time}}{\text{CPU time of benchmark transaction on identical box}} \]
Specific Business Processes: Memory Consumption

Transaction SM04 shows memory consumption for all sessions

<table>
<thead>
<tr>
<th>Int</th>
<th>User</th>
<th>Transaction</th>
<th>Roll</th>
<th>Page</th>
<th>Mem (Total)</th>
<th>Mem (Priv.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>000</td>
<td>MARQUARD</td>
<td>SM04</td>
<td>150</td>
<td>224</td>
<td>32,796</td>
<td>2,184,565</td>
</tr>
<tr>
<td>002</td>
<td>JOSTE</td>
<td>SM04</td>
<td>150</td>
<td>224</td>
<td>6</td>
<td>1,137,600</td>
</tr>
<tr>
<td>002</td>
<td>MARQUARD</td>
<td>SM05</td>
<td>388</td>
<td>640</td>
<td>40,988</td>
<td>3,235,181</td>
</tr>
<tr>
<td>002</td>
<td>JOELO</td>
<td>SM05</td>
<td>212</td>
<td>892</td>
<td>6</td>
<td>2,166,951</td>
</tr>
<tr>
<td>002</td>
<td>MARQUARD</td>
<td>ST05</td>
<td>630</td>
<td>784</td>
<td>1,769,472</td>
<td>39,717,380</td>
</tr>
<tr>
<td>000</td>
<td>SARSYS</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>000</td>
<td>SARSYS</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>000</td>
<td>SARSYS</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Specific Business Processes: Statistic Records

CPU time for each dialog step

<table>
<thead>
<tr>
<th>End time</th>
<th>Task</th>
<th>Program</th>
<th>Time (ms)</th>
<th>Wait (ms)</th>
<th>CPU (ms)</th>
<th>DB req. (ms)</th>
<th>Load/Gen kBytes</th>
<th>Phys. db transfer (ms)</th>
<th>Changes</th>
</tr>
</thead>
<tbody>
<tr>
<td>15:15:50</td>
<td>VAB1</td>
<td>SAPN45A</td>
<td>D 0810</td>
<td>1 MARCO</td>
<td>716</td>
<td>14.922</td>
<td>1</td>
<td>360</td>
<td>435</td>
</tr>
<tr>
<td>15:15:52</td>
<td>VAB1</td>
<td>SAPN45A</td>
<td>D 0810</td>
<td>1 MARCO</td>
<td>198</td>
<td>14.922</td>
<td>1</td>
<td>140</td>
<td>41</td>
</tr>
<tr>
<td>15:15:53</td>
<td>VAB1</td>
<td>SAPN45A</td>
<td>D 0810</td>
<td>1 MARCO</td>
<td>374</td>
<td>14.700</td>
<td>1</td>
<td>310</td>
<td>689</td>
</tr>
<tr>
<td>15:16:00</td>
<td>VAB1</td>
<td>SAPN45A</td>
<td>D 0810</td>
<td>1 MARCO</td>
<td>183</td>
<td>14.922</td>
<td>1</td>
<td>150</td>
<td>38</td>
</tr>
<tr>
<td>15:16:02</td>
<td>VAB1</td>
<td>SAPN45A</td>
<td>D 0810</td>
<td>1 MARCO</td>
<td>236</td>
<td>14.700</td>
<td>1</td>
<td>270</td>
<td>78</td>
</tr>
<tr>
<td>15:16:22</td>
<td>VAB1</td>
<td>SAPN45A</td>
<td>D 0810</td>
<td>1 MARCO</td>
<td>268</td>
<td>14.700</td>
<td>1</td>
<td>290</td>
<td>187</td>
</tr>
<tr>
<td>15:16:30</td>
<td>VAB1</td>
<td>SAPN45A</td>
<td>D 0810</td>
<td>1 MARCO</td>
<td>184</td>
<td>14.552</td>
<td>1</td>
<td>160</td>
<td>37</td>
</tr>
<tr>
<td>15:16:33</td>
<td>VAB1</td>
<td>SAPN45A</td>
<td>D 0810</td>
<td>1 MARCO</td>
<td>397</td>
<td>14.700</td>
<td>1</td>
<td>370</td>
<td>779</td>
</tr>
<tr>
<td>15:16:58</td>
<td>VAB1</td>
<td>SAPN45A</td>
<td>D 0810</td>
<td>1 MARCO</td>
<td>3.246</td>
<td>15.050</td>
<td>1</td>
<td>1425</td>
<td>233</td>
</tr>
<tr>
<td>15:17:14</td>
<td>VAB1</td>
<td>SAPN45A</td>
<td>D 0810</td>
<td>1 MARCO</td>
<td>1.370</td>
<td>14.244</td>
<td>1</td>
<td>700</td>
<td>817</td>
</tr>
<tr>
<td>15:17:30</td>
<td>VAB1</td>
<td>SAPN45A</td>
<td>D 2118</td>
<td>1 MARCO</td>
<td>734</td>
<td>14.027</td>
<td>58</td>
<td>460</td>
<td>292</td>
</tr>
<tr>
<td>15:19:59</td>
<td>VAB1</td>
<td>SAPN45A</td>
<td>D 0810</td>
<td>1 MARCO</td>
<td>11.486</td>
<td>3.217</td>
<td>1</td>
<td>1180</td>
<td>3.457</td>
</tr>
<tr>
<td>15:23:53</td>
<td>VAB1</td>
<td>SAPN45A</td>
<td>D 0810</td>
<td>1 MARCO</td>
<td>13.673</td>
<td>4.800</td>
<td>8</td>
<td>3.300</td>
<td>6.665</td>
</tr>
<tr>
<td>15:24:21</td>
<td>VAB1</td>
<td>SAPN45A</td>
<td>D 0810</td>
<td>1 MARCO</td>
<td>10.837</td>
<td>6.490</td>
<td>1</td>
<td>2700</td>
<td>3.775</td>
</tr>
<tr>
<td>15:26:01</td>
<td>VAB1</td>
<td>SAPN45A</td>
<td>D 0120</td>
<td>1 MARCO</td>
<td>2.865</td>
<td>6.521</td>
<td>1</td>
<td>600</td>
<td>157</td>
</tr>
<tr>
<td>15:26:51</td>
<td>VAB1</td>
<td>SAPN45A</td>
<td>D 0120</td>
<td>1 MARCO</td>
<td>2.865</td>
<td>6.521</td>
<td>1</td>
<td>1920</td>
<td>681</td>
</tr>
<tr>
<td>15:26:03</td>
<td>VAB1</td>
<td>SAPN45A</td>
<td>D 0580</td>
<td>1 MARCO</td>
<td>63.750</td>
<td>13.690</td>
<td>6.682</td>
<td>20.070</td>
<td>662</td>
</tr>
<tr>
<td>15:26:21</td>
<td>VAB1</td>
<td>SAPN45A</td>
<td>D 0120</td>
<td>1 MARCO</td>
<td>1.264</td>
<td>13.920</td>
<td>1</td>
<td>640</td>
<td>169</td>
</tr>
<tr>
<td>15:26:25</td>
<td>VAB1</td>
<td>SAPN45A</td>
<td>D 0120</td>
<td>1 MARCO</td>
<td>3.814</td>
<td>14.072</td>
<td>1</td>
<td>1700</td>
<td>414</td>
</tr>
<tr>
<td>15:26:35</td>
<td>VAB1</td>
<td>SAPN45A</td>
<td>D 0120</td>
<td>1 MARCO</td>
<td>3.814</td>
<td>14.072</td>
<td>1</td>
<td>7000</td>
<td>414</td>
</tr>
<tr>
<td>15:26:53</td>
<td>VAB1</td>
<td>SAPN45A</td>
<td>D 0120</td>
<td>1 MARCO</td>
<td>3.814</td>
<td>14.072</td>
<td>1</td>
<td>10400</td>
<td>3454</td>
</tr>
<tr>
<td>15:26:53</td>
<td>VAB1</td>
<td>SAPN45A</td>
<td>D 0120</td>
<td>1 MARCO</td>
<td>3.814</td>
<td>14.072</td>
<td>1</td>
<td>21000</td>
<td>414</td>
</tr>
<tr>
<td>15:26:53</td>
<td>VAB1</td>
<td>SAPN45A</td>
<td>D 0120</td>
<td>1 MARCO</td>
<td>3.814</td>
<td>14.072</td>
<td>1</td>
<td>40000</td>
<td>414</td>
</tr>
</tbody>
</table>
Specific Business Processes: Calculation Continued

Calculate network load to DB server or presentation server
- Network monitors

Calculate DB size
- Use SE12 (runtime object display) to determine row length / indexes
- For reference use Quick Sizer

Consider disk layout and I/O bandwidth
- Use SQL Trace to check which tables are accessed
  - SELECT, UPDATE, INSERT, DELETE

Determine network load between application servers
- Network monitors
  - IDocs and RFCs use data compression
Customer Performance (Load) Tests

Prove
- Scalability of mySAP.com
- Concepts and performance forecasts from single user tests

Require a dedicated system with sufficient real-life data
- Merely copying master data won’t do
- Scalability depends heavily on data distribution

Are performed by a team of experts
- Cost expensive

Don’t always run smoothly from start to finish
- Article in SAP Professional Journal:
Conclusion

The design of the mySAP.com architecture is the basis for high scalability

Sizing mySAP.com is possible through sizing the various mySAP.com elements while ensuring scalability

Transparent and consistent sizing methodologies

Useful addresses:
http://service.sap.com/performance
http://service.sap.com/sizing
http://www.sap.com/benchmark