Energy efficient servers in Europe

Energy consumption, saving potentials and measures to support market development for energy efficient solutions

Bernd Schäppi, Austrian Energy Agency
Bernhard Przywara, SUN Microsystems
Frank Bellosa, University of Karlsruhe
Thomas Bogner, Austrian Energy Agency
Silvio Weeren, IBM
Robert Harrison, Robert Harrison Associates
Alain Anglade, ADEME

The E-Server Consortium

Supported by
# Table of contents

1 Summary ............................................................................................................................................1

2 Energy consumption and saving potentials for servers and data centres in Europe ..................7
   2.1 Methodology ..................................................................................................................................7
   2.1.1 Electric power consumption ..........................................................................................................7
   2.1.2 Energy saving potentials ..............................................................................................................8
   2.2 Results .........................................................................................................................................12
   2.2.1 Electric power consumption .........................................................................................................12
   2.2.2 Energy saving potentials .............................................................................................................18

3 Current market barriers limiting the use of energy efficient server technology .....................23
   3.1 Introduction ...................................................................................................................................23
   3.2 General results from the survey ......................................................................................................24
   3.3 Evaluation of demand side barriers ..............................................................................................25
   3.3.1 Lack of awareness .......................................................................................................................26
   3.3.2 Lack of information .....................................................................................................................26
   3.3.3 Shortcomings in structures and processes ..................................................................................26
   3.3.4 Limited consideration of costs .....................................................................................................27
   3.4 How to overcome the barriers? ......................................................................................................28
   3.4.1 Awareness raising and education .................................................................................................28
   3.4.2 Information on energy efficient equipment and data centre solutions ......................................28
   3.4.3 Improvement of structures and processes ..................................................................................29
   3.4.4 Holistic consideration of costs, financing models and incentives ..........................................29

4 Technical measures to improve the energy efficiency of servers and data centres ..................31
   4.1 Power Supplies ...............................................................................................................................31
   4.1.1 AC-DC Conversion ......................................................................................................................31
   4.1.2 Energy Savings Potentials ..........................................................................................................33
   4.2 Voltage Regulators .......................................................................................................................33
   4.3 Dynamic Power Management .......................................................................................................33
   4.3.1 Shutdown-Strategies ...................................................................................................................33
   4.3.2 CPU Power Management ..........................................................................................................35
   4.3.3 Memory Power Management .....................................................................................................37
   4.3.4 Disk Drive Power Management .................................................................................................38
   4.4 Thermal Balancing .......................................................................................................................39
   4.5 Consolidation through Virtualization ...........................................................................................39
   4.5.1 Principles of System Virtualization ...........................................................................................41
   4.5.2 Principles of Container Virtualization ........................................................................................45
   4.6 Server Cooling ..............................................................................................................................48
   4.6.1 General aspects and saving potentials .......................................................................................48
   4.6.2 Concepts to improve energy efficiency of server cooling .......................................................49
   4.6.3 General economic aspects ..........................................................................................................54
   4.7 Summary of energy saving options ...............................................................................................54

5 Energy efficiency criteria and benchmarks .................................................................................57
   5.1 Introduction ...................................................................................................................................57
Summary

1 Summary

Project objectives

This report presents a summary of the results of the international project E-Server which has been conducted within the EU programme Intelligent Energy Europe (2007-2009). The project’s central objective was to support the market development for energy efficient servers thereby exploiting energy and cost saving potentials.

In the first stage of the project the EU market for servers, the server energy consumption and energy saving potentials were analysed. Furthermore the total electric power consumption and the saving potentials in data centres were estimated including trends up to 2012.

The second stage was dedicated to best practise cases demonstrating the effectiveness of measures to improve energy efficiency in practise and to the development of tools to support the market development for energy efficient solutions.

Among other major tasks the project contributed to the development of energy efficiency criteria in the Energy Star programme and the development of procurement guidelines for IT managers.

Energy consumption and saving potentials

The annual energy consumption of servers and data centres in EU-27 was calculated based on market data from IDC and energy consumption data for the most popular server models provided by manufacturers.

The general concept for the calculations followed the approach of Koomey and LBNL (2007) used for the estimation of total energy consumption for the US server market, however with some modifications.

Energy consumption in Western Europe (EU 15 plus Switzerland) in 2006 amounted to 14.7 TWh for servers and 36.9 TWh for data centres including storage, network components and infrastructure (cooling, UPS, lighting). The number of servers installed was 6.77 Mio units. The Eastern European countries only account for 7 % of the total EU-27 energy consumption of servers. The total electric power consumption of data centres in EU-27 in 2007 was close to 40 TWh/a.

Electric power consumption of servers has strongly increased since 2003. This increase is largely due to a fast growing number of shipments in the so called “volume server” segment but also to a higher average power consumption of recent server models. Installed base has dramatically increased in the lowest volume server segments while decreasing in all other segments.

In terms of installed server units the Western European market is about 63 % the size of the US market. Correspondingly the Western
European electric power consumption is about 60% of the US-consumption. The total installed base in EU-27 therefore is close to 70% of the US market.

Scenarios on the development of the server market and the energy consumption of servers were calculated on the basis of market forecasts by IDC for 2007–2011 for Western Europe. Energy saving potentials where estimated with a focus on the market segment of volume servers which is most relevant in terms of stock and energy consumption. Saving potentials were calculated based on three future scenarios (business as usual, moderate savings and forced savings). The main options for energy saving measures taken into account were consolidation by virtualization, power management and energy efficient hardware. No specific measures to reduce electric power consumption have been assumed for the mid range and high end server segment, since both product segments are of minor importance regarding energy saving potentials and are less accessible regarding energy efficiency measures. For these segments the standard forecast by IDC was used in all three scenarios.

The business as usual (B.A.U.) scenario is based on a current trend for virtualization suggested by IDC and rather minimal efforts in the area of energy efficient hardware and power management. This scenario suggests an increase of energy consumption of the server hardware from 14,7 TWh in 2006 to 31 TWh in 2011 which corresponds to a doubling of the energy demand.

The second scenario, the moderate savings scenario assumes higher levels of virtualization, higher shipment numbers for energy efficient hardware and a higher degree of implementation of power management. For the total server market including mid-range and high-end servers the energy consumption increases from 14,7 TWh in 2006 to 23,5 TWh in 2011. This is equal to a 60% increase but at the same time means a 24% energy saving compared to the business as usual scenario.

Thirdly a forced efficiency scenario proposes high virtualization levels as well as high shares of shipment numbers for energy efficient hardware and intense use of power management. For the total server market including mid-range and high-end servers energy consumption would increase from 14,7 TWh in 2006 to 17 TWh in 2011. This would correspond to energy savings of 45% compared to the business as usual scenario and energy cost savings of 1,7 billion € in the year of 2011.

Looking at the data centre level in the B.A.U. scenario the electric power consumption would rise from 36,9 TWh in 2006 to 77 TWh in 2011 which corresponds to a 110% increase. In the forced scenario electric power consumption would drop to 31,9 TWh in 2011 which is a 13,5% decrease compared to the 2006 level and a 58% decrease compared to business as usual in 2011. Thus an overall saving of 45 TWh or 5,4 billion € could be realized by the forced saving scenario in 2011. The accumulated energy and cost savings of the forced scenario compared to the BAU scenario would be 101 TWh or 12 billion € respectively between 2008 and 2011.

Current market barriers for the use of energy efficient server technology and energy efficient design of data centres

Current market barriers regarding the use of energy efficient server technology were analysed based on interviews and manufacturers’ experience. Different types of companies covering private and public services took part in the survey.

The main barriers encountered as well as the approaches to overcome them can be summarised as follows:

- **Lack of awareness among experts**

  The awareness of responsible IT experts in companies regarding options for energy efficiency is small. Most experts lack information on the topic and argue that energy efficiency is a new topic which has not been addressed so far. Monitoring for energy efficiency in data centres so far
seems rare. Therefore awareness raising and information activities are important in order to get experts familiar with the objectives and options regarding energy efficiency.

• Missing declaration for energy efficiency or labelling for server hardware

Experts asserted that there is little information provided regarding the energy efficiency of server equipment. Specific product information or labelling could help to eliminate this barrier.

• Structural problems and split incentives

In many companies structural problems are currently limiting the activities for more energy efficiency. Very often experts responsible for the procurement of equipment are not at the same time responsible for infrastructure and energy costs. Therefore in many cases a change of incentives and responsibilities would be important to stimulate effective actions.

• Focus on initial investment and limited application of life cycle costs and TCO

Experts often do not consider TCO (Total Cost of Ownership or Life cycle costs) as a criterion and even if they do energy efficiency is normally not included as a parameter. This leads to the fact that the purchasing prize clearly dominates the buying decision. TCO concepts including energy efficiency therefore should be promoted in procurement guidelines and information concepts.

• Dominant procurement criteria and avoidance of risks

Besides the price, performance and availability are the major factors dominating the buying decision. Experts fear that measures addressing energy efficiency could also affect these major criteria. They tend to stick to already proven concepts and to avoid risks.

Consequently there is a need for demonstration cases showing that energy efficiency measures can be implemented without negatively affecting the primary criteria.

Technical measures to improve the energy efficiency of servers and data centres

Many technical options to improve energy efficiency at the server and data centre level already have been published in literature. Part of the project was dedicated to a compilation and evaluation of the existing options.

Server hardware and software

Measures at the level of the server hardware, as for example the use of efficient power supplies, CPUs, hard disks, motherboards and memory, offer saving potentials in the order of magnitude of 10–30%.

Power management offers another option to improve energy efficiency. While the application of power management for CPUs and hard disks is undisputed, power management at system level involving several standby options for servers is not generally applicable.

Most effective short term optimisation currently can be achieved by consolidation concepts involving virtualisation. Based on this approach energy savings of up to 80% are realistic.

Data centre level

All inefficiency caused at the server hardware level more or less doubles at the level of infrastructure. Increasing energy consumption at the hardware level for example automatically means increasing demand for cooling. Therefore measures operating at the source of energy consumption in data centres respectively at the server hardware level are effective in two ways.

Many measures can be applied to improve energy efficiency of IT infrastructure thereby addressing cooling, lighting, energy supply etc.. Options for optimisation for example include the use of efficient infrastructure hardware components (e.g. compressors), airflow,
temperature and humidity level in the data centre as well as the use of free cooling.

Infrastructure efficiency of many data centres could be improved by 50% and more. However, due to the relatively long service life of infrastructure equipment many options for optimisation often can not be realised in the short term.

**Best practise cases**

A major part of the project was dedicated to the conduction of best practise cases showing the successful application of energy efficient technology in practise. For that purpose several case studies have been documented covering small, medium size and large enterprises in the service sector. The cases included energy efficient hardware solutions, consolidation and power management. The consolidation approaches covered pure hardware consolidation, server virtualisation and consolidation on blades.

Overall the cases showed that energy savings of 60% can be achieved on average based on already well established technologies. Concrete savings in individual projects depend on the initial situation and the specific technology applied. Consolidation approaches may allow energy savings of more than 70-80%.

On the other hand the cases showed that measures for hardware renewal often involve an increase in performance and capacity of the IT systems. New services and growth of the IT system often cause additional requirements. The additional capacity and performance tend to partly compensate energy savings. While for larger IT systems energy savings mostly are still high they may be completely compensated in case of small systems.

Overall experiences from best practise show that adequate analysis and planning is an important prerequisite to exploit efficiency potentials to a maximum. The most promising approach is to consider options to improve energy efficiency as part of the standard renewal of hardware and infrastructure. This allows addressing management issues and issues of energy consumption and costs in combination. Implementation of measures for the purpose of reducing energy consumption and energy costs only will often be much more complicated, may not be accepted by responsible staff and may often be not economic.

**Development of energy efficiency criteria for servers**

Energy efficiency criteria for servers have been have been developed just recently by SPEC and within the Energy Star programme. The E-server project has contributed significantly to the development of the new Energy Star criteria as one of the major group of experts involved at EU level.

The first set of Energy Star criteria for servers has been finalised in May and published recently. It is focussed on the product segment of volume servers (up to 2 CPUs) and addresses energy efficiency of power supplies as well as efficiency of servers in idle mode. Thus the criteria currently can be primarily applied for volume servers used for low load situations. Larger servers as well as equipment used for consolidated systems and significant loads are not covered for the moment but shall be considered for the second version of requirements expected to be available in 2010. The current set of criteria is limited regarding application and to be seen as a first step on the way to a more comprehensive approach.

The intention of Energy Star is to implement a criterion or benchmark which covers on active power under real workloads. The criterion may be based on the SPECpower benchmark if a sufficient number of relevant workloads is addressed by SPEC.

**Development of instruments to support market development for energy efficient technology**

Besides best practise cases and energy efficiency criteria the project involved the development of several other instruments to support the market for energy efficient technology. Among other procurement guidelines were developed supporting IT and infrastructure
managers in the selection of energy efficient hardware and infrastructure equipment. The guidelines have been published as separate brochures which can be downloaded from the projects website.

The evaluation of leasing and guaranteed saving concepts as another potential supply side tool to stimulate the market for energy efficient technology has not led to concepts which are accepted and can be implemented in the market today. Discussions and evaluation with representatives from the target groups have shown that there is currently no demand for such concepts. The following points indicate major barriers to the application of contracting like concepts in the server and data centre market today

- complexity of contracting concepts
- grouping different service components such as consulting, planning, finance, IT management, business drivers
- acquisition efforts and consulting services are disproportionately high
- contract duration is usually longer (8+ years) than the normal amortisation period for IT hardware (3-5 years).

Broader concepts involving complete IT infrastructure including cooling and energy supply may offer new options for the future.
Energy consumption and saving potentials for servers and data centres in Europe

2 Energy consumption and saving potentials for servers and data centres in Europe

2.1 Methodology

2.1.1 Electric power consumption

The annual electric power consumption of servers in Europe was estimated based on market data from IDC and energy consumption data of popular server models available from manufacturers. The concept was similar to the approach of Koomey (2007) however with several modifications regarding data base and sample size. A detailed calculation was done for Western Europe (E-15 plus Switzerland). For Eastern Europe (other countries of EU-27) a somewhat simplified approach has been applied reducing both the number of dominant models considered and the segmentation of the market.

The general market segmentation used by IDC first of all allows to distinguish market segments of volume servers (<25 000 US$), mid-range servers (>25 000 US$, <500 000 US$) and high end servers (>500 000 US$). Volume servers are the dominant market segment in terms of installed numbers and energy consumption. Due to this fact the second level segmentation from IDC was used splitting volume servers into four sub-segments (<3000$, 3000–6000$, 6000–10 000$ and 10 000–25 000 US$). Overall 6 categories or price bands were analysed separately. The estimations for the Eastern European market were done for the aggregated market segments (volume servers, mid-range servers and high end servers) only.

The following sources of market data were used for the calculations:

- **Data source 1**: IDC installed base per model, manufacturer and price band for Western Europe 2000–2005
- **Data source 2**: IDC-shipments per model, manufacturer and price band for Western Europe 1996–2006
- **Data source 3**: IDC shipments per model, manufacturer and price band for Central and Eastern Europe 2001–2006

The installed base for Western Europe in 2006 was calculated by a combination of data source 1 and 2. The calculations for the rest of EU-27 market were based on data set 3.

To calculate electric power consumption for servers in Western Europe the 10 most popular server models per price band have been
analysed for 2006 and the 5 most popular models where considered for 2003. For the estimation of the electric power consumption of servers for the remaining EU-27 countries available data on shipments 2001–2006 were used.

Furthermore it was assumed that the typical service life of servers is 5 and 6 years for volume respectively mid-range and high end servers.

Several sources of information had to be used to assemble the information necessary for the calculation of the electric power of the dominant server models in the market. Energy consumption of a specific server model may vary strongly with configuration and depending on workload. Average values indicating the typical energy consumption of a standard configuration under a typical workload so far are not reported in specifications or datasheets of manufacturers.

The project partners SUN and IBM provided data on typical energy consumption for their dominant server models. Fujitsu Siemens Germany also provided estimations of the typical energy consumption of popular models. For the popular models from HP and DELL online configurators were used to calculate maximum measured power. This was done in two ways. The first approach followed the concept of Koomey (2007), specifying a maximum hardware configuration including maximum number of processors at fastest speed, maximum amount of RAM, largest size and number of disks, n+1 power supplies and processor intensive workloads. The energy consumption derived for this maximum configuration was multiplied by a factor 0,4 for volume servers and 0,66 for mid-range servers and high end servers.

In a second approach medium configurations were specified for a number of models trying to approximate real situation more closely without finally using correction factors. Furthermore workloads were set to an average or 30 % level rather than a power intensive maximum level. The results were compared to the previous method to assess the sensitivity of results depending on assumptions.

Koomey (2007) proposed to use conversion factors of 0,25, 0,3 and 0,4 for volume, mid-range and high-end servers in those cases where rated input power of models is the only information provided by manufacturers. We found that using the 0,25 factor for volume servers can lead to significant underestimates of the typical power consumption for many models. We only used the 0,4 factor for some cases in the high end server segment.

To finally calculate the overall energy consumption of servers based on the per unit power data the following definitions were made. Continuous 365 days/a operation was assumed. Energy consumption of the most popular server models in terms of installed base was weighted according to the market share of the models.

Furthermore some assumptions had to be defined for the calculation of energy consumption at the data centre level. For the estimation of energy consumption of storage and network components as well as infrastructure, no detailed analysis has been done. Calculations have been based on studies by Greenberg et al. (2006), EPA (2007), LBNL (2006). According to these reports the power usage effectiveness (ratio of the power usage of the data centre to the electricity use of the IT hardware) of data centres on average is approximately 2. For every watt of direct IT power an additional watt is needed for cooling, UPS and lighting etc. Furthermore based on assumptions in the literature cited above energy consumption of storage and network equipment was estimated to be in the order of magnitude of 10 % of total data centre electric power consumption. We used these values for the calculation of the overall electric power consumption of data centres.

2.1.2 Energy saving potentials

The calculation of energy saving potentials was based on technical potentials and expected degrees of implementation outlined in Part 2 of the report. These technical potentials
and market potentials were combined with market forecasts done by IDC for the years 2007–2011. The following three scenarios where considered:

1. **Business as usual**

   This scenario is based on the market development expected by IDC. It includes consolidation by virtualisation assuming a factor of 5 but expecting only low degrees of implementation for power management and energy efficient hardware. The increase of energy efficiency by use of highly efficient hardware was set to 25 % for the years 2007–2011, with implementation levels ranging from 5 to 25 %. Efficiency gains by power management of CPUs and disks were set to 20 respectively 5 % with degrees of implementation between 5 and 30 % from 2007 to 2011. For the estimation of total electric power consumption at data centre level a ratio of total electric power consumption to IT power consumption (PUE)\(^1\) of 2.0 was used which is a typical average value currently found in practise.

   This scenario is likely to occur if business as usual will continue without significant measures to support energy efficiency.

2. **Moderate efficiency scenario**

   This scenario is also based on the IDC market trend but involving higher degrees of virtualisation and of implementation of efficient hardware and power management. The virtualisation factor ranges from 5,4 to 7 and implementation of energy efficient hardware and power management goes up to 70 % in 2011. In addition a higher short term market penetration is expected. The total number of shipments for standard servers and logical servers was expected to be the same as in the basic IDC scenario. For the estimation of total electric power consumption at data centre level the PUE was expected to decrease from 2,0 in 2007 to 1,7 in 2011 indicating an increase of efficiency due to for example improved efficiency of cooling hardware and design.

   This moderate efficiency scenario can be expected to occur if a moderate set of measures supporting the trend of virtualisation, use of efficient hardware as well as power management at the CPU and disk level and improvement of some components at infrastructure level is implemented (e.g. procurement guidelines etc.).

3. **Forced efficiency scenario**

   This scenario again is based on the IDC forecast trend but involving a virtualisation factor between 7 and 10, higher energy efficiency of hardware and higher degrees of implementation of efficient hardware and power management. Implementation levels are expected to increase up to 100 % in 2011. The total number of shipments for standard servers and logical servers was expected to be the same as in the basic IDC scenario. For the estimation of total electric power consumption at data centre level the PUE was expected to decrease from 2,0 in 2007 to 1,5 in 2011 indicating an increase of efficiency due to improved efficiency for example of cooling hardware and design. Highly efficient infrastructures will operate at even higher efficiency indices of 1,2-1,3 however this level will not be reached on average.

   This forced efficiency scenario would require a forced mix of measures and instruments to move the market as for example labelling, procurement guidelines, financial incentives, minimum standards etc.

   All three scenarios are based on a couple of common assumptions. First of all a complete retirement of the total volume server stock being in the market in 2006 was assumed to

---

\(^1\) The abbreviation PUE(power usage effectiveness) is proposed and used by The Green Grid
occur within 5 years or by 2011. A retirement proportional to the shipments in 2002–2006 was expected. For this decreasing historic stock no specific measures to increase energy efficiency were considered since it is seen as difficult and unusual to implement efficient hardware components, power management or virtualised solutions on installed respectively older hardware. Consequently measures to improve energy efficiency were applied only to new hardware shipped between 2007 and 2011.

In all three scenarios the average energy consumption of server units in the volume server segment would increase by 6–7% annually. This annual trend of growth corresponds to the development between 2003 and 2006. It is expected that the energy demand of servers will still increase in the future due to increasing performance and configuration of units. This trend will go in parallel to a possible increase of energy efficiency. Thus increased efficiency will partly counterbalance the growing energy demand of server units.

It was furthermore expected that the trend to virtualisation in the volume server segment will be largely based on machines of the upper price band (price band 4), offering enough capacity for virtualisation but requiring more power than entry level servers. For that reason it was assumed that the average power of virtualised servers is typically in the order of magnitude of price band 4. The following table summarizes the specifications of parameters for the three scenarios. More information on the model is provided in chapter 2.2.
Table 2.1 Summary of main parameters used in the three future scenarios. The table shows the setting of parameters for the different energy efficiency options consolidation, hardware efficiency, power management and IT-infrastructure from 2007–2011.

<table>
<thead>
<tr>
<th>IDC Business as usual</th>
<th>Hardware Efficiency</th>
<th>Power Management</th>
<th>IT-Infrastructure</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Share of virtual servers shipped</td>
<td>Average consolidation factor</td>
<td>Shipments of efficient Hardware</td>
</tr>
<tr>
<td>2007</td>
<td>7%</td>
<td>5.4</td>
<td>5%</td>
</tr>
<tr>
<td>2008</td>
<td>15%</td>
<td>5.2</td>
<td>10%</td>
</tr>
<tr>
<td>2009</td>
<td>33%</td>
<td>8</td>
<td>15%</td>
</tr>
<tr>
<td>2010</td>
<td>45%</td>
<td>4.9</td>
<td>20%</td>
</tr>
<tr>
<td>2011</td>
<td>52%</td>
<td>4.8</td>
<td>25%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Moderate Efficiency Scenario</th>
</tr>
</thead>
</table>

|                       | Share of virtual servers shipped | Average consolidation factor | Shipments of efficient Hardware | Energy saving potential per unit | Shipments with CPU power management enabled | Energy saving potential per unit | Shipments with HD power management enabled | Energy saving potential per unit | PUE |
| 2007                  | 7%                  | 5.4              | 5%               | 25%               | 5%               | 20%               | 5%               | 5%               | 2.0  |
| 2008                  | 15%                 | 5.4              | 15%              | 25%               | 20%              | 20%               | 10%              | 5%               | 1.9  |
| 2009                  | 31%                 | 6                | 30%              | 25%               | 30%              | 20%               | 20%              | 5%               | 1.6  |
| 2010                  | 35%                 | 7                | 30%              | 25%               | 30%              | 20%               | 20%              | 5%               | 1.5  |
| 2011                  | 43%                 | 7                | 40%              | 25%               | 35%              | 20%               | 25%              | 5%               | 1.7  |

<table>
<thead>
<tr>
<th>Forced Efficiency Scenario</th>
</tr>
</thead>
</table>

|                       | Share of virtual servers shipped | Average consolidation factor | Shipments of efficient Hardware | Energy saving potential per unit | Shipments with CPU power management enabled | Energy saving potential per unit | Shipments with HD power management enabled | Energy saving potential per unit | PUE |
| 2007                  | 5%                  | 7                | 10%              | 25%               | 5%               | 20%               | 5%               | 5%               | 2.0  |
| 2008                  | 12%                 | 7                | 20%              | 30%               | 20%              | 20%               | 20%              | 5%               | 1.8  |
| 2009                  | 21%                 | 10               | 30%              | 40%               | 30%              | 20%               | 30%              | 5%               | 1.7  |
| 2010                  | 25%                 | 10               | 45%              | 45%               | 40%              | 20%               | 40%              | 5%               | 1.6  |
| 2011                  | 30%                 | 10               | 50%              | 45%               | 50%              | 20%               | 50%              | 5%               | 1.5  |
2.2 Results

2.2.1 Electric power consumption

Fig. 2.1 shows the overall electric power consumption of data centres in Western and Eastern Europe in 2006. In Western Europe (EU-15 and Switzerland) a total of 6.77 million servers were installed with a related energy demand of 14.7 TWh/year. Stocks and energy consumption in the Eastern European countries only amount to roughly 7% of EU-27 or 0.49 million servers and 1.1 TWh.

Based on findings by Greenberg et al (2006) and LBNL (2006) it was assumed that typically 10% of the electric power consumption in data centres are due to storage and network equipment and 50% of total data electric power consumption in data centres are due to server infrastructure (like cooling, lighting and UPS). Adding these amounts to the direct consumption of the server hardware leads to a total annual data centre electric power consumption of 36.9 TWh in 2006 for Western Europe and 39.6 TWh for total EU-27.

A comparison of these findings with recent results from the US Environmental protection agency (EPA report to congress 2007) shows that the electric power consumption of EU-15 is in the order of magnitude of 60% of the consumption in the US. This figure corresponds well with the current installed base of servers. According to calculations based on IDC data the current stock of servers in EU-15 in 2006 is about 63% of the US stock.

![Electric power consumption of data centres in 2006](image)

**Fig. 2.1 Electric power consumption in data centers in EU-27 in 2006.** The chart shows the percentage of server hardware (volume, mid-range, high-end), storage and network equipment and infrastructure (cooling, UPS, lighting etc.) of total power consumption.
Fig. 2.2 Share of electric power consumption in different price bands in EU-27 in 2006. The figure shows the shares on total power consumption for the four price bands within the volume server segment as well as for the mid-range and high-end segment.

Fig. 2.2 and Tables 2.2-2.3 show the typical energy consumption of server models in the market segments of volume servers (<25 000 US$), mid-range servers (25 000–500 000 US$) and high end servers (>5 000 000 US$). The volume server segment has been further split into four subcategories respectively price bands according to the IDC database.

The low end price band within the volume servers segment (<3000 US$ respectively entry level servers) accounts for 35 % of the total stock and 24 % of the total electric power consumption. The weighted average power of the most popular servers in this class amounts to 170 W. The second price band (3000–6000 US$) accounts for 50 % of the total stock and 40 % of the total electric power consumption. The weighted average power of the most popular servers in this class amounts to 200 W. The third price band (6000–10 000 US$) accounts for 5 % of the total stock and 5 % of the total electric power consumption. The weighted average power of the most popular servers in this class amounts to 216 W. Finally the fourth price band (9000–25 000 US$) within volume servers accounts for 6 % of the total stock and for 10 % of the total electric power consumption. The weighted average power of most popular servers in this class amounts to 411 W.

Overall the market segment of volume servers accounts for 96 % of the stock and 78 % of the total electric power consumption.
Energy efficient servers in Europe

![Graph](image)

**Fig. 2.3 Increase of electric power consumption of servers in Western Europe between 2003 and 2006**

The market segment of “mid-range servers” (price band 25 000–500 000 US$) accounts for 4 % of the stock and 12 % of the total electric power consumption. The average weighted power of most popular models is 771 W.

Finally the market segment of high end servers is below 1 % in terms of stock and amounts to 10 % in terms of total electric power consumption.

The results are in good agreement with a recent study for the US market (Koomey 2007) suggesting that volume servers account for 82 % of total server energy consumption, mid range servers for 9 % and high end servers for 7 %.

Tables 2.5-2.6 show the power consumption of popular server models in 2003. The calculated total electric power consumption was 10.7 TWh. Thus the energy demand of the installed base increased by 4 TWh or 37 % between 2003 and 2006. The total stock in the same period increased by 38 % indicating a close relationship between installed stock and energy consumption. The increase in electric power consumption is largely due to a strong market development in the area of small volume servers (<6000 US$).

Table 2.4 shows the electric power consumption of representative server models in the Eastern European countries within EU-27 for 2006.

Overall these results show the high relevance of the volume servers in the context of the Efficient Servers project. Volume servers (especially entry level servers and servers in price band 2) are the by far largest and fastest growing market segment. In contrast there is some trend to a replacement of mid-range servers by smaller servers.
Table 2.2 Representative server models, electric power consumption and installed base in the volume server segments in 2006 in Western Europe (price band 1: <3000$, 2: 3000–6000$, 3: 6000–10 000$, 4: 10 000–25 000 US$)

<table>
<thead>
<tr>
<th>Brand</th>
<th>Model</th>
<th>Typical power (W)</th>
<th>Weighted average power (W)</th>
<th>Total electric power consumption of price band (GWh)</th>
<th>Total installed base in priceband</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Price band 1</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HP</td>
<td>ML350</td>
<td>231</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HP</td>
<td>DL320</td>
<td>143</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HP</td>
<td>ML330</td>
<td>204</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fujitsu Siemens</td>
<td>TX150</td>
<td>125</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HP</td>
<td>ML310</td>
<td>124</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HP</td>
<td>ML110</td>
<td>136</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IBM</td>
<td>206</td>
<td>140</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fujitsu Siemens</td>
<td>Econel</td>
<td>150</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HP</td>
<td>DL140</td>
<td>202</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IBM</td>
<td>205</td>
<td>136</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Price band 2</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HP</td>
<td>DL360</td>
<td>227</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HP</td>
<td>ML370</td>
<td>246</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HP</td>
<td>ML350</td>
<td>212</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dell</td>
<td>2850</td>
<td>219</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IBM</td>
<td>HS20</td>
<td>142</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dell</td>
<td>2650</td>
<td>145</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IBM</td>
<td>346</td>
<td>155</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dell</td>
<td>1850</td>
<td>187</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HP</td>
<td>BL20p</td>
<td>120</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Price band 3</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HP</td>
<td>DL380</td>
<td>227</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fujitsu Siemens</td>
<td>RX300</td>
<td>176</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SUN</td>
<td>V240</td>
<td>300</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fujitsu Siemens</td>
<td>TX300</td>
<td>185</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fujitsu Siemens</td>
<td>F200</td>
<td>140</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IBM</td>
<td>235</td>
<td>277</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Compaq</td>
<td>DL380</td>
<td>227</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HP</td>
<td>ML530</td>
<td>280</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fujitsu Siemens</td>
<td>P200</td>
<td>140</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IBM</td>
<td>345</td>
<td>155</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Price band 4</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HP</td>
<td>DL580</td>
<td>370</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HP</td>
<td>ML570</td>
<td>407</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SUN</td>
<td>V440</td>
<td>520</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SUN</td>
<td>280R</td>
<td>712</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dell</td>
<td>6850</td>
<td>300</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HP</td>
<td>DL585</td>
<td>358</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IBM</td>
<td>366</td>
<td>655</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HP</td>
<td>DL560</td>
<td>200</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dell</td>
<td>6850</td>
<td>381</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dell</td>
<td>6800</td>
<td>300</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### Table 2.3 Representative server models, electric power consumption and installed base in the mid-range and high-end server segments in 2006 in Western Europe (price band 5: 25 000–500 000 US$, 6: >500 000 US$)

<table>
<thead>
<tr>
<th>Brand</th>
<th>Model</th>
<th>Typical power (W)</th>
<th>Weighted average power (W)</th>
<th>Total electric power consumption of price band (GWh)</th>
<th>Total installed base in priceband</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Price band 5</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IBM</td>
<td>i5-520</td>
<td></td>
<td>500</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SUN</td>
<td>V880</td>
<td></td>
<td>1050</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SUN</td>
<td>V480</td>
<td></td>
<td>1023</td>
<td></td>
<td></td>
</tr>
<tr>
<td>IBM</td>
<td>445MP</td>
<td></td>
<td>521</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SUN</td>
<td>V490</td>
<td></td>
<td>1136</td>
<td>771</td>
<td>1796</td>
</tr>
<tr>
<td>IBM</td>
<td>p5-670</td>
<td></td>
<td>864</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HP</td>
<td>Alpha25</td>
<td></td>
<td>840</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HP</td>
<td>rp5400/rp5450/70</td>
<td></td>
<td>420</td>
<td></td>
<td></td>
</tr>
<tr>
<td>IBM</td>
<td>9406-270</td>
<td></td>
<td>330</td>
<td></td>
<td></td>
</tr>
<tr>
<td>IBM</td>
<td>p650_6M2</td>
<td></td>
<td>700</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Price band 6</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HP</td>
<td>SUPERDOME</td>
<td></td>
<td>8976</td>
<td></td>
<td></td>
</tr>
<tr>
<td>IBM</td>
<td>p690-681</td>
<td></td>
<td>12796</td>
<td></td>
<td></td>
</tr>
<tr>
<td>IBM</td>
<td>p5 596</td>
<td></td>
<td>16102</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SUN</td>
<td>15K</td>
<td></td>
<td>12000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>IBM</td>
<td>p680-S85</td>
<td></td>
<td>1335</td>
<td>9398</td>
<td>1335</td>
</tr>
<tr>
<td>Fujitsu Siemens</td>
<td>PRIMEPOWER 1500</td>
<td></td>
<td>3180</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HP</td>
<td>GS160</td>
<td></td>
<td>2640</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SUN</td>
<td>E-25K</td>
<td></td>
<td>14900</td>
<td></td>
<td></td>
</tr>
<tr>
<td>IBM</td>
<td>9406-840</td>
<td></td>
<td>1584</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SUN</td>
<td>12K</td>
<td></td>
<td>6500</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Table 2.4 Representative models, power consumption, installed base in Eastern European countries 2006

<table>
<thead>
<tr>
<th>Brand</th>
<th>Model</th>
<th>Typical power (W)</th>
<th>Weighted average power (W)</th>
<th>Total electric power consumption of price band (GWh)</th>
<th>Total installed base in priceband</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Volume servers</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HP</td>
<td>DL380</td>
<td></td>
<td>227</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HP</td>
<td>ML350</td>
<td></td>
<td>212</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HP</td>
<td>DL360</td>
<td></td>
<td>187</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HP</td>
<td>ML370</td>
<td></td>
<td>246</td>
<td></td>
<td></td>
</tr>
<tr>
<td>IBM</td>
<td>206</td>
<td></td>
<td>140</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Mid range servers</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SUN</td>
<td>V880</td>
<td></td>
<td>1050</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SUN</td>
<td>V480</td>
<td></td>
<td>1023</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SUN</td>
<td>V490</td>
<td></td>
<td>1136</td>
<td>902</td>
<td>295</td>
</tr>
<tr>
<td>HP</td>
<td>rp3440</td>
<td></td>
<td>600</td>
<td></td>
<td></td>
</tr>
<tr>
<td>IBM</td>
<td>445 MP</td>
<td></td>
<td>521</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>High end servers</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HP</td>
<td>SUPERDOME</td>
<td></td>
<td>8976</td>
<td></td>
<td></td>
</tr>
<tr>
<td>IBM</td>
<td>p5 595</td>
<td></td>
<td>16102</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SUN</td>
<td>E25K</td>
<td></td>
<td>14900</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fujitsu Siemens</td>
<td>PRIMEPOWER 1500</td>
<td></td>
<td>3180</td>
<td>10655</td>
<td>128</td>
</tr>
<tr>
<td>IBM</td>
<td>p690-681</td>
<td></td>
<td>12796</td>
<td></td>
<td>851</td>
</tr>
</tbody>
</table>
### Table 2.5 Representative server models, electric power consumption and installed base in the mid-range and high-end server segments in 2003 in Western Europe (price band 1: <3000$, 2: 3000–6000$, 3: 6000–10 000 US$, 4: 10 000–25 000 US$)

<table>
<thead>
<tr>
<th>Brand</th>
<th>Model</th>
<th>Typical power (W)</th>
<th>Weighted average power (W)</th>
<th>Total electric power consumption of price band (GWh)</th>
<th>Total installed base in priceband</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Price band 1</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HP</td>
<td>ML350</td>
<td>172</td>
<td>132</td>
<td>1331</td>
<td>1149261</td>
</tr>
<tr>
<td>HP</td>
<td>ML330</td>
<td>127</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IBM</td>
<td>205</td>
<td>116</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IBM</td>
<td>200</td>
<td>100</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HP</td>
<td>DL320</td>
<td>98</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Price band 2</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HP</td>
<td>DL360</td>
<td>107</td>
<td>138</td>
<td>2652</td>
<td>2114425</td>
</tr>
<tr>
<td>HP</td>
<td>DL380</td>
<td>140</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HP</td>
<td>ML370</td>
<td>215</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dell</td>
<td>2650</td>
<td>145</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fujitsu-Siemens</td>
<td>170</td>
<td>100</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Price band 3</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HP</td>
<td>DL380</td>
<td>141</td>
<td>163</td>
<td>1155</td>
<td>809736</td>
</tr>
<tr>
<td>IBM</td>
<td>5100</td>
<td>160</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fujitsu Siemens</td>
<td>E200</td>
<td>140</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fujitsu Siemens</td>
<td>F200</td>
<td>140</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IBM</td>
<td>235</td>
<td>277</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Price band 4</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HP</td>
<td>DL500</td>
<td>355</td>
<td>454</td>
<td>2190</td>
<td>505752</td>
</tr>
<tr>
<td>SUN</td>
<td>200R</td>
<td>712</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SUN</td>
<td>220R</td>
<td>610</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DELL</td>
<td>6450</td>
<td>382</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HP</td>
<td>ML570</td>
<td>343</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Table 2.6 Representative server models, electric power consumption and installed base in the mid-range and high-end server segments in 2003 in Western Europe (price band 5: 25 000–500 000 US$, 6: >500 000 US$)

<table>
<thead>
<tr>
<th>Brand</th>
<th>Model</th>
<th>Typical power (W)</th>
<th>Weighted average power (W)</th>
<th>Total electric power consumption of price band (GWh)</th>
<th>Total installed base in priceband</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Price band 5</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SUN</td>
<td>450</td>
<td>632</td>
<td>928</td>
<td>2460</td>
<td>302566</td>
</tr>
<tr>
<td>SUN</td>
<td>V330</td>
<td>1050</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IBM</td>
<td>9406-270</td>
<td>250</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HP</td>
<td>rp 7400</td>
<td>1700</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SUN</td>
<td>V430</td>
<td>1023</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Price band 6</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IBM</td>
<td>p690-681</td>
<td>12796</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HP</td>
<td>SUPERDOME</td>
<td>8976</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IBM</td>
<td>POWERSERVER S80</td>
<td>1900</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SUN</td>
<td>10000HE</td>
<td>10931</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HP</td>
<td>rp8400</td>
<td>2900</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
2.2.2  Energy saving potentials

2.2.2.1  Saving potentials at the level of server hardware

As outlined in section 2.2 energy saving potentials to be achieved by various technical measures have been calculated on the basis of market forecasts from IDC and technical options for savings described in Part 2 of this report.

As it has been shown in section 2.2.1 the server market in the Eastern European market of EU27 is still small accounting only for approximately 7–8 % of Western Europe. According to IDC forecasts the Eastern European market will remain small as shipments will be constantly at a level of 7–8 % of Western Europe. For that reason no separate estimation of the saving potentials for the Eastern European countries within EU27 has been done. Scenarios shown below for Western Europe can not necessarily directly be transferred to Eastern countries where approaches to achieve energy savings in this area may be less comprehensive and slower. However in order to get a simplified indication of the overall saving potentials in EU27 potentials shown below for Western Europe may be increased by 5–8 %.

IDC recently have revised their forecasts for the development of the server market for 2007 to 2011 due to the upcoming trends of server consolidation and virtualisation and a growing market share of multi core models with higher capacity. The trend of virtualisation is expected to slow down the annual increase of shipments of physical servers. According to the recent forecasts the annual growth rate of shipments will be in the order of magnitude of 2,8–4,6 %, compared to growth rates of up to 10 % in previous years. So the already ongoing trend to virtualisation will help to slow down the increase of energy consumption in the next years even under business as usual conditions.

The evaluation of energy saving potentials has been focussed on the volume server segment since the midrange and high end segment are small in terms of share of total stock and energy consumption and furthermore offers not much room for measures to improve energy efficiency. Therefore the development of the mid-range and high-end segment was assumed to be identical in the three scenarios using the trend of the IDC forecast without further modifications.

In the standard business as usual scenario (B.A.U.) the electric power consumption of the volume server segment rises from 12,7 TWh in 2006 to 25,8 TWh in 2011 which is equal to a 103 % increase. Although the trend to virtualisation is already slowing down the market growth an extrapolation of the current trend therefore would lead to a doubling of the energy demand between 2006 and 2011 if no counteractive measures are taken. For the total server market including mid-range and high end servers an increase from 14,7 TWh in 2006 to 31 TWh in 2011 would be expected (Fig. 2.4). For the B.A.U. scenario the same virtualisation rates have been applied as have been forecasted by IDC and very moderate levels for shipments of energy efficient hardware have been assumed ranging from 5 % in 2007 to 25 % in 2011 (Tab. 2.1). Fig. 2.5 shows the shipments of total physical as well as standard and virtualised servers for the different scenarios from 2007 to 2011.

The second scenario (Moderate), introducing moderate measures to support energy efficiency, allows slowing down the annual growth of energy consumption significantly. This results in a maximum electric power consumption in the volume server segment of 19,3 TWh in 2011 and an energy saving of roughly 25 % compared to the B.A.U scenario respectively a saving of energy costs in the order of magnitude of 0,8 billion €. For the total server market including mid-range and high-end servers the energy consumption increases from 14,7 TWh in 2007 to 23,5 TWh in 2011. This is equal to a 60 % increase but to a 24 % energy saving compared to the B.A.U. scenario. In the moderate scenario a higher virtualisation level of 6-7 is assumed as well as higher levels for implementation of energy efficient hardware.
and power management ranging between 5–10 % in 2007 and 70 % in 2011 (Tab. 2.1).

In the third scenario (Forced) electric power consumption of the volume server segment in 2011 is decreasing to 12,6 TWh which would be a 14 % reduction compared to the level in 2007 and a more than 50 % reduction compared to the B.A.U. level in 2011. For the total server market including mid-range and high-end servers energy consumption would increase from 14,7 TWh in 2006 to 17 TWh in 2011. This would correspond to energy savings of 45 % compared to the business as usual scenario in the year of 2011. For this scenario virtualisation factors between 7 and 12 have been assumed and a use of energy efficient hardware and power management of up to 100 % in 2011.

Fig. 2.4 Electric power consumption 2007-2011. Development of power consumption is shown for the BAU (IDC 2007) as well as a moderate and forced energy saving scenario.

In terms of saved energy costs this means an annual cost reduction of 1,7 billion € in the year of 2011 (assuming average costs for electricity of 12 cents per kWh) and total cumulated cost savings of 3,3 billion € between 2008 and 2011. The forced scenario however is likely to be realised only if a mix of strong supporting measures will be driving the market as for example the introduction of an efficiency label, financial incentives, procurement guidelines, campaigns and promotion.
2.2.2.2 Savings at the level of data centres

Energy savings also have been estimated at data centre level. For this purpose no detailed model was used but more general a change of the overall power usage effectiveness (PUE) was assumed which was specified for the three scenarios. The efficiency of storage and network equipment was expected to increase linearly to changes of efficiency achieved for the server hardware.

Since PUE reflects the overall energy efficiency of the data centre infrastructure in relation to the IT-hardware this simplified approach provides a reasonable approximation for the future scenarios. For the BAU scenario the PUE has been kept at the current average value of 2 for all years. In the moderate scenario PUE it is gradually decreasing to 1.7 by the year of 2011. For the forced scenario a decrease to 1.5 by the year of 2011 was expected.

Fig. 2.6 shows the three scenarios. In the business as usual scenario the electric power consumption would rise from 36.9 TWh in 2006 to 77 TWh in 2011 which corresponds to a 110 % increase. In the moderate savings scenario energy demand only increases to 49.9 TWh in 2011 suggesting a 35 % saving compared to the business as usual scenario.
Scenarios of electric power consumption of data centres in Western Europe 2007-2011

Fig. 2.6 Electric power consumption of data centres in Western Europe in 2007–2011 for the scenarios “BAU”, “moderate savings” and “forced savings”.

Annual energy costs for data centres in the different scenarios

Fig. 2.7 Costs for electric power in data centres in Western Europe in 2007–2011 for the scenarios “BAU”, “moderate savings” and “forced savings”.

In the forced scenario electric power consumption would drop to 31.9 TWh which is a 13.5 % decrease compared to the 2006 level and a 58 % decrease compared to business as usual in 2011. These results indicate the additional saving or efficiency potential in the area of infrastructure and especially cooling.

In terms of saved energy costs the forced scenario would lead to an annual cost reduction of 5.5 billion € in the year of 2011 and total
cumulated cost savings of 12.1 billion € between 2008 and 2011.

2.2.2.3 Comparison of the results with scenarios for the US

EPA (2007) defined four scenarios to describe the energy saving potentials for the US server market. Calculations were also based on forecast data provided by IDC. However besides this common source for the definition of the basic pattern of market development the assumptions used in the US study differ from the approach in the Efficient Servers project in several respects. However the trends resulting from the two different approaches are similar in the long term perspective of 2011 suggesting overall similar energy saving potentials. EPA defined three scenarios in addition to the baseline (current trend according to IDC) which have been called “improved operation”, “best practise” and “state of the art”.

According to the results of the US study energy demand of US data centres would decrease by more than 21% in the “improved operation” scenario, by 55% in the “best practise scenario” and by almost 70% in the “state of the art” scenario. So the overall savings forecasted for 2011 seems to be in good agreement with the trend calculated in the two scenarios for EU27 (35–60%). However the short term trend for 2008–2009 differs strongly between the US and the EU approach. Here the US scenarios seem very optimistic in their assumptions regarding quick wins. For example the following parameters have been specified for the best practise and the state of the art scenario:

- 100 % enabling of power management on servers where applicable (all scenarios). This is very unlikely to occur in the short term.
- “Energy efficient” servers represent 100 % of volume server shipments 2007 to 2011 (best practise and state of the art). This is also a rather unrealistic goal in the short term.

These assumptions lead to a disruptive trend starting with a sudden increase of energy efficiency already in 2007. In the Efficient Servers project also the forced scenario starts at only moderate levels for hardware efficiency and enabled power management of 10-20%.

Overall both the US and the EU scenarios suggest high energy savings to be by achieved by 2011 if a mix of supportive measures is applied.
3 Current market barriers limiting the use of energy efficient server technology

3.1 Introduction

To characterise the typical procurement and use of servers as well as current market barriers for energy efficient server technology a small market survey was conducted. The objective of the survey was not a broad quantitative investigation but an analysis of procurement and management behaviour and demand side trends based on a limited qualitative sample of interviews in companies. About 30 companies and institutions in Austria and France were involved.

The sample comprises companies and institutions, covering SMEs with 50–250 employees resp. 5–15 servers up to several thousand employees and a few hundred servers.

The following types of institutions were addressed:

- NPO/NGO: Chamber of Labour Vienna, WWF, Austrian Consumer Association
- Energy utility: BEWAG, Energy AG, Total Oil Distribution
- Research: CEMEF, CEP Public Laboratories, Alpheeis, MD3E Private Consultants
- Trade and Retail: Auchan General, Quicksilver Clothes, Cdiscount e-commerce

The questionnaire used as a basis for the interviews is provided in the Annex.
3.2 General results from the survey

Typical services provided by the server IT

Companies and institutions addressed in the survey mainly host their servers internally, outsourcing of servers or services is only used to a small extent. The following IT-services were dominant in most cases:

- Collaborative (Email, Workgroup)
- IT Infrastructure (File and Print, Networking, Proxy Caching, Security, Systems management)
- Web Infrastructure (Web serving, streaming media)
- Business processing (Enterprise resource planning – ERP, Customer relationship management – CRM, Online transaction processing – OLTP, Batch)
- Decision support (data warehouse / data mart, data analyses / data mining)

Typical hardware used and criteria for the selection of hardware concepts

The majority of all companies run rack based servers (mainly servers with 2 rack units (RU) and higher), some smaller companies also use PC derived servers. Blade servers so far are of minor relevance.

According to answers from IT experts the current IT system often is the result of a long term evolution. It is a common objective to consolidate hardware towards more homogeneous systems, being advantageous for operations and maintenance. Many companies seem to follow a dual vendor strategy.

Typical service life of server hardware

Most Austrian companies involved in the survey regardless of size and branch stated that the typical service life of their servers is 5 years. Few companies seem to use their hardware only 4 years or more than 5 years. It is assumed that the probability for any hardware failure rises after a service life of 5 years.

The typical service life of the servers in France according to the survey is 3 to 5 years. The service life is also partly driven by warranties provided by manufacturers. IT experts are facing distinctly increasing costs for service contracts after a three years period.

Estimation of the energy consumption

Discussing the energy consumption of the server systems about two thirds of the interview partners in Austria were able to roughly estimate the order of magnitude of the energy consumption for their server systems, mainly based on the data for the dimensioning of the UPS. However with a single exception no company had so far really considered and monitored energy consumption which has not yet been a relevant management or procurement criterion.

In the French survey only 25% of the interview partners could give a rough estimation of the IT related energy consumption. Energy consumption is generally not addressed and monitored.

Energy consumption related to network infrastructure as well as air conditioning was rarely known. However this is partly due to the fact that often specific staff is responsible for the IT infrastructure which has not been addressed in the survey.

Overall the interviews showed that the knowledge and concern about energy consumption of IT hardware and infrastructure currently is low.

Share of TCO in energy costs

Many of the IT experts stated that they don’t consider energy costs within TCO but rather focus on hardware, software, service costs and manpower costs for maintenance. There was no commonly accepted definition for TCO in this context. Consequential one third of the experts could not provide any rough assessment at all, one third assumed that energy costs would amount to 10–30 % of total TCO. A few experts assumed that this share could rise up to 30–60 %. Regarding the French situation TCO if used at all never included energy costs.
Potential for energy savings

IT experts were asked about the potential for energy savings they do expect for highly energy efficient server solutions compared to standard hardware. More than 50% of the experts in Austria expected saving potentials of 10–50%. Some experts did not provide an estimation or made extreme estimations below 10% or above 50%.

In the French survey saving potentials between 10–25% were expected.

However in general it was underlined that energy consumption has not been a relevant topic or criterion in the past.

Primary options for energy efficient server systems

IT managers were asked for their opinion on options for optimization of energy efficiency of servers. Companies most frequently mentioned the following potential measures:

- Consolidation and virtualisation
- Efficient CPUs and hard disks
- More efficient HW design (e.g. Blades)
- Power Supplies (PS): Efficient PS, reduce redundancy of PS, central PS instead of server PS
- Air conditioning (AC): New concepts for AC, AC rack based instead of data centre cooling, free cooling

Procurement guidelines and procurement criteria

Companies were asked which of the following aspects are covered by their procurement guidelines or criteria.

- Hardware procurement costs
- TCO
- Systems’ energy consumption for the entire service life
- Long term security regarding system maintenance
- Service level agreements

The sequence of priorities identified in Austria was hardware procurement costs followed by total cost of ownership and long term security regarding system maintenance and finally service level agreements. TCO as used here by IT experts excluded a consideration of energy costs. The system energy consumption was never considered as a specific criterion. On the other hand IT experts mentioned flexibility and high availability as important criteria.

According to the results from France procurement guidelines include mainly the hardware procurement cost, long term maintenance or services agreements, but very rarely TCO. The technical performance of the systems is ranked as a highly relevant criterion. It was emphasised that technical solutions to save energy have to be assessed in terms of possible decrease of the performance (e.g. some concerns regarding virtualization vs. performance).

Vendor-independent product information scheme and tools supporting the procurement of efficient servers

IT experts were asked what kind of tools to support energy and cost savings they would appreciate in the future. In general a vendor-independent product information scheme for efficient products (e.g. label, benchmark) would be welcome. Furthermore many companies would appreciate supporting tools like guidelines for procurement or a data base for efficient hardware.

3.3 Evaluation of demand side barriers

The following evaluation is based on further results from the survey and on experiences from industry.

Several barriers identified in the discussions have been clustered into the categories awareness and information, structures and processes in companies as well as cost concepts. Some barriers are cross-linked or interdependent to others.
Energy efficient servers in Europe

3.3.1 Lack of awareness

The majority of companies stated that energy efficiency has not been considered as a relevant criterion so far and awareness of energy saving potentials has been quite low.

It has been mentioned that more recently energy efficiency of servers and data centres has become a topic in the media but transparent and easy to apply efficient solutions still seem to be limited. It has been observed that vendors currently use energy efficiency in marketing strategies but this trend is mainly regarded as a promotional activity. There was the impression that energy efficiency is just at the beginning of the process to be transformed from marketing messages to customer interest and finally practical usage.

Many experts have the impression that the hardware currently on the market is rather similar in terms of energy efficiency, not offering much potential for energy savings.

3.3.2 Lack of information

Energy efficiency data at product level

IT managers generally emphasized a significant lack of information regarding the energy efficiency of server hardware currently on the market. So even if they aim at purchasing energy efficient equipment, they do not get the relevant product information.

TCO considerations have to be based on reliable product data. In this context responsible experts are seeking for standardized information and common definitions for energy efficiency.

Monitoring of energy consumption

Tools for the monitoring of energy consumption at different levels in the data centre are rarely implemented. Normally only one central electricity meter is installed. In most cases no detailed data on energy consumption is available to be correlated with performance parameters and workload. In other words the necessary tools enabling IT managers to effectively monitor the server related energy consumption are generally lacking.

Demonstration and best practise

More complex and demanding options supporting energy efficiency in data centres as for example consolidation by virtualisation impose significant changes to structure and processes. Such changes are always assessed regarding risks. To proceed with a known and proven architecture often is the preferred safe option. Consequently potentials for energy savings are not explored since the priorities of the IT department are on business requirements (e.g. service levels, performance, availability).

It was a general finding of the Austrian and French survey that many virtualization projects have been planned already but not yet have been conducted. This also reflects some reluctance regarding new technologies and indicates a lack of best practise cases demonstrating successful energy efficient solutions in practise.

Currently well documented best practice cases demonstrating strategies for energy efficiency in practise are rare. At the level of data centre facilities to date no complete demonstration cases also addressing energy efficiency of IT hardware do exist.

From this perspective best practices and demonstration cases are crucial to convince enterprises of the feasibility of theoretical concepts, of potential cost savings and of the compatibility of energy efficiency with central business demands.

3.3.3 Shortcomings in structures and processes

Lack of procurement guidelines and traditional buying behaviour

Energy efficiency is so far not included as a criterion in most procurement processes and guidelines in enterprises. IT experts tend to maintain IT systems as homogenous as possible because of service and maintenance requirements. Renewal projects are often addressing subsystems. So changes of hard-
Current market barriers limiting the use of energy efficient server technology

ware concepts to increase energy efficiency may be out of the scope. The missing consideration of energy efficiency as a procurement criterion also may be related to the fact that it is currently almost impossible to evaluate energy efficiency of the equipment offered on the market (see below).

**IT Management and Facility Management are different worlds**

IT managers are generally not informed about the energy consumption of the equipment they are responsible for nor about the related costs. This is mainly due to the fact that they don’t have to include these costs into their budget. Consequently IT-specific energy costs are often not considered and reported.

The departments responsible for the IT hardware on one hand and for the data centre facilities and infrastructure (e.g. energy supply, cooling, lighting etc.) on the other often are “different worlds with different goals”. Although facility managers are responsible for the overall operation and costs of the infrastructure, consideration of energy consumption and energy costs is rather limited. The focus is clearly on providing a high capacity infrastructure allowing the data centre to be run without technical problems. Only recently the awareness has increased that the demand for cooling can reach upper limits where it can not be served anymore within the technical boundaries of a given infrastructure. It has become clear that the efficiency and effectiveness of cooling can be optimised. However this is still a consideration of a part of the saving potentials excluding the major saving potentials at the source respectively at the IT hardware level. Very rarely IT managers themselves feel a responsibility to consider energy efficiency in the data centre although they are at the source of energy demand.

Communication between IT and infrastructure departments often is limited. In many cases the only occasion to discuss requirements between IT and infrastructure departments is at the stage of planning of new server facilities.

Independent of company size it seems as a general rule that systems remain largely unchanged until the cooling facilities reach their limits. Energy efficiency measures in many cases are conducted only if the infrastructure reaches its upper capacity.

**Over sizing of capacity**

Many data centres require high availability and concepts to ensure the maintenance of service levels. Especially for larger facilities this implies considerable standby losses and unused server capacity. Modern concepts of scalability would allow for new architectures with much less headroom to be provisioned for the case of failure. However migration respectively change towards these concepts is seen as additional costs and risks.

**3.3.4 Limited consideration of costs**

**Shortcomings regarding Total Cost of Ownership**

As already discussed above the consideration of costs for IT projects often does not include energy costs for the operation of the equipment. Regarding Total Cost of Ownership different definitions are applied depending on the aspects included (hardware, manpower resp. support, etc.). As already stated above energy efficiency is rarely considered as a part of TCO.

The improvement of energy efficiency at data centre level or hardware level often leads to higher initial costs. However these additional costs are difficult to be justified due to the lack of an overall TCO perspective - even if the calculated pay back period is only few months. Consequently it is less risky for vendors to offer less expensive relatively inefficient solutions than to try to convince customers that higher initial costs will pay off through energy savings.

**Lack of information and motivation for service providers**

A large number of small volume servers – mainly responsible for the current overall
increase of server energy consumption – is used by service providers in the area of IT solutions to SMEs and private consumers. Although service providers normally pay for the electricity costs needed for operation of the equipment they are often not aware of efficiency potentials and consequently install and use relatively inefficient low price equipment.

**Budget allocation – fixed versus volatile budgets**

A specific aspect encountered in data centres is the discussion of fixed vs. volatile budgets. Investigations by companies specialized in optimizing energy usage show that the level of energy consumption in most data centres is nearly constant all over the year. One of the reasons for that is the typical budget planning process of enterprises. There are only few and partial concepts to handle volatile spending – for example depending on load or usage – versus a fixed planned budget. Consequently incentives for facilities and IT operations to achieve energy savings are oft missing.

**3.4 How to overcome the barriers?**

**3.4.1 Awareness raising and education**

*Awareness raising at top management level (CEO, CIO, CFO) and campaigns to sensitize all relevant target groups*

IT managers often are not in a position to take central decisions and thus may not be able to support energy efficient strategies independently. Therefore it is very important that energy efficiency is also perceived and supported as a relevant goal at the higher management level. Consequently awareness raising is not only needed at the operational level but at the higher management level as well.

In order to sensitize target groups to the issues of energy efficiency, campaigns and measures for awareness raising crucial.

The most relevant target groups are IT managers (CIOs), consultants, vendors. Ongoing initiatives like the *Climate Savers Initiative* and the *Green Grid Initiative* give good examples for initiatives and campaigns.

**Education of experts**

The education of IT managers about energy efficiency is just about to be started. It is necessary to

- present potentials and opportunities
- provide examples
- explain risks and safe options
- offer workshops
- disclose reference implementations
- identify problems and obstacles

**3.4.2 Information on energy efficient equipment and data centre solutions**

**Information tools**

Information on energy efficiency of products is crucial for IT-managers and staff responsible for hardware procurement. Consequently information tools such as databases offering information on efficient equipment as well guidelines and recommendations for energy efficient efficient procurement are important. To be effective such tools shall be easy to use and up-to-date.

**Labelling Schemes**

Labelling schemes currently under development as for example the Energy Star requirements for servers may act as a powerful tool supporting IT experts in decision making and purchasing of energy efficient equipment.
Benchmarking

Benchmarking concepts at data centre level considering energy consumption in relation to workload and services may be helpful for a self assessment of overall efficiency and providing a basis for a further in-depth analysis.

Best practice cases

Best practice cases demonstrating the applicability of energy efficiency concepts in practice will encourage IT managers to consider the existing saving potentials and technical options. Neutral bodies shall support best practice / benchmark programs to share experience and facilitate best practice efforts.

Research and development

The already known options regarding energy efficient solutions as for example efficient hardware and software, innovative cooling concepts etc. is still not exhaustive. There is also a need for further research and development as newly emerging technologies are also likely to at least partly increase power demand. In this context it is necessary to support research and development in companies as well as in scientific bodies.

3.4.3 Improvement of structures and processes

Split incentives and competences

As outlined above different goals and responsibilities in the IT departments and the facility management departments often hamper effective solutions and pose a significant structural problem at company level caused by split incentives. Measures at the upper management level are necessary to solve this problem. Common cost effective goals should be set across departments in combination with a reward for energy and cost efficient solutions.

Establishment of efficiency requirements in procurement guidelines

The incorporation of energy efficiency criteria in companies procurement guidelines provides another important basis for energy efficient solutions. Besides that communication and improved to support know how transfer and decisions.

3.4.4 Holistic consideration of costs, financing models and incentives

Transparency of costs

Management should make energy costs transparent across departments and reassign budgets to the relevant business units.

Total cost of ownership concepts

Energy costs should be included in an overall TCO approach. Best practice examples should be considered.

Third party financing models

Third party financing concepts could help to overcome some cost issues. This option needs to be further investigated in order to develop widely accepted solutions in the long run.

Incentive programs

Incentive programs run by governments or energy utilities may help to accelerate the market transformation process. Examples are already available in the US (California Pacific).
Chapter 5 is dedicated to an analysis of technical options to improve energy efficiency both at the level of servers and whole data centres. Therefore the spectrum of options considered covers efficient server hardware, power management, consolidation by virtualisation as well as efficient infrastructure and cooling. Several findings of this chapter have been used for the definition of the scenarios described in chapter two.

4.1 Power Supplies

Power supplies have been subject to investigation in the context with energy efficiency for a longer time. At EU-level a voluntary Code of Conduct has been established to support the market development for energy efficient power supplies.

To assess and compare the power supply and delivery efficiency several factors have to be considered:

- Conversion from AC (240V/120V) to single DC (mostly 12V) or to multi DC (more typical for desktop derived server)
- Conversion from DC to DC as used by the different components (board, CPU, RAM, Fans etc.)
- Resistance of power distribution (cables, on board wiring)

4.1.1 AC-DC Conversion

To discuss power supply efficiency thoroughly, the efficiency over the effective load has to be regarded. Many vendors will soon ensure that more efficient power supplies (> 80 % efficiency ) will be used in their servers systems. There are several industrial [PS07, SSI07, SUN07] and governmental [LBL07] initiatives to increase the efficiency of power supplies above 80 % for all levels of system load.
Energy efficient servers in Europe

Fig. 4.1 Power supply efficiency depending on load. The figure shows a 82% efficiency power supply typical for Intel/AMD server installed in 2006/2007 and higher for the PS of a modern Blade system.

Server systems designers assume typical loads of 40-70%. However, many customers buy rather low configured systems and run them with low utilizations. Assemblers often use overrated power supplies. Thus, the servers may well run idle at 20% of the maximum load of the power supply. Further, many servers are equipped with redundant power supplies, which cut the load in half.

We assume 200W DC for typical Intel/AMD dual socket servers, 5 years life time, factor 2 for data centre cooling & UPS (PUE=2, see Green Grid [GG07]) and 0,14€/kWh power costs.

Table 4.1 shows an example of energy and cost efficiency of different power supplies

<table>
<thead>
<tr>
<th></th>
<th>PS1</th>
<th>PS2</th>
<th>PS1</th>
<th>PS2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>single</td>
<td>redundant</td>
<td>single</td>
<td>redundant</td>
</tr>
<tr>
<td>DC at 20 % [W]</td>
<td>200</td>
<td>200</td>
<td>200</td>
<td>200</td>
</tr>
<tr>
<td>DC at 10 % [W]</td>
<td></td>
<td>100</td>
<td></td>
<td>100</td>
</tr>
<tr>
<td>Efficiency</td>
<td>70 %</td>
<td>86 %</td>
<td>54 %</td>
<td>83 %</td>
</tr>
<tr>
<td>AC [W]</td>
<td>286</td>
<td>233</td>
<td>370</td>
<td>241</td>
</tr>
<tr>
<td>5y costs</td>
<td>€3.504</td>
<td>€2.852</td>
<td>€4.542</td>
<td>€2.955</td>
</tr>
<tr>
<td>Delta</td>
<td>€652</td>
<td></td>
<td>€1.587</td>
<td></td>
</tr>
</tbody>
</table>

As servers are typically operated 24h*7days, the resulting energy consumptions and 5 years costs differ significantly. Especially, in case of redundant power supplies the delta costs may well approach or exceed 50% of the server investment costs. This may be even more pronounced with power supplies providing lower conversion efficiencies especially at the steep slope below 20%.
4.1.2 Energy Savings Potentials

With recent server developments having significantly reduced the idle power, more attention should be paid to the low end part of the conversion efficiency curves of power supplies or options to avoid operating below 20% effective load. Therefore the idle power and potential peak power should be evaluated.

The maximum power of the power supply has to cover the worst case for any system configuration. Therefore an over provisioning of I/O-slots and drive bays increases peak power. There are no technical or administrative hurdles but possible barriers due to higher prices. When installing and replacing power supply energy efficient components should be chosen despite their higher price. With financial incentives energy efficient components might become mass products which lowers the price to an attractive level.

4.2 Voltage Regulators

Current methods for IT Equipment efficiency metrics tend to deal only with power supply efficiency and don’t address or consider the total power system efficiency. Consideration of the total system in IT equipment is important, as the power system may consist of an AC/DC conversion, multiple DC to DC conversions, and various strategies to condition the power to meet the exacting requirements of the IT components. There are many issues with just using power supply efficiency due to the different architectures that can be found in computer products. The IBM paper “End to End Power” [IBM07] investigates four of the most common architectures found in today’s servers, in order to illustrate the complexities and issues involved in assessing total power system efficiency. Bottom line, the system efficiency including all AC/DC and DC/DC conversions (in power supplies and onboard) are rather comparable being typically at 73%.

4.3 Dynamic Power Management

In the following section dynamic strategies are discussed that can be applied by the system software, to bring individual components or the whole server system to an energy-efficient power state. An overview of dynamic power management for server and server clusters is given in [BR04].

4.3.1 Shutdown-Strategies

In underutilized cluster-systems shut-down strategies can be applied. They require a scenario with client applications that can be made state-less (examples are web-front-ends) and that can be re-established within a short period of time. Vary-on-vary-off strategies [PBCH01] conserve power by keeping active a number of servers that suffice to service the current load, and shut the remaining servers down.

Shutdown Strategies save energy by putting some servers into a low-power mode, while a reduced number of servers stays active. If, after some time, the utilization increases, additional servers are brought back from low-power mode, e.g., via wake-on-lan [AMD95] functionality. Parameters for the low-power modes are the time it takes the server to get from the active mode to the low-power mode (1), the time it takes to get from the low-power mode back to the active mode (2) and the amount of power consumed in the low-power mode (3). Parameters (1) and (2) strongly correlate, i.e., if it takes a long time to get into the low-power mode, it also takes a long time to get back to active mode. Opposed to that, there is a negative correlation between parameter (3) and parameters (2) and (3), i.e., the less power the server consumes in a certain mode, the longer it takes to reach that mode.

Typical server systems support three low-power modes: suspend-to-ram, suspend-to-disk and complete shutdown, which can be seen as an extremely power saving mode.

In suspend-to-ram mode (ACPI state S3), the state information of the chip-set, I/O devices, and the state of the processor (i.e., the contents of the registers and the caches) are saved into the system's main memory. Since the caches contain only some megabytes of data at most, getting into and out of suspend-to-ram mode does not take much time. (For
example, getting the Windows Vista compatibility Logo requires a System to get out of S3 in less than 2000ms for desktop systems [St06].) Once its state is saved, the processor can be switched off to conserve power, as can peripheral devices like disks. However, the memory must remain in a mode in which its content is retained, which implies that the RAM must stay powered on. Depending on the size of the RAM, this means power consumptions of up to 20W. Servers designed for remote maintenance and monitoring require an additional power budget of 10W for the embedded service processor.

The state transition requires less than 2 seconds for desktop systems with an HW/operating system setting that is optimized for fast state transition. However entering/leaving S3 mode can last much longer if the state of multiple complex I/O components has to be saved/restored and if the operating system is not optimized for fast state transition as it is in most common server operating systems (Windows XP, Linux).

In suspend-to-disk mode (ACPI state S4), the system's state, which means the state of the processor and the main memory, is written onto hard disk. Since the hard disk is persistent storage, the complete system can be turned off after the state has been saved, so the only power consumed in suspend-to-disk-mode is the power required for wake-on-lan functionality. On re-entering active mode, the state has to be reloaded from the disk into main memory.

However, storing and reloading the complete contents of the main memory to/from disk is a time-consuming process, considering that typical servers are equipped with several gigabytes of memory. Assuming a transfer rate of 100Mb/s and a system with 8Gb of memory, it takes one minute and 20 seconds to suspend to disk.

A complete shutdown (ACPI state S5) of the system means that the system is turned off and consumes no power. However, the system's state is not saved completely when shutting down: Only the file system cache is written back to disk, and all applications and services are shut down, which in turn may write some of their state information to disk. Therefore, shutting down the system can be faster than suspend-to-ram, especially if main memory is large. However, re-starting a system that was powered down means booting the operating system and restarting all applications and services, which can take a long time. Additionally, after the restart, all caches are cold, which means additional penalties when services are first accessed. Slowdowns as high as 20 % have been observed when using vary-on-vary-off with complete shutdown [PBCH01].

### Table 4.2 Transition times and energy demand in different power saving modes

<table>
<thead>
<tr>
<th>Mode</th>
<th>suspend-to-ram</th>
<th>suspend-to-disk</th>
<th>complete shutdown</th>
</tr>
</thead>
<tbody>
<tr>
<td>from active to low-power mode</td>
<td>1 to 2 seconds up to one minute</td>
<td>minutes</td>
<td>~ one minute</td>
</tr>
<tr>
<td>from low-power to active mode</td>
<td>1 to 2 second up to one minute</td>
<td>minutes</td>
<td>minutes, plus penalties from cold caches</td>
</tr>
<tr>
<td>power consumed in low-power mode</td>
<td>5W to 20W (power for service processor is not included)</td>
<td>2W (power for service processor is not included)</td>
<td>2W (power for service processor is not included)</td>
</tr>
</tbody>
</table>

#### 4.3.1.1 Energy Saving Potentials and possible market barriers

Especially for server that show long periods of idleness (e.g., print and file servers idling during the night or weekend) suspend to RAM is a promising approach. Studies in Switzerland show savings of about 50 % for these scenarios often found in small and medium size enterprises [Hu04]. However there are no
studies showing the general applicability of shutdown strategies in a general server environment. Consequently this option offers no general saving potential.

Shutdown strategies require precise planning of shutdown and wake-up procedures. This implies significant administrative overhead. Furthermore the administrator takes the risk of a server that might not wake up at all or with significant delay under some conditions. Testing of shutdown/wake-up procedures in large and heterogeneous server installations with many dependencies is complex because there is no tool support. Therefore the introduction of shutdown strategies is critical and imposes a significant barrier. To date it can not be recommended as a general strategy.

4.3.2 CPU Power Management

Most servers show only a modest CPU utilization. This offers the opportunity for a reduction in clock speed or a temporary operation in halt- or sleep mode. This operation can be supported by batching of requests to increase the average duration of idle times. In the case of multi-core or multi-processor systems the temporary transition of some CPUs/cores into deep sleep mode further reduces the energy consumption and improves the efficiency of the system.

Contemporary processors offer a multitude of mechanisms that trade reduction in computation speed for reduction in power consumptions. These are the hlt-instruction, clock modulation, instruction decode throttling and frequency/voltage scaling.

The hlt-instruction is the most radical method for conserving power, yielding minimal power consumption, but halting computation completely. After executing the hlt-instruction, the processor goes into a power conserving sleep mode, deactivating parts of its functional units. Some processors also reduce frequency and voltage after executing the hlt instruction. The only way for a processor to get reactivated after having executed the hlt-instruction is the arrival of an interrupt.

Almost all of today’s operating systems issue the hlt-instruction when there is no task ready to execute at the moment. After the arrival of an interrupt (e.g., the timer interrupt), the operating system checks if a task has become ready, and otherwise issues another hlt-instruction. Besides its use for saving power when there is no work to be done, the hlt-instruction has also been used to reduce power consumption when there is work to be done, by alternating the execution of a runnable task and of the hlt-instruction. However, this allows only coarse-grained control over power consumption and is not used in practice, since the hlt-instruction suspends the CPU until the arrival of the next interrupt, which can take some milliseconds.

On multithreaded systems, all threads must execute hlt for the processor to be halted. On multi-core systems, typically individual cores can be halted, but a halted core consumes more power if a sibling is active than when all siblings are halted, too. The Intel Core2 E6700, for example, has a thermal design power of 65W. It offers a halt state consuming 20W, which can be entered independently by different cores, and an extended halt state, in which the frequency and voltage is lowered, consuming 12W. However, the extended halt state can only be entered if all cores have executed the hlt-instruction [Int07].

In contrast to the hlt-instruction, clock modulation operates more fine grained. The clock signal gets modulated, i.e., periodically switched on and off for some microseconds (see figure).
Energy efficient servers in Europe

Fig. 4.2 Clock Modulation

Since most of the CPU's power consumption is caused by switching activity, during the time the clock signal is switched off, the power consumption is reduced considerably. The portion of time the clock signal is activated is called the processors duty cycle. Processors like the Intel Pentium 4 or Core2 support duty cycles of 1/8, 1/4, 3/8, 1/2, 5/8, 3/4, 7/8, and 1.

The Pentium 4 and the Core2 use clock modulation to prevent thermal emergencies: If the CPU temperature rises over a predefined limit, the hardware triggers clock modulation (Thermal Monitor 1 [Int06, Int07]). The duty cycle can also be set by software. Linux, for example, offers an interface for user space programs to set the duty cycle, which can be used for more sophisticated temperature control policies. However, it has to be said that clock modulation makes inefficient use of energy: while the clock signal is deactivated, the processor still consumes power (although less than with activated clock signal), but does no work. In multi-core systems, clock modulation can typically be applied to individual cores.

Instruction decode throttling [SKO+97], also called fetch toggling, reduces power consumption by limiting the number of instructions that can be processed per time unit by a certain pipeline stage, e.g., the fetch stage. This leads to other stages being under-utilized, so they can be deactivated for some time. Instruction decode throttling is used in PowerPC processors.

Table 4.3 Transition times, granularity and usage of options for CPU power management

<table>
<thead>
<tr>
<th></th>
<th>hlt-instruction</th>
<th>clock modulation</th>
<th>decode throttling</th>
<th>freq./volt. scaling</th>
</tr>
</thead>
<tbody>
<tr>
<td>transition time</td>
<td>Few microseconds</td>
<td>Few microseconds</td>
<td>Few microseconds</td>
<td>10 to 100 microseconds</td>
</tr>
<tr>
<td>granularity</td>
<td>very coarse, halts until next interrupt</td>
<td>fine, e.g., 8 steps on the Core2</td>
<td>fine</td>
<td>coarse, only two frequencies available on server and desktop processors, finer granularity for mobile processors</td>
</tr>
<tr>
<td>usage</td>
<td>save power during idle times</td>
<td>prevent thermal emergencies</td>
<td>prevent thermal emergencies</td>
<td>save power required for computations, prevent thermal emergencies</td>
</tr>
</tbody>
</table>
In contrast to clock modulation and decode throttling, frequency and voltage scaling offer the potential of actually saving energy. Hence, these techniques were first introduced in mobile processors to lengthen battery time, but have by now also advanced into the server sector. The power consumption of a processor is proportional to its operating frequency and to the square of its operating voltage. By lowering the frequency, the power consumption can be reduced, at the cost of performance being reduced by the same factor. However, operating at lower frequencies also means that the voltage required for driving the processor is lower; the required voltage is roughly proportional to the frequency. Hence, by scaling frequency and voltage, the amount of energy required for a certain computation can be reduced.

However, changing frequency and voltage are complex operations; frequency multipliers have to be adapted, and voltage has to be changed continuously over a period of time. Compared to clock modulation or halting, which can be engaged within some microseconds, frequency and voltage scaling is time consuming: Earlier processors had to be halted completely until the new frequency and voltage was stable; current processors are just halted to change the frequency and can then continue their calculations while the voltage is adapted. While the first P3 mobile processors had to be halted for 250 microseconds when changing frequency, Intels SpeedStep technology used in recent processors halts the processor for less than 10 microseconds [GRA+03].

Processors like the Pentium 4 or the Core2 offer two frequency/voltage combinations, maximum speed and reduced speed. Like clock modulation, hardware makes use of the reduced speed to prevent thermal emergencies (Thermal Monitor 2 [Int06, Int07]), but the frequency can also be set by software. In recent multicore processors, all cores have to operate at the same frequency, but multi-cores that can do frequency and voltage scaling for individual cores have been announced.

The different performance levels established by certain frequency/voltage configurations can be accesses with the help of the ACPI without the need to support these CPU-specific levels by low-level operating system code. Examples are the 6 frequency/voltage combinations of AMD Opteron processor that are mapped to 6 APCI P-states [AMD 07].

Running at lower frequencies can be especially beneficial for memory intensive applications, for which the memory bus is a bottleneck: executing those applications at a lower frequency does hardly harm performance, but saves energy [WB02]. In situations when the CPU is not fully utilized, frequency scaling can also be used to save energy by stretching active phases and reducing idle times [GCW95, FM02].

### 4.3.2.1 Energy Saving Potentials and possible market barriers

For systems with low utilization (e.g., < 20 % utilization) CPU power management can save up to 75 % energy on the CPU level which results in about 20 % saving of the system level [AMD07].

The dynamic management of the CPU power is done by policies, which are available in advanced operating systems (Linux, Windows XP). These policies have to be switched on and configured by the system administrator with the help of system monitoring tools. The worst case scenario is a server with a low performance due to a inadequate configuration. Therefore the activation of CPU power management is recommended and does not impose a significant barrier.

### 4.3.3 Memory Power Management

Although the increase in power consumption over the last decades has not been as dramatic for memory than for processors, memory power consumption still makes a considerable contribution to the overall power consumption especially of servers that are equipped with large amounts of main memory. Estimation tools for servers like power calculators from SUN, IBM or HP charge 4.5W to 6W per...
DIMM-module installed in the server, with an integrated memory buffer (AMB) in FBDIMM modules in the range of 10 Watt [Mic07a]. Depending on the capacity of one module, this translates to 3W up to 9W per Gb of memory.

Memory chips like Rambus’ RDRAM support low-power states that can be entered when the memory is not accessed and that reduce memory power consumption to 10 % to 1 % of the power consumed in the active state. However, transition from the low-power to the active state requires clock re-synchronization and is therefore time consuming. Policies determining when to enter a low-power state are typically implemented in the memory controller and not exposed to software. Nevertheless, software strategies have been proposed that aim at clustering memory accesses to a subset of memory banks, so the controller detects the remaining banks as being idle and switches them to low power mode. This is achieved by the operating system concentrating memory pages that belong to one application into the same memory bank.

4.3.3.1 Energy Saving Potentials and possible market barriers

Although theoretical memory savings of up to 80 % of memory power have been reported [HPS03] with memory clustering, this approach can only be achieved with novel operating system policies and appropriate hardware support which might come in the medium term. Short term savings in the range of 25 % on the memory level are possible with the deployment of advanced low-power memory technology [Mic07]. The best option to save energy in servers with a high number of memory modules is to reduce the number of modules and minimize the number of chips per module [Mic07b, Mic07c]. Giving the fact that memory power is less than 10 % of system power for commodity servers, savings of 2-3 % are only marginal.

Memory power management is just coming up in the research community and is far away from an introduction in the normal operation of a company.

4.3.4 Disk Drive Power Management

Disk drives used in Desktop and portable systems are often powered down during periods of inactivity to save power. For server systems, this strategy is not applicable for several reasons: Firstly, for servers, there are few scenarios where the disk is idle long enough to justify a shutdown [GZS+03, CPB03]. Secondly, spinning up the disk again after a shutdown introduces penalties of one second and longer which cannot be accepted in a server from which the user demands low response time. Thirdly, and most importantly, server disks are by design optimized for continuous operation, and are therefore only able to abide a certain number of start cycles, considerably fewer than notebook or even desktop disks can.

However, disks contribute considerably to server power consumption. A server disk like the Hitachi Ultrastar™ 15K300 [Hit07] is consuming between 8W and 19W. The fact that server systems are typically equipped with RAIDs of many disks makes disk power management an important issue.

Dynamic rotations per minute (DRPM, [GSKF03a]) has been proposed as a means to conserve power in server disks. DRPM disks offer multiple rotational speed settings, allowing to reduce the rotational speed and thus the power consumption of the disk in times when the frequency of disk accesses is low. Desktop disk drives supporting multiple rotational speeds are available from Hitachi [Hit04]. However, reduced speed is only offered for saving power when the disk is idle, since serving requests requires the disk to resume full speed operation.

4.3.4.1 Energy saving potentials and possible market barriers

Even moderately loaded servers show disk reference patterns that make shutdown strategies for disk drives ineffective. A promising way to save energy is to combine the consolidation of servers with a consolidation of storage. Storage consolidation drastically reduces the number of active disk drives. However
often storage consolidation is achieved with the introduction of a storage area network (e.g., Fiber Channel or iSCSI) or network attached storage. Both technologies require a network interface on the client side that requires about the same energy as a single disk drive (10W) thus compensating the disk energy savings from a consolidation of servers with 1-2 drives. Assuming that disk power is less than 15 % of total system power the expected savings are below 10 %.

Beyond hard disks, solid state drives have recently been introduced for servers, e.g. with IBM Blade Center, delivering 4x higher reliability compared to RAID 1 disks thus being able to save about 18-20W (for two 10000rpm drives). Capacities of 4-32GB – quickly growing – allow extremely fast OS and application access with data stored on the SAN as typical in professional consolidated infrastructures. However, the energy savings are by far not able to compensate the significantly higher costs. For I/O performance sensitive database applications like SAP, which access dozens of disks in parallel using only a very small portion of the available space, the improved I/O performance may justify wider application of SSDs in the near future.

The major drawback of disk drive power management is the long latency to wake up a disk. On the other side it imposes only minor administrative overhead and comes with no risk of data loss. Especially for servers that are not utilized for long periods of time (during the night or the weekend) disk drive power management should be activated.

4.4 Thermal Balancing

Ventilation and cooling of a computer room can contribute 30 % up to 200 % of the power Consumption of the complete IT-infrastructure. By a migration of energy intensive applications to server systems that are close to the outlets of the cooling system (typically at the bottom of a rack), the outlet temperature can be increased without affecting the allowed temperature margins of the server systems. A requirement for this procedure is the knowledge of the system temperatures by the means of sensors and a characterization of applications concerning their heat production [MSS+04, SBP+03]. This approach can also be applied to attenuate the impact of cooling system failures: If the cooling system fails in some part of the computer room, migrating energy intensive applications to locations where the cooling system is still functional prevents thermal emergencies and ensures getting the best performance possible under the given conditions.

4.4.1.1 Energy Saving Potentials and possible barriers

Thermal balancing makes an increase of the inlet temperature to 24°-26° Celsius feasible. A CRAC running at this high temperature halves the cooling costs compared to a CRAC with an inlet temperature of 18° Celsius [MCRS05]. Assuming that cooling cost can make up to 50 % of the total energy consumption, the expected savings are below 25 %.

Thermal balancing requires fine grained measurement of the temperature distribution in a computer room and mechanisms to shift the load from one server to another.

The installation of a temperature measurements infrastructure brings additional costs. The balancing of the load is simple to achieve in a web server cluster with request routing policies. In a general heterogeneous server environment load balancing is a complex administrative procedure that goes fare beyond the know how of a system administrator. Therefore thermal balancing imposes additional cost, administrative overhead and brings the risk of instability or latency because of migration events.

4.5 Consolidation through Virtualization

End user application services are frequently comprised of components that are distributed across multiple servers. The figure below depicts a typical non-virtualized server. A single operating system per server directly initializes and controls all hardware. The result
of this simplistic approach is often server sprawl - a large number of servers that are typically underutilized, are difficult to manage effectively, and increase requirements for data center space, cooling, and power. Studies from IBM and VMWare have found that the typical utilization of Intel/AMD based servers are about 10% on a daily average and high utilizations of about 70% or more are rarely achieved and most often only over a few seconds or less.

To reduce costs – and space, cooling, and power - businesses are eager to apply virtualization to be able to consolidate these applications onto fewer servers, while being careful to maintain isolation between applications and operation environments.

Virtualization is the process of presenting computing resources in ways that users and applications can easily get value out of them, rather than presenting them in a way dictated by their implementation, geographic location, or physical packaging. In other words, it provides a logical rather than physical view of data, computing power, storage capacity, and other resources.

By virtualizing storage and server infrastructure, you can move away from a vertical approach to a more integrated IT environment. There are really two dimensions in Virtualization: Make a big thing (server) look like many small (logical servers, partitioning) to accommodate and isolate many applications and make many things (the logical servers) look like one.

The objective is to bring more flexibility in allocating resources and tools, to integrate the hardware part of the infrastructure and give a horizontal end to end application or transaction view.

The table below summarizes key features of established virtualization techniques.

Nearly all of these solutions involve some form of partitioning and containment. Containment is a concept — the prevention of spreading material or effects beyond a barrier or boundary. Examples include physical server partitioning, such as IBM's dynamical physical partitions (DLPARs) and Sun's Dynamic System Domains, and software-based solutions such as VMware, Microsoft Virtual Server 2005, and
Solaris™ Containers. The underlying concept of all these technologies is that of server virtualization, the partitioning of network servers into several independent execution environments.

The purpose of these virtualization techniques is to enable safe and efficient workload management on a shared resource. **Hardware partitioning** technologies like Sun’s Dynamic System Domains and IBM’s Dynamic Logical Partitions (LPARs) are designed to partition servers along physical boundaries (i.e., CPUs or System Boards) and to work in conjunction with resource management and partition reconfiguration services.

Software-based containment solutions use a concept called virtual machine monitor (VVM). In these solutions, a primary software layer (or even an operating system) runs directly on the system hardware, and permits multiple guest operating systems, such as Linux, Microsoft Windows, or the Solaris OS, along with their applications, to run simultaneously in a contained manner on the host system.

However, for all these technologies operating systems and applications are fully contained within their respective partitions, and effectively run on separate hardware servers, albeit within the same physical system enclosure. This means that all virtual machines on a physical server use all the resources (CPU, memory, disk) necessary to run a separate operating environment. In fact, especially VVMs can consume significant CPU resources as they re-write or redirect guest operating system code, especially when they need to intercept and redirect privileged guest operating system instructions.

### 4.5.1 Principles of System Virtualization

Originally invented in the late 1970’s for the mainframe world, virtualization technology has nowadays become available also for commodity server systems, and is being eagerly embraced by research and industry as to overcome all sorts of deficiencies in legacy operating systems (OSes). Virtualization technology allows multiple OS instances to be consolidated onto a single physical machine. A virtualization layer thereby provides, for each running OS instance, an isolated virtual computer environment, which closely resembles its physical counterpart. Any OS and application designed for real hardware can therefore execute in a virtualized system as well. The virtualization layer for typical Intel/AMD servers consists of a virtual machine monitor catering for the processor and device virtualization, and a host OS controlling the actual physical hardware; the host OS itself may be a fully-fledged OS such as Linux or Windows (as in VMware Workstation), or a minimal, hypervisor-based OS (as in Xen or VMware’s ESX server).

The most well-known application of virtualization technology is the consolidation of multiple OSes onto a single server, in order to improve utilization. Further applications are dynamic resource provisioning via software partitioning or live migration. Finally, virtualization can be used to enhance a system’s reliability and security, since it allows the privileged and trusted part of the operating system to become much smaller compared to a full monolithic legacy operating system kernel.

#### 4.5.1.1 Hardware Partitioning

Hardware partitioning is making a big server look like many small logical ones: creating several instances of operating systems in one server box, it is the most widely known dimension of virtualization. IBM has pioneered in that matter with the mainframe (S390 to System z) and VM, which for decades brings that functionality. One can create on demand lots of logical servers without dealing with the needed resources, automatically allocated by the system on the fly. Hardware partitioning is now a standard feature of nearly all UNIX servers in the midrange and high end server space.

Dynamic hardware partitioning allows server resources to be grouped into logically separate systems within the same physical server. The word dynamic indicates that the configuration of the logical servers can be modified in operation without interrupting the running applica-
tions. VMware allows to create logical partitions but no to perform dynamic reconfiguration.

In partitioned environments where more business-critical applications are consolidated on different operating systems with the same hardware, additional availability and service-ability is needed to ensure a smooth recovery of single failures and allow most of the applications to still be operative when one of the operating systems is out of service. Furthermore, high availability functions at the operating system and application levels are required to allow for quick recovery of service for the end users.
Table 4.4 Key features of established virtualization technics

<table>
<thead>
<tr>
<th>Vendor Family</th>
<th>IBM</th>
<th>Sun</th>
</tr>
</thead>
<tbody>
<tr>
<td>Solution</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Implementation</td>
<td>LPARs</td>
<td>LPARs</td>
</tr>
<tr>
<td>Operating system</td>
<td>Solaris</td>
<td>Domains</td>
</tr>
<tr>
<td>LPARs</td>
<td>SW</td>
<td>SW</td>
</tr>
<tr>
<td>LPARs</td>
<td>Hypervisor</td>
<td>HW</td>
</tr>
<tr>
<td>LPARs</td>
<td>Hypervisor</td>
<td>Hypervisor</td>
</tr>
<tr>
<td>LPARs</td>
<td>SW</td>
<td>SW</td>
</tr>
<tr>
<td>LPARs</td>
<td>SW</td>
<td>SW</td>
</tr>
<tr>
<td>LPARs</td>
<td>SW</td>
<td>SW</td>
</tr>
<tr>
<td>SW</td>
<td>SW</td>
<td>SW</td>
</tr>
<tr>
<td>SW</td>
<td>SW</td>
<td>SW</td>
</tr>
<tr>
<td>Max. number partitions</td>
<td>64</td>
<td>18</td>
</tr>
<tr>
<td>Granularity (min. number of processors per partition)</td>
<td>254</td>
<td>24</td>
</tr>
<tr>
<td>Granularity (min. number of processors per partition)</td>
<td>254</td>
<td>32</td>
</tr>
<tr>
<td>Granularity (min. number of processors per partition)</td>
<td>30</td>
<td>8192</td>
</tr>
<tr>
<td>Granularity (min. number of processors per partition)</td>
<td>&gt;10000</td>
<td>1</td>
</tr>
<tr>
<td>Granularity (min. number of processors per partition)</td>
<td>1/8</td>
<td>1/64</td>
</tr>
<tr>
<td>Granularity (min. number of processors per partition)</td>
<td>1/10</td>
<td>1 or no constraint</td>
</tr>
<tr>
<td>Dynamic reconfiguration</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>Shared resources</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>Interpartition communication</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>CPUs, memory, I/O</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>CPUs, memory, I/O</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>CPUs, memory, I/O</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>CPUs, memory, I/O</td>
<td>yes</td>
<td>yes</td>
</tr>
</tbody>
</table>
In the IBM System p5 systems, there are several mechanisms that increase overall system and application availability by combining hardware, system design, and clustering.

Virtual I/O separates partitions from I/O. It enables aggregated Ethernet and Storage resources to be shared between many partitions, rather than requiring fixed allocation of a limited physical I/O resource. Adapters are shared among the partitions for optimized usage.

Virtual LANs enable high speed communications between partitions. It allows to re-create the network viewed by applications and to take advantage of memory like speed.

Sun SPARC Enterprise servers feature redundant, hot-swap power supply and fan units, as well as the option to configure multiple CPUs, memory DIMMs, and I/O cards. Administrators can create redundant internal storage by combining Sun SPARC Enterprise server hot-swap disk drives with disk mirroring software. High-end servers also include redundant, hot-swap service processors, and Sun SPARC Enterprise M9000 servers include degradable Crossbar Units and redundant Clock Control Units. If a fault occurs, these duplicated components can enable continued operation. Depending upon the component and type of error, the system may continue to operate in a degraded mode or may reboot — with the failure automatically diagnosed and the relevant component automatically configured out of the system. In addition, hot-swap hardware within the Sun SPARC Enterprise servers speeds service and allows for simplified replacement or addition of components, without a need to stop the system.
4.5.1.2 Full-/Para-Virtualization

The basic principle of OS virtualization is to permit the guest OS to execute innocuous operations directly on bare hardware, but to emulate all sensitive operations that would affect and potentially compromise other, concurrently executing guests OSes. Examples of such sensitive operations are the changing of the processor’s privilege level, or the issuing of a DMA request to a disk device. The virtual machine monitor (VMM) must intercept the guest OS code whenever it tries to issue such a sensitive operation, and then transfer control to an appropriate emulation routine. For that purpose, VMMs typically run all guest OS code at a low privileged processor mode (user mode), with the result that the processor will raise an exception as soon as the guest tries to execute a sensitive instruction.

However, relying solely on processor exceptions to detect of sensitive instructions turns out to be a problematic approach on x86 processors prevalent in modern computer systems. The root cause lies in a few instructions not generating a processor exception when executed in low-privileged mode; rather, the instructions silently fail or pass with changed behavior. To successfully virtualize a physical processor, however, the VMM must reliably detect and emulate all cases where the guest executes sensitive code. Several methods have been proposed to perform that task reliably, even in presence of not fully virtualizable processors. Software methods, as they are used in VMware and VirtualPC (now being part of Microsoft) products [VM06], leverage binary translation or binary rewriting to detect all sensitive guest kernel code, and insert the emulation routines directly into the instruction stream. Hardware methods, like in Intel’s VT or AMD’s SVM processor extensions [Nei06], change the processor’s functionality to now reliably trap into the hypervisor on all sensitive instructions.

4.5.2 Principles of Container Virtualization

If a collection of processes and resources can be defined and bounded to match the re-
requirements of a contained server environment on a single instance of an operating system, this type of containment has been described as container virtualization [IDEAS-VIRT]. Container virtualization can be accomplished efficiently without the use of a separate virtual machines (or partitions).

### 4.5.2.1 Solaris Container

One of the first approaches to container virtualization was taken with Solaris™ Containers [SOL-CONT]. Solaris Containers create an execution environment within a single instance of the Solaris OS and provide full resource containment and control, software fault isolation and security isolation.

The figure below illustrates the concept of Solaris 10 Containers. The Solaris Operating System runs directly on the hardware, manages the boot process, and initializes interfaces to the CPUs, memory, host bus adapters, network interface cards (NICs), storage, and device drivers in the system, and it is referred to as the global zone. The administrator defines one or more non-global zones that contain virtual server environments. A non-global zone appears to all users — end users, applications, developers, and the zone administrator — as a fully realized server with its own host name, IP address, process and name space, root and user names and passwords, and network devices and file systems [OPENSOL-ZONES].

Only one instance of the Solaris OS runs on the hardware, allowing all local Zones or Containers to share central resources of the operating environment: CPUs, memory, disks, devices. Zones even share the central tasks and processes of an operating system which leads to much less memory and disk usage inside a virtualized server. Zones (Solaris Containers) do not require dedicated CPU resources, dedicated I/O devices such as HBAs and NICs, or dedicated physical memory.

![Fig. 4.8 Concept of solaris 10 containers](image)

### 4.5.2.2 Further Containment Techniques

Beside of Solaris Containers there are some other implementations of OS virtualization in the market. IBM AIX 6 on system I and system p Power6 systems allows application virtualization environments within one OS instance. SWsoft Virtuozzo is a commercial product for Linux and Windows. SWsoft’s Virtuozzo modifies (Linux or Windows) OS kernel to support multiple virtual private servers, in an approach comparable to Solaris Containers. Virtuozzo for Linux can be installed on Red Hat, Fedora (Red Hat’s development platform), or CentOS.
Technical measures to improve the energy efficiency of servers and data centres

to support most Linux flavors as guest OS's. Virtuozzo for Windows offers similar capability to run legacy Windows OS's. IBM's acquisition of Meiosys includes a product similar to Solaris Containers called MetaCluster UC. IBM is productizing this capability into AIX, and could also offer it on Linux. Other commercial products exist in this area, including Novell's AppArmor, Microsoft's Softricity, and Trigence AE. However, these are more focused on security than consolidation. Further the community of open source developers has contributed several containment and virtualization solutions, including FreeBSD Jails, User Mode Linux, and OpenVZ (based on Virtuozzo), and Linux V-Server.

4.5.2.3 Energy Saving Potentials and possible market barriers

Energy savings of consolidation/virtualization go beyond the pure savings from a reduced number of servers which results from a reduced static power because of the fewer power supplies, fans, chip sets or I/O interfaces:

- The communication between two systems in a virtualized environment on one physical server is more efficient than the communication via network between two physical servers.
- The utilization of memory is improved. With a rising contribution of main memory to the energy consumption in the next years, a higher utilization implies increased energy efficiency.
- Virtual Machine concepts allow a higher degree of resource utilization by supporting sub-CPU granularity of the virtualized servers, leading to more energy efficient usage of server hardware.
- Container Virtualization does allow the same increased resource utilization and even further reduce the amount of necessary physical devices (CPUs, memory, disks, HBAs) by sharing the same operating environment. This leads to even less energy usage within a virtualized server.

- Hypervisor-based policies and scheduling plans can be used to keep up energy constraints at the software layer, to complement hardware cooling facilities and avoid their over-provisioning. First scientific publication [NuSc07, SLB07] show benefits from dynamic load balancing in combination with dynamic CPU power management in virtualized environments. But theses technologies are not yet adopted by HW/SW vendors.

Furthermore enhanced resource provisioning capabilities of virtualized systems may be used to conserve power more efficiently: live migration can be used to dynamically consolidate virtual machines during phases of underutilization, and to re-allot them during phases of high load. Idle machines can then be put into low-power sleep states.

The factor that determines the saving potential for consolidation is the number of servers that can be eliminated. The EPA report to the congress [EPA 07] expects a consolidation factor of 2 to 1 for most of the best practice cases. This results in 50% of total energy savings. With aggressive volume server virtualization a factor of 5 to 1 can be achieved. However a small percentage (5% in the EPA report) of servers cannot be eliminated because of legacy applications.

Hardware partitioning is a technology that is in production for decades and does not impose major administrative risks. The major barrier for the introduction is the fact, that hardware partitioning is only available for the mid and high end server market and perceived as more expensive although TCO considerations including energy costs can proof differently.

Full-/Para-virtualization require additional software for the virtual machine environment plus additional support to adapt the operating system in a para-virtualized environment. Additional software requires license fees (e.g., 5750$ per 2 CPUs for VMware enterprise edition) and requires support by the OS-vendors for software updates does. Furthermore the virtualization of I/O components
implies additional CPU (w/o HW support (Intel/AMD server) overhead that limits the energy savings due to consolidation. Finally the sharing of many hardware resources requires administrative know-how. Therefore full-/para-virtualization exhibit financial and administrative barriers on low end servers.

Application and container virtualization is technical approved for several years and requires only minor administrative efforts. The technology is integrated or is part of the operating system and requires no additional software or license fees. However it does not allow to run different guest operating systems on the same hardware. Therefore application and container virtualization has fewer barriers beside its limited flexibility compared to the other virtualization concepts.

4.6 Server Cooling

4.6.1 General aspects and saving potentials

Cooling respectively air conditioning of server facilities causes a major part of overall energy consumption in data centers. Several studies on energy consumption of hardware and cooling equipment have been conducted in the past investigating the subject from somewhat different perspectives. US and Swiss studies [MJ01] and [Al01] reported that the energy consumption of the cooling equipment on average is between 60 and 75 % of the energy consumption of the IT-hardware. Studies in several data centers in the US indicated an energy consumption of the cooling equipment of about 35 % of total energy consumption in the data centre [LBNL02]. However this study also showed that this share in total consumption varies between 20 and over 50 % in different data centers. Shamshoian et al. [Sh05] report a variation factor of ten in power requirements per unit floor area and a 2.5:1 variation in the effectiveness of the HVAC (heating, ventilation and air-conditioning).

Energy demand for cooling and other server infrastructure (UPS, power distribution, light etc.) on average is in the same order of magnitude as the energy demand of the hardware [Ko07] thus a typical power usage effectiveness (energy demand of data center to energy demand of IT) is 2. But also regarding power usage effectiveness at data center level there seems to be strong variability. For example Tschudi at el [Ts06] report from case studies a variability of typical power usage effectiveness by a factor of five.

Forecasting the future trends it is expected that electricity demand for cooling will strongly increase as heat load in data centers will be steadily rising. Current heat loads per square meter are already at an order of magnitude of 1500 W/m2 varying roughly between 500 and 2500 W/m2 [Sc07]. For the future even higher average heat loads in the order of magnitude of 3000 W/m2 and above are expected. On a rack basis average power has increased from 10 kW some years ago to 20 and even 30 kW nowadays [Pa07]. Estimations from Gartner [Ga06] indicate that the effort and costs for cooling increases tremendously with increasing power density in racks. For example cooling power increased from 60 to 200 % when power density per rack is increased from 6 to 12 kW.

There are several approaches to optimize server cooling in terms of energy efficiency. According to different studies it should be possible to reduce energy consumption for cooling well below 20 % of total energy consumption [Al03]. Under specific conditions including free cooling and optimized server facility structure even a further optimization of energy efficiency would be possible lowering the share of cooling even below 15 %.

The most prominent concepts to improve cooling efficiency to be discussed in the following section are:

- Optimizing the structure of the server and the cooling system thereby optimizing air flow and overall efficiency
- Optimizing temperature of cooling air and cooling liquids
- Use of free cooling
- Rack based cooling
The following chapters provide an overview on concepts for energy efficient air conditioning at different levels of the system. However server internal cooling by integrated fans or liquid cooling units is not addressed in this section.

### 4.6.2 Concepts to improve energy efficiency of server cooling

#### 4.6.2.1 Small server rooms or so called server closets

Small server rooms or so called server closets are often cooled with small standard split system air conditioners. If no measures are undertaken to optimize the airflow this is a rather inefficient way of cooling since warm and cold air flows are mixing in the room and homogeneous cooling of the server hardware can not be achieved. The location of the servers in these rather small rooms is often very much dependant on room space and architecture, and therefore the options for optimization are limited.

If the air conditioning units are placed in an inefficient way and no measures are undertaken to optimize the airflow the air conditioning equipment has to be operated at very low temperatures to achieve homogeneous cooling of the equipment. This is of course inefficient.

It is advised that for small server rooms split systems are installed in a way that air mixing is minimized. External components (compressor unit) of the split system should be moved to a favorable position – e.g. shaded location offering the possibility for significant airflow around the unit. Short circuiting and mixing of the cold and warm airflow should be avoided as far as possible for example by ensuring that the baffle plate of a split cooling system is properly fitted for example.

#### 4.6.2.2 Efficiency of cooling/air-conditioning units

Besides optimization of the air flow also the efficiency of the cooling units is of relevance fan efficiencies vary between 55 and 65 %. Current efficient chiller technology uses direct driven fans being 30 % more efficient than standard fans. Further approaches to increase energy efficiency in small server rooms are similar to the ones that can be applied for larger facilities also, including higher temperature settings and use of free cooling. These concepts are discussed in the following section on larger server rooms and data centers.

#### 4.6.2.3 Larger server rooms and data centers

##### 4.6.2.3.1 Design of cooling system and air flow

a) Simple air conditioning without specific measures to optimize air flow and temperatures

Simple cooling systems in larger server rooms and data centers are not specifically designed with an optimised air flow. Also the location of the server hardware in the rooms mostly is not optimized in these cases. Often the equipment is placed in a row and cold air is introduced from the ceiling or from the side. Such systems normally do not allow an efficient airflow since cold (incoming) and warm (outgoing) air is strongly mixing. The warm air is not dissipated most efficiently. Already heated air maybe sucked into adjacent hardware. Fig. 3 shows the mixing of cold and hot air if the server room is laid out as shown, where the cold is vented through ducts between the racks. These “designs” are often found in server facilities with an installed power of less than 2-3kW per rack.

Measures to improve airflow and efficiency are individual but also limited in such cases since very efficient solutions can only be achieved by largely changing the location of the hardware as well as redesigning the cooling system. A possible option to improve the cooling efficiency in such cases without completely changing the system is the use of doors equipped with fans on the backside of the servers. Such doors support the dissipation of the hot air. Additional fans first of all mean additional energy consumption and overall savings can only be achieved if for example a
Energy efficient servers in Europe

higher cooling temperature can be used due to this measure.

**Fig. 4.9 Server room with simple air conditioning placing of servers in a row**

b) Advanced cooling designs

Larger server facilities respectively with a higher power density per rack are often equipped with a cooling system in a raised floor and servers are placed in a back to back concept with so called cold and hot aisles or corridors (Fig. 4). This type of server cooling in principal allows an efficient air flow. The cold air accesses the cold aisle or corridor via the raised floor is sucked through the server and flows back to the chiller via the hot aisle and a hot air return plenum. For high performance systems and highly powered racks further measures should be undertaken to optimize the airflow. For example cold aisles can be closed on top to avoid mixing of cold and warm air, Outlets for cold air should be carefully designed and located to ensure proper cooling of hardware in the whole rack.

The energy consumption is also highly dependent on air velocity. Optimized systems run at low velocities which lead to better distribution of the cold air and thus reduce the risk of hot spots. High velocities cause a substantial increase in electricity demand.

Airflow dynamics should allow air to cover the floor space in a homogeneous way to access all heat sources. To achieve this, dynamic pressure in the raised floor and air velocity must be kept low (< 3 m/s). Air outlets have to be well defined.

To optimize air flow and energy efficiency in server facilities it is also important to minimize bypass air and recirculation air flow. Bypass air is returning directly to Computer Room Air Conditioners (CRACs) without cooling the server. Recirculation air is flowing directly from the server outlet to the server inlet.

If the cold and warm airflows mix due to bypass, recirculation or other effects, then the temperature of the cold air from the cooling system has to be reduced. Power consumption for cooling increases by 1–3 % for every degree the temperature of the cold air has to be reduced due to inefficiencies. Frequently, the temperature of the cold air delivered to the racks does not need to be lower than 10–15°C. Using typical cold air temperature levels significantly below 10 °C therefore means that the power consumption is 10–50 % higher than it needs to be [Vi04].

Air flow is often also limited by inefficient wiring on the back side of the servers limiting the dissipation of the hot air from the server.

Circulating air through racks is a very inefficient cooling method because with this approach cooling the upper servers in the rack sufficiently requires massive cooling at the
bottom. A large part of the cool air dissipates or is heated up by the lower servers before it reaches the upper ones. Servers are often mounted in the rack in a way that the cold air is not easily drawn into the servers higher up the rack. Consequently the cooling system has to operate at a very low temperature, which also leads to high power consumption.

In general it has been also indicated above an important aspect to improve efficiency is separating cold and hot aisles in a way that air mixing is prevented. In best practice cases reported by Garday and Costello [GC06] effective separation of aisles was achieved by a lightweight wall between each hot and cold aisle. These walls were extending from the top of each row of cabinets to the ceiling to block airflow across the top of the cabinets, and from floor to ceiling at each end of the aisles. Hot return air was forced to flow through grills in the hot aisles into a return air plenum above the drop ceiling. In addition instead of standard perforated air diffusers special diffusers were used increasing the air flow and decreasing velocity. Consequently air temperature and chilled water temperature could be reduced to 20 °C and 14.44 °C respectively. The hot aisle enclosures helped to achieve a very high level of air conditioning airflow efficiency, defined as the amount of heat removed per standard cubic foot of airflow per minute. Also good effectiveness of air conditioning was achieved described by the ratio of air conditioning power to computer power (0.23).

![Fig. 4.10 Server room with raised floor and hot aisle cold aisle concept](image)

c)

Cooling systems for high power systems without raised floor

Garday and Costello [GC06] also tested cooling designs for larger systems avoiding the raised floor concept. In this case studies air was supplied above the aisles putting cooling equipment on a kind of mezzanine level. This is only possible if a room height of more than 7 m (25 feet) at minimum is available. However the effectiveness of the concept has not been proven yet.

Tschudi et al [Ts06] among others summarize the following typical inefficiencies regarding air flow in data centers based on best practice cases:

- short circuiting and mixing of cold and hot air
- misplaced floor air discharge
- poorly located air conditioning units
Energy efficient servers in Europe

- undersized ceiling height or hot air return plenum
- air blockages due to piping and cabling
- openings in racks allowing short circuiting
- poor air flow due to rack structure
- inappropriate placing of IT equipment
- inappropriate under floor pressurization

Garday and Costello [GC06] publish experiences from efficient design approaches for several types of data centers including blades.

4.6.2.3.2 Efficiency of cooling units for larger systems

Tschudi et al [Ts06] propose to use central chiller and air handling systems preferably to local units due to the following reasons: fewer large motors and fans are more efficient than a high number of small ones, central systems are designed for variable volume operation using variable frequency drives etc. Local air handling units are difficult to coordinate/control in an optimized way, the centralized water cooling unit is more efficient.

4.6.2.3.3 Free cooling

The strategy of free cooling offers basically two options that do not require refrigeration: cooling with ambient air and cooling with water at ambient temperature. The potential of airside free cooling is seen as higher by some authors since air is available in unlimited amounts and higher temperatures [To07]. Other authors however indicate that free air cooling should be applied with some caution. It is not recommended to use standard economizers without analyzing local climatic conditions (temperature and humidity fluctuation) and possible pollutants in the air. However the use of liquid based heat exchangers (water/air) is generally recommended for climates where the wet bulb temperature is below 13 °C for about 3000 hours/year.

Free liquid based cooling involves a heat exchange system. The compressor is cooled via a liquid circuit with an externally located heat exchanger. In addition, an extra heat exchanger is connected to an evaporator, which is also connected to the liquid circuit. The unit can operate as a standard cooling system or with free cooling under low ambient air temperature conditions. Combined cooling with both free cooling and compressor cooling is possible.

Free cooling with water can be done in a sequential or mixed way thus either switching from free cooling to refrigeration at defined conditions or to use a mix of free cooling and refrigeration. Energy saving potentials due to free cooling can be between 50 and 75 % [Al03], [Ts06]).

Minimum savings achieved by free cooling have been demonstrated in practice for example by [Sc07] being in the order of magnitude of one third of the cooling electricity consumption at least. In this example a 1000 kW cooling equipment was compared with and without free liquid cooling leading to energy savings of 33 % or 460000 kWh/a and 46000.- EUR respectively.

4.6.2.3.4 Temperature and humidity management

Depending on the architecture and air flow temperature levels in server facilities can be kept higher than the typical settings often found in practice. Temperatures of 26 - 30 °C or even above are possible. Optimized layouts with raised floor and hot aisles/cool aisles concept (section 3.2.2.1c) allow the temperature of the cold air delivered to the IT equipment to be high, which increases the total cooling efficiency. [Sc07] indicates that servers are designed for environment temperatures of 30 - 35 °C. However the higher the average level and the temperature of the in-flowing air the more optimized the air circulation has to be to avoid unacceptable hot spots.

At low room air temperatures the electricity demand for cooling and ventilation can differ by a factor of 2 between an optimized and an inefficient layout. At 25 – 28 °C differences between efficient and inefficient equipment are still in the order of magnitude of 40 - 50 % [Ae04].
System modeling done by [Al03] showed that a system without free cooling, a cold water temperature of 6/12°C, a computer room temperature of 20°C and an inefficient ventilation system demands about 64 % to the electricity consumption of the IT hardware. By use of free cooling an increase in the cold water circuit temperature and an optimized ventilation system as well as a rise in computer room air temperature from 20°C to 26°C demand for cooling and ventilation could be strongly decreased to 17.5 % of IT hardware consumption.

Regarding air humidity ASHRAE [As04] proposes a wider range for humidity settings than the widely applied standard targets. In the meantime earlier recommendations to operate systems at very strict humidity levels are put into perspective. Operating the cooling units above the dew point (chilled water above 10 °C) eliminates unnecessary dehumidification.

4.6.2.3.5 Correct sizing of infrastructure and modular cooling

Over-sized IT systems and consequently oversized cooling equipment is a typical phenomenon leading to unnecessary electricity demands and undesired sources of heat within a facility. Oversizing is often due to overestimates of current and future loads and necessary computing capacity. Often an undesired redundant over-sizing by multiple parties can be observed (owners, process engineers, electrical engineers, HVAC engineers, etc.) each adding “safety factors” to their sizing calculations. This results in much higher first cost but also higher operating costs due to excessive and inefficient on-off cycling and/or part-load operation.

Traditional rules of thumb regarding the sizing of cooling facilities can also lead to considerable inefficiencies. For example it has been generally accepted that datacenters can require an electrical grid connection to support 2,7 kW/m² (or more) watts per square foot of space. Benchmarking by Shamshoian et al.(2005) in contrast to this has shown that average values in practice are around 0,27, maximum values being at a level of roughly 0,7 kW/m².

Especially for larger data centers cooling infrastructure is often designed for maximum system capacity which is close to the maximum room capacity. The expected load is about 90 % of total capacity. However this load often is not reached even after years. Initial loads are often below 20 % and even after years only 30 % of maximum load is reached [Re07]. Thus the cooling equipment is clearly oversized.

Oversizing is a frequent problem to be addressed. One option to avoid this effect is modular cooling equipment that can be upgraded and extended according to specific needs. This for example can be achieved by using blade systems in cooled racks.

4.6.2.3.6 Liquid rack cooling

Liquid rack cooling involves cooling directly on the rack with only little heat dissipation into the server room. The principle is based on closed server racks and on air to liquid heat exchanger. There are different types of approaches for rack based cooling:

- Completely integrated rack cooling where the fans and the heat exchanger are part of the rack
- External heat exchanger coupled to the rack (also a closed rack concept)
- Heat exchanger integrated in backdoor of open racks.

The last of the three options is sometimes used when existing data centers are upgraded. It is an easy option to upgrade the cooling facility with additional cooling capacity.

There is a certain trend observed to rack based water cooling. According to IBM water cooling in the "Cool Blue systems" allows energy savings of 15 %. According to Pawlowski (Pa07) rack- and water-based cooling can reduce energy consumption in a data center by up to 25 %.
4.6.3 General economic aspects

Energy consumption in data center facilities is 10 - 30 times higher than in standard office facilities already indicating the ecological and economic potential in this area. As field investigations clearly show measures to increase the energy efficiency implemented at the time of the design of a data center always clearly pay-off within a short period of time significantly reducing cost of ownership [Sh05]. However according to the same authors also retrofitting of data centers offers significant energy savings with a pay back time of less than three years. It is evident that cooling/air conditioning can be optimized considerably in many data center facilities allowing considerable energy savings. However looking into the future the central problems are rather in the area of the roots of increased cooling demand which is the enormous increase of energy consumption by server equipment.

A study by Brill [Br07] analyzed how electricity costs and TCO will be increasing during the next years. The example shows that an investment of 1M$ for server hardware in 2006 causes 0,5M$ of electricity costs or 0,9M$ of TCO in a three years period. For the future scenario it is expected that the same 1M$ investment causes 1,9 M$ for electricity and 7M$ for TCO in 2012 indicating that energy costs and TCO soon will become the much more relevant cost factors than hardware.

Other recent studies suggest a more moderate growth of energy consumption for example of about 100 % from 2000 to 2005 (Koomey et al 2007). It is difficult to estimate when exactly electricity costs will exceed hardware costs under given trends but this is expected to be the case some time between 2009 and 2012. The cause behind this is the fact that computing power currently is developing considerably faster than energy efficiency. Although hardware is becoming more and more energy efficient these improvements in energy efficiency are over-compensated by the increase of computing power and computing density. The only factor that is currently seen which could significantly retard the undesired trend is virtualization.

There is currently a significant lack in the consideration of costs in data centers. Pure hardware costs are still dominating purchasing decisions, TCO is often not considered. Responsible decision makers are still not aware that the relevance of TCO has strongly changed from a minor factor in the past to highly relevant one. The same situation is observed for electricity costs per se.

Despite the general economic potentials many specific barriers are currently limiting the implementation of energy efficient systems and infrastructure in practice. Among others these are

- the separation between capital and operating budgets
- differences in incentives for owners and tenants
- lack of trained staff
- fragmentation among the stakeholders interacting to create and maintain facilities
- potential adverse impacts on uptime
- undesirable prolonging of facility construction or renovation time
- discounting incentives

4.7 Summary of energy saving options

Table 4.5 shows a summary of energy saving potentials in the areas of hardware, power management and consolidation by virtualisation. The energy saving potentials range between a few percent in the area of single components up to 80% if total server virtualisation is considered.
### Table 4.5 Summary of energy saving potentials

<table>
<thead>
<tr>
<th>Measures to Increase Energy Efficiency</th>
<th>Saving Potentials (system power)</th>
<th>Barriers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Efficient Power Supply</td>
<td></td>
<td>- Initial cost</td>
</tr>
<tr>
<td>Efficient Cooling</td>
<td>50 %</td>
<td>- Strongly dependant on already existing infrastructure, long life time of equipment</td>
</tr>
<tr>
<td>Power Management</td>
<td></td>
<td>- Complex administration - Limited applicability</td>
</tr>
<tr>
<td>Shutdown Strategies &lt; 50 %</td>
<td></td>
<td>- Not yet approved</td>
</tr>
<tr>
<td>CPU &lt; 20 %</td>
<td></td>
<td>- Infrastructure for temperature and load monitoring - Complex administration - License for migration software</td>
</tr>
<tr>
<td>Memory &lt; 3 %</td>
<td></td>
<td>- Not available for commodity HW</td>
</tr>
<tr>
<td>Disk Drive &lt; 10 %</td>
<td></td>
<td>- License for virtualization software - Overhead for virtualization - Complex administration</td>
</tr>
<tr>
<td>Thermal Balancing &lt; 25 %</td>
<td></td>
<td>- Limited choice in operating system</td>
</tr>
<tr>
<td>Consolidation</td>
<td></td>
<td>- License for virtualization software - Overhead for virtualization - Complex administration</td>
</tr>
<tr>
<td>Hardware Partitioning &lt; 80 %</td>
<td></td>
<td>- Limited choice in operating system</td>
</tr>
</tbody>
</table>
5 Energy efficiency criteria and benchmarks

5.1 Introduction

Benchmarks are a common approach to assess the characteristics of a system. For the assessment of the energy efficiency and the improvement in energy efficiency by the introduction of specific measures, a bunch of criteria has to be satisfied. We classify benchmark environments to assess the energy efficiency, judge their applicability for cases with efficiency improvements and describe a measurement protocol to assess the energy efficiency in general. An outcome of the analysis is that common benchmark environments are suitable for the assessment of single machines with single applications but they are not applicable in advanced consolidation projects which represent the state of the art in energy-aware provisioning and administration.

In practical situations IT managers or other responsible staff are more interested in overall energy saving potentials under real operating conditions respectively workloads rather than energy efficiency of single hardware units under standardised (lab) conditions.

Therefore we close with a recommendation for a procedure to assess the gains in energy efficiency for best practise cases. This approach may be useful when energy saving potentials in server systems shall be demonstrated.

In the future standardised benchmarks currently developed by SPEC (SPEC Power and SPEC Virtualization) will offer additional tools for assessing and comparing hardware. However SPEC data normally is published only for selected models and not for the server market in general.

The energy efficiency of a server is not only dependent on the characteristics of the hardware but also on the usage profile of the energy consuming functional blocks of the hardware that are active at runtime. This profile highly depends on the workload and the operating system that provides services to the running applications. Therefore we have to characterize the workload under aspects that impact the energy efficiency as well as the ability to assess the energy efficiency of a server with a specific HW/OS configuration. Further we have to describe procedures to effectively measure the energy consumption of a server with reasonable effort.

5.2 Workloads

The analysis of the energy efficiency can be performed with the workload of the normal operation of a server or with a standardized benchmark. For comparison of a wide range of assessments a benchmark is favourable. However many real workload scenarios cannot easily be mapped to a standardized benchmark suite.
5.3 Classification of benchmarks

Standardized benchmarks can be characterized according to their timing behaviour, the heterogeneity of the benchmark suite, requirements to hardware and system software as well according to the patterns of the usage of hardware and system software.

5.3.1 Timing

*Duration based benchmarks* measure the time to complete a certain amount of work. The amount of work is predefined before the run (e.g., Linpack, SPEC CPU2006).

*Throughput based benchmarks* measure the throughput for a given time (e.g., SPEC jbb2005). Here the workload is tuned at runtime by injecting requests that have to be answered according to quality of service demands. The workload driver injecting the requests can be part of the system under test (local driver) or located on another computer (remote driver). Power measurements have to be correlated with a performance value.

Duration based benchmarks always run the system at 100% utilization. Therefore the energy efficiency cannot be accessed at predetermined points at a certain throughput or utilization. Therefore the duration based benchmarks are only applicable for scenarios of extremely high utilization like in the case of grid or scientific computing. For scenarios with varying workload they are not applicable.

Throughput based benchmarks are best suited for the assessment of the energy efficiency because they provide the necessary information over the full range of utilization from 0% to maximum throughput. A remote driver is favourable for power/performance measurements because the driver software for the allotted injection is independent of the system under test (SUT) and the acquisition of power data can be co-located on the driver system.

5.3.2 Heterogeneity

Homogenous benchmarks consist of a single class of application that runs the complete time of measurement (e.g., SPECjbb2005) while a heterogeneous benchmark consists of several applications that are evaluated independently and where the individual benchmark results are summarized according to a given metric (e.g., averaging). The individual applications of a heterogeneous benchmark can run sequential (e.g., SPEC CPU2005) or run in parallel (e.g., SPEC Virtualization).

Homogenous benchmarks exhibit a regular execution behaviour that allows consistent measurements at arbitrary points in time while the assessment of the energy efficiency with a sequential heterogeneous benchmark requires the correlation of measurements with the runtime points of the individual application. The situation becomes more complex in a parallel environment with time multiplexing of the applications like it is in a virtualized environment (e.g., SPEC virtualization). Here the only viable solution is to run the benchmark for a longer time (e.g., hours) and to sample and average the energy consumption at many arbitrary measurement points.

Today the most promising solution to increase the energy efficiency of servers is the consolidation, where several servers running applications at very low utilization are virtualized and consolidated on a single system. To assess the efficiency of a virtualized environment we need a collection of representative benchmarks with multiple drivers each fulfilling the requirements for power/performance analysis. The development of VM performance benchmarks has just begun [SV08]. Power/performance analysis for VM environments is still an open topic.

Summarizing up we have to state that homogenous benchmarks are an approved workload to assess the efficiency of a server while parallel heterogeneous benchmark suites ideally resemble the realistic runtime conditions of a consolidated environment. However suitable heterogeneous benchmarks to be applied for consolidated, virtualized systems
are not yet approved and still under development. They will not become available before 2008 or 2009.

5.3.3 Use of HW/Components

Benchmarks can use a subset of the available hardware components depending on their internal structure and their interface to the operating system.

- Benchmarks that don’t issue I/O operations at all or just submit a small number of I/O requests are called CPU bound. Because the CPUs don’t have to wait for I/O devices with high access latency, the throughput and energy efficiency of systems can easily be compared. The results can be reproduced without the need to establish exactly the same I/O configuration as in the reference system. An example for this class of benchmarks is SPECpower_ssj2008 [SP08].

- Benchmarks can run in parallel on an arbitrary number of processors/cores. This scalability is required to cover a broad range of servers and to configure different levels of CPU utilization. In SPECpower_ssj2008 this scalability is achieved by instantiating a specified number of threads and Java virtual machines (see [SP08a]).

- Benchmarks may make substantial use of I/O devices (e.g., disk drives, network interfaces). A typical representative of this class is the system file server benchmark SPEC SFS [SF97].

5.3.4 System Requirements

- The benchmark runs in a standardized runtime environment (e.g., SUN Java Runtime Environment (JRE) or Microsoft Common Language Runtime (CLR))

- The benchmark runs as native code and is OS and architecture dependent.

5.3.5 Hardware Requirements

Benchmarks may require specific types of hardware components or a minimum amount of hardware resources such as

- CPU Instruction set architecture
- Minimum/maximum number of CPUs
- Minimal main memory capacity
- Number of network interfaces
- Number of disks
- Size of file system space

5.4 Examples of relevant benchmarks for assessing energy efficiency

5.4.1 SPEC Power and SPEC jbb2005

SPECpower_ssj2008 [SP08] was released in late 2007 and is based on SPEC jbb2005. “SPECjbb2005 is a software benchmark product developed by the Standard Performance Evaluation Corporation (SPEC), a non-profit group of computer vendors, system integrators, universities, research organizations, publishers, and consultants. It is designed to measure a system's ability to run Java server applications. SPECjbb2005 is closely based on SPECjbb2000, which was inspired by the TPC-C benchmark and loosely follows the TPC-C specification for its schema, input generation, and operation profile. SPECjbb2005 replaces database tables with Java classes and replaces data records with Java objects. The objects are held in memory as Java Collection instances or other Java data objects. SPECjbb2005 is implemented as a Java 5.0 application emulating a 3-tier system with emphasis on the middle tier. All three tiers are implemented within the same JVM (though the benchmark may spawn several JVM instances).

The system modelled is still a wholesale company, with warehouses that serve a number of districts. There are a set of operations that customers (also known as users) initiate, such as placing new orders or requesting the status of an existing order. Additional operations are
generated within the company, such as processing orders for delivery, entering customer payments, checking stock levels, and requesting a report on recent activity by a given customer. There is only one user active per warehouse. A warehouse is a unit of stored data. It contains roughly 25 MB of data stored in Java Collection objects. Users map directly to Java threads. Each thread executes operations in sequence, with each operation selected from the operation mix using a probability distribution. As the number of warehouses increases during the full benchmark run, so does the number of threads. SPECjbb2005 is totally self contained and self driving (generates its own data, generates its own multi-threaded operations, and does not depend on any package beyond the JRE). SPECjbb2005 is memory resident, performs no I/O to disks, has only local network I/O, and has no think times.” [SPEC jbb]

The workload architecture of SPECpower_ssj2008 is a 3-tier system with emphasis on the middle tier. These tiers are comprised as follows:

1. Threads: Random input selection
2. Business logic
3. Warehouse: Tables of objects (Java Collections), rather than a separate database

The workload at different load points between 0% and maximum throughput in 10% steps will be reached by injecting request from an external driver. For finding the maximum throughput three calibration intervals are prepend. The measurements are completed with measuring active idle power which is defined as the state where the system is ready to accept transactions, but none are being issued [SP07].
The benchmark harness running on a monitoring system consists of a collection of daemons for acquiring power values from a power analyser and temperature values from a sensor. Furthermore an external driver (control and collection system CCS) controls the load points via a workload driver (JVM Director) that is running on the system under test (SUT). This way the measurements of AC power can be correlated with workload information of the SUT [SP07].

To reach the maximum throughput in scalable servers multiple JVM instances have to be established. The load driver (JVM Director) is responsible for controlling the JVM instances and for injecting the workload.
SPECpower_ssj is a homogenous benchmark that exhibits a regular execution behaviour. Because of the use of a Java runtime environment and the focus on memory resident operations, SPECpower_ssj has no special requirements concerning CPU architecture or I/O capabilities beside a network interface. The benchmark requires at least 256 MB heap for each instance so there is a memory requirement of about 512 MB of main memory to run the benchmarks that can easily be fulfilled by modern commodity server systems. Therefore SPECpower_ssj is a good benchmark to assess the energy efficiency for a single server dedicated to a CPU bound application scenario. However the results from SPEC Power are not representative for I/O intensive scenarios and consolidated environments.

5.4.2 SPEC Virtualization/ VMmark

To assess the performance of hardware and to capture key performance characteristics of the VMware virtualization environments, VMware released the VMmark benchmark [VM06]. Furthermore the SPEC group is working on a release of SPEC Virtualization to assess, and to compare a variety of virtualization environments [SV08].

VMmark is based on workload tiling. A tile is a collection of typical data center workloads (Mail, Java, Web, Database, File, Stand-by). Depending on the performance of the server, several tiles are executed. Each tile requires an own client machine that submits requests. All the workload runs at less than full utilization of their virtual machine. The score for each workload is collected, normalized according to a reference system, and aggregated. Some of the software required to run VMmark is free, some is available in no-cost evaluation versions, and some requires paid licenses. A sophisticated VMmark harness is required to submit requests to the target server and to score the performance of individual virtual machines.

Virtualization benchmarks are very well suited to execute the workload for energy efficiency analysis. However a general supplier independent benchmark that can assess a variety of virtualization environments (e.g., Xen,
VMware, Microsoft Virtual Server) does not yet exist. SPEC has just started the development of a virtualization benchmark. However prototypes are not to be expected before late 2008 or 2009 and are therefore not relevant within the time frame of the project “Efficient Servers”. Furthermore the development of a measurement harness for complex application scenarios is costly. Therefore we can’t expect a benchmark to assess the energy efficiency of virtualized/consolidated environments in the near future.

5.4.3 Energy Star

ENERGY STAR for office equipment is a joint international program originally initiated by the U.S. Environmental Protection Agency and the U.S. Department of Energy. In 2002 the EU has joined the programme by signing an official agreement. The development of energy efficiency criteria for servers has been started in 2006 when desktop derived servers have been addressed for the first time as part of the requirements defined for computers. Since 2007 a separate specification has been developed for so called “volume servers” covering the most important product segment on the market in terms of numbers and energy consumption. A first version of this specification has been published recently (EPA2009).

5.4.3.1 Energy Star for desktop derived servers

The product category is specified as follows:

A desktop-derived server is a computer that typically uses desktop components in a tower form factor, but is designed explicitly to be a host for other computers or applications. For the purposes of this specification, a computer must be marketed as a server and have the following characteristics to be considered a desktop-derived server:

- Designed and placed on the market as a Class B product per EuroNorm EN55022:1998 under the EMC Directive 89/336/EEC and has no more than single processor capability (1 socket on board);

- Designed in a pedestal, tower, or other form factor similar to those of desktop computers such that all data processing, storage, and network interfacing is contained within one box/product;

- Designed to operate in a high-reliability, high-availability application environment where the computer must be operational 24 hours/day and 7 days/week, and unscheduled downtime is extremely low (on the order of hours/year);

- Capable of operating in a simultaneous multi-user environment serving several users through networked client units; and

- Shipped with an industry accepted operating system for standard server applications (e.g., Windows NT, Windows 2003 Server, Mac OS X Server, OS/400, OS/390, Linux, Unix and Solaris).

- Desktop-derived servers are designed to perform functions such as processing information for other systems, providing network infrastructure services (e.g., archiving), data hosting and running web servers.

5.4.3.2 Energy Star for volume servers

Energy Star has developed new criteria for volume servers with one or two CPUs. Energy Star further differentiates between non managed and managed servers (equipped with a service processor etc.). The requirements currently only cover power supply efficiency and energy efficiency in idle mode and are to be seen as a first step on the way to more comprehensive criteria. More information is provided in chapter 5.7.3.
5.5 Measurement Requirements for Benchmarks (Measurement Protocol)

5.5.1 Metering

The measurement of the energy consumption of servers is a non-trivial task, as we are facing high currents of high variance. Therefore we need meters that sample at a high rate and that can be attached to the equipment without affecting warranty. There is the option for DC and AC metering:

- DC power can be measured with a current meter embedded in the power supply (e.g., power supply management interface PSMI [PSMI07]).
- DC power is measured with the help of a current meter attached to the service processor (e.g., IBM Power Executive [IBM07]).
- AC power is measured with a power meter that is able to measure true power (True RMS, true root mean square). The meter is connected to an AC line voltage source. Normally the attachment requires a shutdown and reboot of the server.
- If a measurement including the UPS is acceptable, we can switch to battery power, attach the AC meter and switch back avoiding an additional shutdown/reboot.

The meter should report periodically (e.g., once per second) the average power over the last period to a data collection computer (e.g., a PC attached via USB, serial interface, network to the meter).

5.5.2 Test conditions

The test conditions (temperature, humidity) should be recorded. Especially the temperature can have an impact on energy consumption. However in the normal operation conditions of room temperature (18°-25° Celsius) we don’t expect significant variations in energy consumption.

5.5.3 Measurement Points

To assess the efficiency a different performance levels, the benchmark has to run at intermediate measurement points between 0% and 100%, as defined by the appropriate benchmark. Throughput based benchmarks are applicable for efficiency measurements because a certain utilization can be adjusted by injecting an appropriate number of requests. In some protocols the idle power (= 0% performance) has to be measured twice (before and after running the benchmark).

- The test driver can be located externally with a network connection to a stub in the SUT.
- The test driver is located internally, but might consume a significant amount of energy thus affect the measurement of the energy consumption.
- The maximum performance has to be determined.

The benchmark has to be adjusted to run at a number of performance levels (5% or 10% intervals between idle and max performance). In this way we receive a number of power performance points that are the basis for a power/performance metric.

5.6 Power/Performance Metric

Most of the benchmarks map the measurements sampled at multiple points to a single value that is labelled with a name reflecting the benchmark. Depending on the metric the single value reflects more the efficiency at low, average or high utilization.

A prevalent metric for power and performance with a focus on average utilization is defined as follows:

- The total performance metric for each of the distinct measurement segments is computed and these totals are summed
- The average power measured for each benchmark segment, including the AC-
Energy efficiency criteria and benchmarks

5.7 Assessment of Common Benchmarks

5.7.1 Relevance of Benchmarking

The value of standardized benchmarks should always be considered with the necessary care with regards to real results which can be achieved in the field.

Gartner Consulting discusses the area of benchmarking for server selection in their White Paper [GA]. Gartner states:

“Benchmarks have been established to set up a standard, meaningful way of measuring a computer system’s performance (or energy usage within the context herein) for comparative purposes during the evaluation of vendors’ server products.

Despite this noble goal and improvements to the actual benchmarks themselves, vendors continually have been able to tweak benchmark results that can favour their products.”

After describing benchmark approaches and shortcomings of specific benchmarks, Gartner proposes a road map to best practices for server selection:

Because a reliance on benchmarks could lead to an unsatisfactory purchase decision, it is suggested that enterprises look at server evaluation from a much broader perspective. Gartner has recognized a growing trend to deploy servers in specific application-dependent situations (such as enterprise resource planning, data warehousing and Web serving), which means that a server may be well qualified for one kind of deployment but not for others. It is therefore recommended to use an evaluation tool or model that allows multiple different application deployment types.

The evaluation is concluded with the statement:

“The goal of benchmarking organizations is to help enterprises evaluate products on a level playing field. In reality, all server vendors do what it takes to put out benchmark numbers that are most favourable to their own products, whether or not the resulting scores are achieved under realistic, real-world conditions.”

Looking into the future one should not expect that all vendors will report benchmark ratings for all of their server products. Typically only products which are best in class will find their way into public benchmarks reports. All other benchmark values eventually will be available to customers under Non Disclosure Agreements only.

5.7.2 SPEC Power

The SPECpower benchmark is an important example for a common approach of IT industry to the benchmarking and evaluation topic. Many server vendors and CPU developers are represented in the SPECpower committee, so all necessary stakeholders are involved in the process of defining the new metric for power resp. energy efficiency.

Not only a single number calculated according to the power/performance metric is published for the SPECpower benchmark but also the individual measurement points (see published results in [SP08b]). Therefore all relevant data is published to support an individual analysis with a focus on a specific load condition. For a server with high load, the emphasis might be on measurements of 80%, 90% and 100% load. For a dedicated server running on low load with long periods of idleness the idle power and the power at 10% load might be of great importance to assess the energy efficiency.

5.7.2.1 Sample Benchmark Results

The significance of published results can be shown with some sample results published at the SPEC web site [SP08b]. We selected three
servers using quad-core processors with 1, 2 and 4 chips.

- Fujitsu Siemens Computers PRIMERGY TX150 S6 (2.4 GHz Intel Xeon X3220)
  143 ssi_ops@100%

- Fujitsu Siemens Computers PRIMERGY RX300 S4 (2.8 GHz INTEL Xeon E5440)
  303 ssi_ops@100%

- Hewlett-Packard Company Proliant DL580 G5 (1.86 GHz, Intel Xeon processor L7345)
  359 ssi_ops@100%

Fig. 5.5 SPECpower_ssj2008 Results FSC TX150_S6, FSX TX300_S4, HP DL580_G5 [SPRa,b,c]

The overall value \( \left( \frac{\Sigma\text{ssi_ops}}{\Sigma\text{power}} \right) \) published in the results can just express the average power performance ratio. For the assessment of the energy efficiency in a real world scenario, the value of the idle power and the maximum throughput are relevant as well. The FSC servers TX150 and TX300 have a similar overall value (669 vs. 690 ssi_ops/watt) but differ in idle power (75 W vs. 166 W) and peak performance (143 ssi_ops vs. 303 ssi_ops) by a factor of 2. For a real-world scenario with low performance requirements the single socket system (FSC TX150) is superior by a factor of two because of its low idle power which is an implication of the modest chipset requirements for supporting just a single socket. For a higher load the dual socket system (FSC TX300) offers the best Performance to Power Ratio. If maximum scalability and peak performance is required a 4 socket server is needed. However the architectural requirements to support 4 sockets (HP DL580_G4) come a the cost of a high idle power (271 W) which cannot be compensated totally by the high energy efficiency of the Low Voltage CPU.

5.7.2 Future Work

After publishing the first version of SPECpower_jbb this benchmark may develop further to include I/O and will take more and different workloads (file server, mail and messaging server, VM environments) into account. Especially the technique of consolidation and virtualization promise the highest savings results (see section 3 of the EPA report to the congress [EPA07]). Therefore SPECpower should respect this in the future.

5.7.2.3 Applicability for Best Practice Cases

SPECpower_ssj is just qualified to assess the efficiency of a single system under a single homogenous workload. SPECpower_ssj does not cover consolidated/virtualized scenarios. It is therefore only of very limited value for the best practice cases to be carried out in the E-Server-project at a later stage.

5.7.3 ENERGY STAR

As has been outlined above a first set of Energy Star specifications has been published in May2009 [ES09]. These criteria are currently
focussed on energy efficiency of power supplies and on idle mode power only and therefore to be seen as a first preliminary step to a more comprehensive set of requirements.

The proposed requirements for power supply efficiency require a minimum efficiency for several load points (Tab.5.1, 10 %, 20%, 50%, and 100%).

Today in many areas, where for specific reason consolidation is not possible or desired (e.g. dedicated Servers etc.), workloads are quite low and extended periods of operation in idle mode are encountered. For that reason Energy Star addresses energy consumption in idle mode in its first set of criteria.

Tab. 5.2 shows the maximum allowed energy consumption in idle mode for servers with 1-2 CPUs according to the new Energy Star criteria.

### Tab.5.1 Minimum efficiency levels of power supplies

<table>
<thead>
<tr>
<th>Type of power supply</th>
<th>Nominal rated power</th>
<th>Load level (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>10</td>
</tr>
<tr>
<td>Multi-Output (AC/DC, DC/DC)</td>
<td>All power levels</td>
<td>-</td>
</tr>
<tr>
<td>Single Output (AC/DC, DC/DC)</td>
<td>&lt;500W</td>
<td>70</td>
</tr>
<tr>
<td></td>
<td>&gt;500-1000W</td>
<td>75</td>
</tr>
<tr>
<td></td>
<td>&gt;1000W</td>
<td>80</td>
</tr>
</tbody>
</table>

The requirements for low load (e.g. 10% and 20%) will be relevant for example for blade servers with low configuration in off-peak hours.

Since for technological reasons it is easier to achieve higher efficiency for larger power supplies, a classification into power level categories has been introduced. To ensure that small power supplies are not put at a disadvantage lower requirements have been set for the category <500W.

Overall efficiency of a power supply in a specific server depends on the concrete workload requirements and consequently on maximum and average workloads. Such a more comprehensive consideration however only can be done when a benchmark, addressing energy consumption at different loads, has been introduced.
Energy efficient servers in Europe

For servers with 3-4 CPUs only qualitative criteria have been defined. Quantitative criteria will be part of the version 2 requirements. The four server categories shown in Fig. 5.2 differ regarding number of CPUs, the availability of a service processor and the capacity for redundant power supplies.

The values given in the table have been defined for basic configurations covering one power supply, 4GB RAM and a hard disk. For larger configurations including additional disks, memory, I/O slots etc. functional adders are defined specifying the additional allowance for power demand (Tab. 5.3).

In contrast to SPECpower Energy Star provides concrete requirements regarding the configuration and the measurement procedure for the determination of idle power. It is also requires that the idle power is determined with the operating system normally used for the specific server type.

The server configuration and the energy data have to be provided by manufacturers in the product data sheets (see www.energystar.gov, www.eu-energystar.org) and in the web based information.

As discussed in previous chapters the idle power is highly dependent on scalability and peak performance issues. Therefore different idle power limits are required depending on number of CPUs as well as hardware configuration in general.

For the second tier of the specifications expected for 2010 a power/performance benchmark shall be introduced. The benchmark may be based on SPECpower if by that time this benchmark will address different relevant workloads and therefore will be more flexible for practical application.

### Table 5.2 Limits for the power demand in idle mode

<table>
<thead>
<tr>
<th>Server Type</th>
<th>Limit for power demand in Idle-Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>Category A: Standard Server 1 CPU</td>
<td>55 Watt</td>
</tr>
<tr>
<td>Category B: Managed Server 1 CPU</td>
<td>65 Watt</td>
</tr>
<tr>
<td>Category C: Standard Server 2 CPUs</td>
<td>100 Watt</td>
</tr>
<tr>
<td>Category D: Managed Server 2 CPUs</td>
<td>150 Watt</td>
</tr>
</tbody>
</table>

(basic configuration: 1 hard disk, 4GB RAM, 1 power supply, 1Gbit Ethernet)

Standard Server: no service processor or management controller, no option for redundant power supply

Managed Server: Service processor or management controller available, option for redundant power supply

### Table 5.3 Functional adders for additional hardware components

<table>
<thead>
<tr>
<th>Additional component</th>
<th>Allowed additional power demand</th>
</tr>
</thead>
<tbody>
<tr>
<td>Power supplies &gt;1</td>
<td>20W/PSU</td>
</tr>
<tr>
<td>Disks &gt;1</td>
<td>8 Watt/disk</td>
</tr>
<tr>
<td>RAM &gt; 4GB</td>
<td>2 Watt/GB</td>
</tr>
<tr>
<td>I/O components &gt; 1 Gbit</td>
<td>No add. allowance</td>
</tr>
<tr>
<td>- additional Ethernet &lt; 1Gbit</td>
<td>No add. allowance</td>
</tr>
<tr>
<td>- Ethernet &gt; 1Gbit</td>
<td>2W per port</td>
</tr>
<tr>
<td>- Ethernet &gt; 10Gbit</td>
<td>8W per port</td>
</tr>
<tr>
<td>- Fibre Channel</td>
<td>5W per component</td>
</tr>
</tbody>
</table>

Overall the first version of Energy Star requirements for servers provides a basic tool addressing some aspects of energy efficiency for volume servers used in low load scenarios. The requirements do not yet cover hardware appropriate for consolidated systems as energy consumption at relevant load levels is not addressed.

A more comprehensive set of criteria which is expected to address larger servers as well as blades and storage units will be developed for tier 2 which is expected for 2010.
6 Best practise cases

6.1 Assessment of best practice cases

The most valuable method to convince customers to choose a specific (energy efficient) technology is the use of reference cases describing similar projects. This observation also has been made in the small market study conducted as a part of the project. Customers are interested to see real results in real environments.

So, the most valuable input for parties interested in energy efficiency and energy savings can be given by best practice cases from

- real customers
- real application environments
- a comparison of energy usage before and after applying measures to improve energy efficiency.

6.2 Restrictions for best practice cases

Suppliers are dependent on the requirements and boundary conditions set by the customers. It is not possible to conduct field demonstration projects only for the purpose of demonstrating energy savings. Because of costs and time.

For real best practice cases, suitable customers and renewal or modification projects will be identified allowing the implementation of the defined energy saving measures. Such projects are always in context of business processes of the customer and in each case they are under pressure of time and budget. Therefore such a measurement has to be as non-disruptive as possible.

In fact these restrictions lead to the best and most valuable approach: assessment of real energy savings in real customer environments. Based on real customer workload, a comparison of an old and a new IT infrastructure can be done. This gives other interested parties a much better view on possible energy savings in a real environment than benchmarks ever could. According to Gartner, such kind of assessment is a much better criterion for server and technology selection than just benchmarks.

6.3 Settings for IT-renewal projects from a generic point of view

In the following possible framework conditions for renewal projects are discussed:

- Case A: Complete replacement of IT hardware
- Case B: Partial replacement of IT equipment (subsystems)
Energy efficient servers in Europe

1. Case B1: Direct replacement of hardware
2. Case B1: Consolidation or virtualization of hardware

As the first case appears to be rather unchallenging as well as seldom in most of the business cases, we won’t discuss this item in detail.

Most likely according to ongoing modernization activities only subsystems will be affected by renewal projects. This implies that energy measurement must take place on server level instead of data center level. For Case B1 as the number of physical server will remain unchanged, direct correlation of workloads will be a simple task.

Presumably the significant majority of projects will apply consolidation and virtualization of servers. Consequential appropriate concepts for workload correlation must be considered as there is no chance to map workloads on a server to server basis.

There is also need to clarify the next issue. Where to apply the actual power measurement?

In principle we come upon this power supply chain:

- Mains power supply
- Fused outlet line
- UPS
- One Server (or several servers powered by the same UPS)
- Redundant Power supplies for one server

Starting from below the most comfortable and non-invasive method is to undertake measurement on DC level with a current meter embedded in the power supply, provided that server’s power supply is equipped with this facility. Thus this would not require any additional equipment.

More often measurements will be conducted on the AC side. In principle there are two alternatives available for the integration of the measurement equipment: between the power supplies and the UPS or between the UPS and the mains. The latter enables to insert or withdraw the power meter without any shut down of the server.

Quite common only one central UPS powers several servers or more general the entire IT hardware. So in this case it would be difficult or rather impossible to monitor a subsystem only for metering the power draw for the UPS.

6.4 Measurement equipment

Since the characteristics and chronological sequence of the workloads of the initial and final system will be more or less similar in the best case but definitely not identical there is a special need to compare patterns of utilization for the initial and final system as well as to cancel out some time dependent effects. Thus we have to monitor power consumption for a longer period, not only accumulating the energy consumed but to log a time series for power consumption values.

In some of the best practice cases an energy measurement and logging device newly availability on the market will be used: the Energy Logger 3500 from Voltcraft [EL3500].

![Energy Logger 3500](http://www.conrad.de)
The energy cost measuring device serves to measure and analyse consumption details of electrical devices. The measuring device is simply connected between the mains socket and the electric device.

The measuring device has an internal, non-volatile memory where the data for performance factor, current and power can be saved for up to 6 months. Via an SD card slot this data can be transferred to a computer for analysis. An enclosed analysis program allows a graphic illustration and further processing of the recorded data of the energy cost measurement device.

Freely available software to convert the logged data to .csv and Excel format is available via http://energycheck.meterstand.info/.

Technical Data Energy Logger 3500:

- Operating voltage: 230 V/AC
- Performance measurement display: 0.1 - 3500 W
- Accuracy:
  - 5 - 3500 W: (± 1% + 1 count)
  - 2 - 5 W: (± 5% + 1 count)
  - < 2 W: (± 15% + 1 count)
- Dimensions (LxWxH): 164 x 82 x 83 (mm)

In other cases rack based power distribution units (PDU) will cover the requirements, too, if they are equipped with a power monitoring facility.

6.5 Recommended Test Procedure

The evaluation of the case compares the energy consumptions and workload of the initial and final states of the System under Test (SuT). The final result of the evaluation is the percentage of energy consumed finally over energy consumed initially for the same workload. In typical consolidation projects with SME customers many types of workloads may be combined onto the final SuT which would make adjustments of throughput changes rather difficult.

The effort to assess the workloads applying “exact scientific procedures” may be unacceptable to the customer within a real life project. The in depth description might be hindered by confidentiality requests of the customer.
6.5.1 System description
The existing and the new hardware configuration (e.g., server, disk, peripherals) are described in detail. This includes all IT components (server and where necessary switches).

6.5.2 Runtime Environment
The configuration of the system and application software is described in detail.

6.5.3 Workload
The workload is described qualitatively (general description, CPU usage, memory requirements, I/O utilization) and quantitatively with the help of representative traces of system utilization (processors time, system time, user time, context switches), number of I/O operations (e.g., raw disk I/O operations), memory utilization, and in-/outgoing network packets. The sampling rate and the duration of samples should respect changes in the workload. E.g., to cover a complete load period, we have to sample one typical week at a high sampling frequency (e.g., one sample per second). The system monitoring tools from Windows (perfmon) or Unix/Linux provide the required information. However the monitoring has to be configured and might need some support by scripting to store the required data.

The workload should also be quantified on the application level. Here the services have to be quantified by counting the number and type of answered requests (e.g., delivered web pages, forwarded emails, processed messages, processed queries, …). The collected data has to be accumulated without interfering with the load characteristics of the SUT (e.g., the accumulation should not require too much CPU-time) and may not contain sensitive data that is not allowed to leave the customer site.

- The workload of the final system has to be equivalent to the initial system in structure and timing (same application characteristics but different OS and application possible to allow consolidations with a change of OS). In some setups, this could be achieved by copying the requests to another port of a router connected to the final SuT and to discard the reply of the final SuT.

6.6 Measurement procedure – step by step
The acquisition of power and performance has to be done for the initial as well as for the final system. If there are dependencies among multiple servers, they have to be analysed in parallel. Parallel data acquisition requires synchronization. This can be done by adding time stamps to all readings or by a consolidation of all data acquisition to a single machine.

The following procedure is recommended:

- **Measurement of the idle power of the final SuT**
  which is not yet under normal operation and workload. Therefore the server has to run without serving any requests from the outside.
  A measurement with active and disabled power management should be conducted, if power management is supported at all.

- **Measurement of the initial SuT under the customer workload**
  Energy consumption has to represent a typical weekly consumption. Normally, the SuT should be measured over a whole week to record the complete cycle of load conditions. The measurement accuracy has to be documented (power meter) and is expected to be <=5%.

  The throughput, CPU and memory utilization, as well as disk/network activity has to be reported with system monitoring tools.

  The sampling frequency of the system monitoring should be the same as the measurement frequency to correlate both information (e.g., one sample of utilization and one measurement of power per second)

  The accumulated energy consumption should be recorded periodically (e.g., once per minute).

- **Switch from initial to final system**
Best practice cases

- **Measurement of the final SuT under the customer workload**

Energy consumption has to represent a typical weekly consumption. Normally, the SuT should be measured over a whole week, at least over 24h. The measurement accuracy has to be documented (power meter) and is expected to be <=5%.

The throughput, CPU and memory utilization, as well as disk/network activity has to be reported with system monitoring tools.

In case of consolidation efforts the overhead of virtualization (e.g., CPU time spend in the VMM layer) should be reported if this does not impose technical or administrative hurdles.

The sampling frequency of the system monitoring should be the same as the measurement frequency to correlate both information (e.g., one sample of utilization and one measurement of power per second).

If additional other workloads are added to the final SuT, they have to be assessed in addition to the initial workloads and the conversions have to be agreed individually.

- **Measurement of the idle power of the initial SuT**

which is already out of normal operation and that does not serve any requests anymore. A measurement with active and disabled power management should be conducted, if power management is supported at all.

It is not clear in advance if such formal methodology could be completely applied in all best practice case. Since measurement of real customer environments interfere with customers business operations it should be expected that in most cases the methodology could be applied in parts only. Despite of this situation at customer side the consortium partners are working on delivering sustainable data from real measurements.

<table>
<thead>
<tr>
<th>System / APV feature</th>
<th>VIOSS V1.1, V1.2, V1.3</th>
<th>VIOSS Linux</th>
<th>ISOS</th>
<th>AIX SL V5.2 ML2</th>
<th>AIX SL V5.3</th>
<th>RHEL AS V3</th>
<th>RHEL AS V4</th>
<th>SLES V9</th>
</tr>
</thead>
<tbody>
<tr>
<td>Partitions with dedicated processors</td>
<td>X</td>
<td>X</td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Micro-partitions</td>
<td>X</td>
<td>X</td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Virtual ITY</td>
<td>X</td>
<td>X</td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Virtual console client/server</td>
<td>X</td>
<td>X</td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Virtual Ethernet</td>
<td>X</td>
<td>X</td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Boot from virtual Ethernet</td>
<td>X</td>
<td>X</td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Shared Ethernet Adapter</td>
<td>X</td>
<td>X</td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Ethernet bridge with STP support</td>
<td>X</td>
<td>X</td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Virtual SCSI server</td>
<td>X</td>
<td>X</td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Virtual SCSI client</td>
<td>X</td>
<td>X</td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Boot from Virtual SCSI client disk</td>
<td>X</td>
<td>X</td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Virtual Tape</td>
<td>X</td>
<td>X</td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Boot from Virtual CD</td>
<td>X</td>
<td>X</td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Partition Load Manager</td>
<td>X</td>
<td>X</td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Integrated Virtualization Manager</td>
<td>X</td>
<td>X</td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Dynamic LPAR CPU</td>
<td>X</td>
<td>X</td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Dynamic LPAR RAM</td>
<td>X</td>
<td>X</td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Dynamic LPAR physical I/O</td>
<td>X</td>
<td>X</td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Dynamic LPAR virtual adapters</td>
<td>X</td>
<td>X</td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>On-line scan of Virtual SCSI devices</td>
<td>X</td>
<td>X</td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
</tbody>
</table>
6.7 Best Practise Cases in the e-Server project – summary of central results

6.7.1 General aspects

As one major objective of the E-Server project a number of best practise cases has been conducted to show the feasibility of energy efficiency measures in typical IT projects. It was intended to cover different types of hardware and software solutions for different environments in small, medium size and large service companies.

The cases should show successful practical solutions and serve as a tool for motivation and information of IT managers. All case studies involved a comprehensive documentation of energy consumption and workload patterns prior and after setting of measures. Detailed technical information is available in a separate brochure on best practise cases (see also download at www.efficient–servers.eu).

6.7.2 The case studies

**STRATO AG**

STRATO AG (Berlin) is the second largest European web hosting provider and operates more than 30000 servers at two locations in Germany. STRATO conducted a project addressing hardware consolidation in combination with load balancing. The issue here was to replace existing standard hardware in the area of web-services by energy efficient hardware. The case study was based on hardware from SUN Microsystems.

The specific case study showed extreme energy savings of more than 90% in the new system compared to the old system. The high savings could be achieved due to the specific IT environment and the IT services addressed. Thus savings of this order of magnitude can not generally be expected for any type of IT environment.

The project also showed very good economic efficiency since pay back for the new server hardware directly from the energy cost savings is less than 3 years. Thus the measures are highly efficient both from an ecologic and an economic point of view.

**German Ministry of Environment**

The German Ministry of Environment (Berlin) conducted a larger project for a complete IT renewal. Part of the project was monitored and analysed in the E-Server project. The overall goal also specified in the call for tender was to increase energy efficiency by a minimum of 40% applying consolidation and virtualisation technologies. The case study was based on hardware from IBM.

The consolidation and virtualisation measures applied lead to energy savings of 60% compared to the original system could. It was estimated that additional measures addressing also network hardware could increase energy savings up to more than 80%.

Experiences from the project underlined the fact that a thorough analysis and planning prior to the project can strongly improve effectiveness and efficiency of measures. As in most projects addressing renewal of IT hardware also in the German Ministry of Environment several measures to improve performance and capacity of the system have been implemented. Additional performance and capacity often is desired due to the fact that new services are implemented. It normally compensates for part of the theoretical energy saving potential.

**City of Bad Soden**

The City of Bad Soden conducted a case which may be typical for IT infrastructures in the public service area of municipalities and thus provides a good example for smaller projects in the public service sector.
In addition to server consolidation and virtualisation also desktop virtualisation was addressed. The case study was based on hardware from SUN Microsystems.

To meet requirements regarding safety and availability redundancy has been introduced in the system concept which partly compensates the energy saving potentials possible in theory.

Nevertheless the Consolidation of servers combined with the application of desktop virtualisation based on a thin client concept allowed energy savings of 61%.

The City of Bad Soden received an award for this innovative IT project in the area of public services.

**Sozialwerk Nürnberg**

The Sozialwerk Nürnberg conducted a case which was based on consolidation from standard servers to blade servers. Furthermore energy efficiency potentials at the level of cooling have been addressed. The case study was based on hardware from IBM.

The case involved a switch from standard server technology (1U rack servers) to energy efficient blade servers and a new storage system. Energy savings of about 43% could be achieved.

The new hardware allows temperatures in the server room of up to 35°C. Consequently cooling is only necessary on very hot Summer days and energy consumption and costs for cooling could be reduced by 32%.

**Wincor Nixdorf**

Wincor Nixdorf in Paderborn is a worldwide supplier of software solutions for banks and insurance companies. At Wincor Nixdorf a consolidation from standard blade server technology to most energy efficient blade server technology was conducted. The case study was based on hardware from Fujitsu.

The project suggests energy savings of 75% as soon as the full capacity of consolidation is used. In the partial consolidation conducted during the project energy savings by more than 40% could be achieved.

**Austrian Energy Agency**

At the Austrian Energy Agency (Vienna) a case has been conducted addressing both consolidation options and power management of virtualised systems. The case study was based on hardware from Fujitsu.

The overall energy savings in the project were well beyond 40%. New power management features of virtualisation software allow migration of virtual servers when not in use. Thus virtual servers can be consolidated on fewer server hardware allowing temporary shut down and energy savings. The experiences from the project showed that additional energy savings in the order of magnitude of 7-10% can be achieved by use of these power management options.

However to date this type of power management is not yet provided as certified standard feature but still in an experimental stage. Improved features with high reliability shall be available in the next versions of the virtualisation software. At present software licences providing the relevant features are costly suggesting that they will be used only by larger companies.

Overall power management options primarily used at weekends and at night provide significant further saving potentials to be addressed in a second step after consolidation and virtualisation.

**Encontrol**

Finally at Encontrol (Zürich) an example for a case in a very small enterprise has been conducted to check the potentials for savings in small IT systems. The case study was based on hardware from HP.

The case study showed that the saving potentials to be accessed by virtualisation and consolidation in very small companies are quite limited and tend to be compensated by additional investments or increases in performance and capacity.
In a 1:1 consolidation savings of 24% would have been achieved. However this increase in energy efficiency has been compensated by additional capacity for new services as well as for redundancy.

6.7.3 Conclusions from cases

Overall the best practice cases demonstrated energy savings of up to 90% and of approximately 60% on average. The case studies consequently confirmed the average energy saving potentials indicated in the market study.

Overall the results showed that high saving potentials can be exploited in data centres or server rooms of very different types of companies. New technologies providing the highest saving potentials are consolidation and virtualisation. However power management at server unit level shows good potential as well especially if applied for already consolidated, virtualised systems.

Experiences from best practice cases show that the degree of improvement of efficiency clearly depends on a proper analysis and planning of projects. The experiences from the cases also suggest that measures to improve energy efficiency whenever possible should be linked to projects of hardware or infrastructure renewal. Measures to improve energy efficiency often can not be initialised for this specific purpose but can easily be linked to “standard” IT projects. In this case optimization of processes, hardware and energy efficiency go hand in hand.

A major shortcoming encountered in many companies today is the missing documentation of energy consumption at an appropriate level. The best practice cases also showed the high importance of the concrete documentation of the energy consumption to be able to show specific effects and benefits of measures.

All external project partners which have contributed best practice cases confirmed their positive experience and benefit. In some cases this was an initial step to further exploit energy efficiency potentials for example in the area of storage, network equipment and infrastructure.
7 Recommendations for hardware procurement

Based on the current technologies and options for assessment of technologies regarding energy efficiency a number of recommendations can be given for the procurement of new equipment.

In IT hardware renewal processes there are different ways to improve energy and cost efficiency, primarily:

- Selection of energy efficient hardware
- Consolidation and virtualisation
- Consideration of power management capabilities

Consolidation aims at a more efficient usage of the server hardware as well as a facilitated server management. Measures to improve process efficiency thus can be easily combined with the goal of higher energy efficiency.

Power management means to power down servers or server components at times of low workload. The practical application of power management is possible if the applications allow slightly delayed response times and performance of hardware components and if power management features are supported by the hardware and software. Virtualisation offers other new options of power management involving migration and shut down of virtual servers (see section on power management).

In the following the options in the three areas of measures will be shown and practical recommendations will be provided.

7.1 Energy efficient server hardware

Energy efficient server hardware is based on energy efficient server components such as power supplies, CPUs, hard disks; memory and I/O-slots. Currently also more efficient motherboards are being developed. Here the idea is to minimize transformation losses on the board from input power to power at component level.

**Tab. 7.1: Examples for the power draw of different hardware components (according to [FAW07])**

<table>
<thead>
<tr>
<th>Hardware component</th>
<th>max. power (W)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CPU</td>
<td>80</td>
</tr>
<tr>
<td>RAM</td>
<td>36</td>
</tr>
<tr>
<td>Disks</td>
<td>12</td>
</tr>
<tr>
<td>I/O Slots</td>
<td>50</td>
</tr>
<tr>
<td>Motherboard</td>
<td>25</td>
</tr>
<tr>
<td>Fans</td>
<td>10</td>
</tr>
<tr>
<td>Power supply</td>
<td>38</td>
</tr>
</tbody>
</table>
Table 7.1 shows the approximate power draw of individual hardware components indicating that CPUs, power supplies, RAM and I/O slots are the components with the major power demand.

To consider energy efficiency in procurement the IT-manager needs tools to assess the energy efficiency of the hardware available on the market. Currently the selection of energy efficient equipment is still a challenging task since the development of internationally accepted standards has just started. Criteria and benchmarks available so far cover only part of the efficiency aspects. A general transparency regarding energy efficiency of server equipment thus is still to be developed.

**Recommendations for practise:**

- **Consideration of the Energy Star Label**

  For the procurement of smaller volume servers especially for low load situations Energy Star labelled products can be considered. Labelled equipment will contain energy efficient power supplies corresponding with the specifications given above. As of May 2009, Energy Star compliant servers will be labelled in the technical documentation of manufacturers and special technical data sheets will be available, providing specific information on configuration and energy consumption. As indicated above Energy Star currently addresses only the product segment of smaller volume servers. It is therefore not generally possible to base procurement on labelled equipment. In many cases however, the Energy Star efficiency criteria may be used as a starting point to define individual procurement criteria (see also following points).

- **Consideration of Energy Star criteria in procurement tenders**

  Efficiency criteria for power supplies defined by Energy Star can be used as a basis for procurement tenders. This is effective in particular if a large number of equipment is being ordered and suppliers therefore see an incentive for offering specific configurations. Energy Star minimum efficiency levels are met by about 25% of the products available on the market. If a specific focus is to be put on energy efficiency issues the criteria can be set more even demanding than required by Energy Star, for example addressing a level which is only met by 10% of the equipment.

- **Adequate sizing and configuration**

  Server hardware configuration and upgrading options should be selected according to the real practical requirements. Overcapacity and extreme options for upgrading always lead to oversized energy supplies and lower energy efficiency. The necessary maximum power is finally defined by the CPU, the number of I/O-Slots, disks and memory. As hard-disks normally (except for very small systems) are outsourced to separate storage units, the disk storage capacity in the server can be kept relatively low.

  During the average operation cycle of volume servers (4-5years) upgrades are rare, except for additional memory. Normally the upgrade goes with the renewal of the server. Upgrading options far beyond the capacity currently needed therefore are not necessary.

- **Consideration of complete SPEC profile**

  For the procurement of volume servers especially for non I/O-intensive applications SPECpower can be used as a general indicator of energy efficiency. In any case the complete SPECpower profile i.e. the performance and the power demand at the different load levels has to be considered. The average SPEC value is less meaningful for practical application.

- **Detailed consideration of hardware configuration**

  Depending on hardware configuration SPECpower values will be different. SPECpower-values published are often determined at low configurations, so in practise for typical average configurations a higher energy consumption has to be expected. The configuration of the hardware tested is provided in detail in standardised SPEC data sheets. This
Recommendations for hardware procurement

information has to be considered in detail to arrive at meaningful interpretations.

- **Area of application and limitations**

The current version of SPECpower is a CPU-oriented benchmark. SPECpower results are therefore not to be extrapolated to all types of applications. I/O intensive applications for example are not covered. This should also be considered in practical considerations. A more comprehensive SPECpower benchmark set, which will be broadly applicable, has been announced for late 2009.

- **Currently limited availability of data**

SPECpower is only published for a minor part of the server models available on the market. Consequently data for specific models have to be requested from manufacturers in procurement tenders.

- **Consideration of energy efficiency at different loads**

Depending on the area of application of the specific server and thus the typical workload, different parts of the SPECpower profile may be considered. The upper load levels are more relevant in context with consolidation approaches whereas efficiency at lower loads is to be considered for equipment where low workloads and extended periods of idle operation can not be avoided.

- **Consolidation as the primary tool of choice**

In case of low workloads the option of consolidation on a lower number of servers should be evaluated (see section 4).

- **Energy Star criteria as a basic tool for the procurement of volume servers for low load applications**

For the procurement of volume servers addressing applications with low workloads and extended periods of idle operation (where no consolidation is possible or desired for some reason) the first criteria set from Energy Star can be a useful tool for hardware selection.

- **Extended SPECpower benchmark set and Energy Star Version 2 as tools for consolidated systems and higher workloads**

For the selection of hardware for consolidation concepts, where higher workloads are achieved and different types of applications are addressed which can not be covered by the currently existing tools the new version 2 benchmark and criteria sets of SPECpower and Energy Star becoming available in 2009-2010 will be the most relevant instruments to be considered. For further information regarding consolidation scenarios see chapter 3.3.

### 7.2 Energy efficient storage solutions

Storage capacity and storage use also strongly contribute to the energy demand of IT systems. In the past storage capacity and related power consumption have been steadily growing due to several driving factors. One is the fastly growing storage density of standard storage media coupled with a cost reduction per gigabyte. Second the growing number of IT-based processes in companies as well as data intensive applications contribute to a dramatic increase in demand for storage. Third a non economic use (for example redundant storage of data) enforces the demand for storage. Today often storage capacities at the level of several terabytes are used.

The energy consumption of storage components in data centres can amount to 30-40% of the total IT-equipment. The energy consumption of a hard disk of several GB capacity amounts to roughly 20W. Overall larger (and slower) hard disks consume significantly less power per GB of storage. For example a consolidation of three 300GB (10k rpm) disks to one 1 TB (7,2k rpm) disk can allow energy savings in the order of magnitude of 75%.

However optimization of storage solutions should not start at the level of the selection of specific hard discs. Concerning the optimization and consolidation of storage first of all the performance requirements as well as many
other criteria have to be clarified. In many cases only 20% or less of the data volume stored is frequently accessed. Consequently it is important to assess and categorise used storage capacities according to the service level required which includes response time /access, performance, security, confidentiality, downtime, recovery etc.

Often a significant part of the storage capacity (e.g. 20% or more) is blocked by duplicate data (not defined and used as backup) which is a very inefficient way of utilization. Thus an important step in the improvement of storage efficiency is de-duplication. A data reduction of up to a level of 1:30 may be possible. De-duplication of data also allows check for consistency, easier administration of backups and use of free disk space for part of the backups.

De-duplicating is not the only important measure to be applied prior to the design of new storage concepts. In many companies there are lots of data kept in archives or even on disks which will never be used again. Consequently a rigorous clear out of obsolete data also is essential for exploiting efficiency potentials.

Based on the basic analysis of storage categories and a consolidation by de-duplication and removal of irrelevant data appropriate solutions for primary storage, archiving etc. can be evaluated and selected. The selection of the preferred technology has to be based on application requirements. High-Performance storage systems (with higher energy consumption) are often only required for a relatively small part of the system.

Table 3.7 shows different types of storage technologies and the specific advantages and disadvantages. Storage area networks (SAN) and network attached storage (NAS) allow an improved utilization of storage capacity, reduced administration costs and simple and fast backup and expandability compared to direct attached storage (DAS). Total cost of ownership (TCO) of SAN and NAS concepts are often less than 50% of DAS [ML07]. It furthermore has been found that the majority of storage costs is due to administration while hardware costs are often less than 20% [Ga07]. Consequently additional investment in efficient hardware is easily charged off by savings in management and administration.

Specific software for storage resource management [SRM] can improve the utilization and archiving to 50%-70%. SRM provides transparency how much storage capacity is required and what the areas of demand are.

<table>
<thead>
<tr>
<th>Storage Type</th>
<th>Advantage</th>
<th>Disadvantage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Direct attached storage (DAS)</td>
<td>Low purchase cost</td>
<td>Bad efficiency regarding utilization of capacity</td>
</tr>
<tr>
<td></td>
<td>Simple administration</td>
<td>High operation cost</td>
</tr>
<tr>
<td>Network attached storage (NAS)</td>
<td>Fast solution for multiple clients</td>
<td>Lower performance</td>
</tr>
<tr>
<td></td>
<td>Simple administration</td>
<td></td>
</tr>
<tr>
<td>Dedicated storage area network (SAN)</td>
<td>Low cost centralised storage</td>
<td>Restricted bandwidth</td>
</tr>
<tr>
<td></td>
<td>Disaster recovery option</td>
<td></td>
</tr>
<tr>
<td>Fibre channel storage area network (SAN)</td>
<td>High performance and availability centralised storage</td>
<td>Higher purchase costs</td>
</tr>
<tr>
<td></td>
<td>Disaster recovery option</td>
<td>More complex administration</td>
</tr>
</tbody>
</table>
An important strategy to improve energy efficiency is the storage of less frequently used data on tape. In most companies large amounts of data are kept accessible on disk. This requires an unnecessary high and steady expansion of disk capacity.

Another option to improve energy efficiency are MAIDS (massive arrays of idle disks for the purpose of archiving). Rarely used disks can thereby be powered down. MAID stands for a disk array which allows to operate individual disks in an energy efficient way. This is done by parking the read head or by a reduction of the drive speed. MAID is applied in areas where data is accessed rather rarely. Examples are Virtual Tape Libraries and systems for archiving.

To avoid potential problems due to rare activation of disks, disks are activated from time to time based on a monitoring mechanism.

**Recommendations for practise:**

The following recommendations should be considered in IT renewal processes and the design and procurement of new storage solutions:

**Analysis of initial situation**

- analysis of current use of storage in the company is an important prerequisite. Do a segmentation into different types of data (frequently used, duplicate data, rarely accessed data, unused data etc.)

**Consolidation of unused data**

- Reduce data volume by de-duplication and removal of irrelevant old data.

**Selection of appropriate general concept**

- Select appropriate concepts for specific situation considering NAS, SAN, DAS as well as backup on tapes etc. Consider TCO as well as energy consumption and costs.

**Use of more energy efficient hard disk technology**

- use smaller number of high capacity disks and slower disks (7,200, 10,000 rpm).

**Appropriate storage solutions for back-up**

- store less frequently used data and backup on tape
7.3 Consolidation and virtualisation

7.3.1 General aspects
Consolidation of server hardware today is often synonymously referred to as virtualisation. However virtualisation is not the only approach to consolidate equipment. In many cases for example unused server capacities simply can be turned off without any loss of service.

For that reason in any case it is advisable to analyse the whole equipment in the company with regard to the degree of utilization (CPUs, hard disks, network). Often some servers can be identified which can be easily shut down without any losses.

On the basis of the degree of use of server hardware it can be assessed what capacity is really needed to maintain the required service levels as well as certain flexibility in performance and demand. This provides the basis for a concentration of applications in consolidation projects.

Overall virtualisation and consolidation projects are rarely carried out to improve energy efficiency only. It is generally a major goal to improve the use of capacities and to optimize the management of services to reduce infrastructure and administration.

Consolidation and virtualisation therefore provides multiple benefits which have to be considered and evaluated in an overall assessment

In the past especially with Wintel based systems one server per application was used to guarantee system stability and performance. With modern technology and virtualisation the parallel operation of several applications on one server is possible without any problems. Virtualisation nowadays is used for very different applications such as database, high performance systems as well as desktops.

The energy saving potentials based on virtualisation are in the order of magnitude of >70-80%. The effective savings in virtualisation projects depend on the degree of virtualisation, as well as on the degree new applications and features are implemented. Often consolidation goes together with expansions and upgrades which partly compensate the energy savings.

One problematic side effect of virtualisation which may lead to inefficiencies in the course of time is in fact the very simple and fast way new virtual server instances can be implemented. Thus there ist the potential problem of a rapidly growing number of virtual servers, what consequently again leads to an increasing demand for resources. Thorough planning and management is therefore highly critical to avoid undesired counter acting effects.

There are also software tools available now which support the evaluation of the economic efficiency of consolidation processes. Such tools (e.g. “Data Center Intelligence Software”, CIRBA) besides an analysis of TCO and ROI also allow addressing the aspects energy consumption and energy costs. Such software tools however often are costly (e.g. for 100 server instances annual costs in the order of magnitude of 25000 € can be expected).

In general a comprehensive cost assessment which includes energy costs should be part of any procurement process.

7.3.2 Basic preparation for consolidation and virtualisation projects
Projects for consolidation and virtualisation as indicated above allow a substantial improvement of energy efficiency and an optimisation of the management of IT resources. As already outlined it is important to check options for consolidation whenever a substantial renewal of hardware is planned.

To address efficiency potentials effectively first of all an inventory of the IT services used in the company should be established. Typically a categorisation into mission critical services,
Recommendations for hardware procurement

production services etc. can be done which is an important basis for the design of the consolidation approach. For mission critical services maximum protection and minimum downtime are critical and clearly more important than degree of utilization and energy efficiency. Consequently these applications are not primary candidates for consolidation.

The first step in the planning process is an inventory of the computing resources and the application workloads. Several types of information are relevant in this context:

- The IT services provided / applications running
- Responsibility for the server
- Requirements regarding service level and protection
- Location of the server
- Typical workload
- Technical information:
  - Model
  - Processor types and detailed specification (sockets, speed, cache)
  - Memory specification (size, modules)
  - Storage specification
  - Operating system
  - Network components (number and speed of ports)

This type of information should be collected for all hardware equipment under consideration. Some aspects may limit options for consolidation and virtualisation in some areas as for example:

- highly variable, unpredictable workloads
- strong requirements regarding security and privacy

- technical incompatibilities or potential problems at hardware and software level regarding virtualisation
- etc.

A next step in the planning process is the assessment of the computing power necessary to cover the services already implemented or new services. This is often done by simply multiplying the computing capacity available by the typical loads encountered: number processors x number cores/proc. x max. frequency x max. percentage utilization.

According to a broad survey regarding capacity planning [GG09] the typical computing demand for a server varies between roughly 250 and 600 MHz from low load to peak load. For a modern server with four cores running at 3,2 GHz the maximum computing power would be 12,8 GHz. Requiring that the computing demand should not exceed 50% of maximum capacity (e.g. 6,4 GHz) about 10 of the typical workloads indicated above could be consolidated on one server.

Requirements regarding memory also have to be estimated for the new consolidated system. A typical consolidated system with for example 10 virtual servers can normally easily do with 16GB RAM. Utilisation of memory often is at a level of 40-50%.

Network interface capability is not really a decisive factor in this context since current utilization rates are often very low (<1%) and thus even very high consolidation ratios are unproblematic regarding this criterion.

In practise consolidation ratios of 10-15 are unproblematic and safe for many areas and applications.

A more detailed analysis of the workloads on the servers to be consolidated allows to assess the typical different load patterns and to optimize combinations and consolidations of complementary loads (combining loads which have their peaks and lows at different times of the day or week). To get a statistically comprehensive workload pattern sampling over a week including the weekend is necessary. Sampling over a month in some cases may
provide useful additional information on variable loads if necessary. Per minute sampling usually is sufficient. Inexpensive equipment allows data logging in minute intervals over a week or even a month.

In any case the target utilization levels as well as the service level of the consolidated solution have to be defined. For many areas utilisation levels of 50% or higher may be possible.

Regarding availability also concepts for redundancy are an important issue. As consolidation and virtualisation also increases the risk that a higher number of services is affected in case of failure and appropriate redundancy concepts have to be implemented to guarantee availability.

**Recommendation for practise**

In case of hardware renewal and procurement the following points are generally recommended to assess options for consolidation and virtualisation:

- generally consider consolidation and virtualisation as a primary option to improve efficiency at several levels (management, infrastructure, energy consumption etc.)
- assess and categorize servers and applications according to central criteria which allow an evaluation of consolidation/virtualisation approaches:
  - Mission critical applications, production applications etc.
  - Privacy, security
  - Workload variability and degree of utilization
  - Technical limitations at hardware or software level
  - etc.
- monitor and assess workload profiles over a minimum of a week as a basis for optimised consolidation of workloads
- choose servers/applications for consolidation/virtualisation where no specific risks or requirements are indicated. In initial stages servers with low service level objectives and high consolidation ratios are to be preferred.
- Define utilisation levels for the new system and ratio of consolidation (25%, 50%, 10:1, 15:1 etc.)
- Assess demand for computing power and memory etc. as indicated above (rough assessment based on MHz and GB)
- Consider service level requirements and design appropriate redundancy as security measure
- Assess and monitor improvements of efficiency prior and after consolidation (cooling reduction, floor space reduction, reduction of energy consumption, consolidation of load etc.)
7.4 Consideration of power management in procurement

7.4.1 General aspects
Power management for servers can be implemented at several levels. First of all the hardware component level offers power management features like frequency and voltage scaling of the CPU and power management for disks. Moreover power management on demand for complete server units is possible.

Standby-modes and low energy modes for servers in contrast to PCs and Laptops in the past often have been undesirable since delays for reactivation were not accepted. However in practise there are many applications where response times without any delay are not a real requirement. Many companies are operating a large number of servers where small delays for reactivation are completely non critical and consequently the use of saving potentials by power management could be easily justified. Standby or even shut down of servers should be possible for many services in small and medium enterprises especially during night time and at weekends.

For virtual servers there is a new option to migrate virtual instances during times of low load and temporarily shut down some physical servers. This concept also allows significant energy savings during weekends and at night. Further details are given below.

7.4.2 Power management at the level of components
Many processors used in servers nowadays according to information of manufacturers in principal do support power management (Optimized, Power Management, Speed Step Technology etc.).

However the proper function of the power management therefore is often not dependent on the CPUs and disks only but on chipset and the operating system. It is therefore important to address these issues already at the level of procurement and to make sure that hardware and software are delivered in a configuration which allows to the take advantage of power saving options. It should therefore be included to tenders, that the type of power management options is specified and that the proper configuration and compatibility of relevant hardware and software components is guaranteed.

The new Energy Star programme requirements for servers explicitly require that the options for power management are declared in the technical documentation of products and power management features are enabled per default when equipment is shipped. However this requirement currently is only required for volume servers with 3 and 4 CPUs but not for smaller and larger servers.

Systems must be shipped with power management enabled in the system BIOS, and/or the BMC or the service processor. All systems shipping with a preinstalled supervisor system (operating system or hypervisor) must also have this functionality enabled by default. In order to meet this requirement, all processors must be able to reduce their power use in times of low utilization by, either:

- Reducing voltage and/or frequency through Dynamic Voltage and Frequency Scaling (DVFS), or
- Using processor or core reduced power states when a core or socket is not being used.

As with all computer servers covered under this specification, partners must disclose all power management techniques enabled using a specific power and performance data sheet. Dual-Node servers with three or four sockets per node must also meet this requirement.

Power management for hard disks in the future is less relevant at the level of servers but will more effectively be addressed for storage units (SAN). Power management for memory is
Energy efficient servers in Europe

currently a field of extensive investigation but not yet addressed in practise.

7.4.3 Power management at server level

Virtualisation techniques offer new options for power management at the server level which are briefly described in the following section. The examples shown here are based on the features of VM-ware ESX 3.5.

When running a cluster consisting of at least two servers, resource scheduling and an load distribution can be administered beyond the limits of a single hardware resource. For these purposes logical resource pools can be defined. Using Distributed Resource Scheduler the workload of these resource pools will be constantly monitored and shared amongst all virtual machines of the cluster regarding to available resources and to predefined rule sets, reflecting the needs and priorities of a particular organisation. As soon as the workload of a virtualized system increases, the system automatically allocates the needed additional resources, eventually shifting the virtual machine to another server within the resource pool.

![Image](image.png)

**Fig. 7.1** This chart shows virtual machines being moved by Virtual Infrastructure DRS service in a cluster consisting of 3 physical servers.

Optimizing energy consumption of a resource pool can be achieved using the Distributed Power Management feature, which allows for consolidating workloads on a reduced set of servers at off-peak times and for bringing idle physical servers in standby mode. This new feature of VMWare Distributed Resource Scheduling, although experimental in ESX 3.5, provides recommendations, or makes decisions on turning physical servers on and off, regarding to resource claims of virtual machines, available server resources and predefined sets of rules providing spare server capacity for specialised applications e.g. web servers.

If overall resource demand including predefined spare capacity of all virtual machines in a cluster lies much below the available resources of all running physical servers depending on whether automatic or manual mode for Distributed Power Management is in use, servers are brought directly to standby mode or appropriate recommendations are given.

When resource needs of virtual machines in a cluster increase, DPM brings physical servers back online using the Wake On LAN feature of the servers, or gives appropriate recommendations, when automatic mode is not in use.
Decisions for bringing physical servers off- and online should be made according to load histories of at least 1 to 2 hours for standby and to 5 to 15 minutes for Wake on LAN, thus avoiding changing the number of active servers constantly due to just short term load changes. Some testing on the settings of these parameters and functions, as well as the Wake On LAN feature, which requires some special settings on the network infrastructure devices as well, is strongly recommended to meet the needs of the organisation in an appropriate way.

According to special high availability rules of an organisation additional affinity rules have to be defined, which provide for starting certain virtual machines on certain physical servers if they need some special hardware and for leaving at least two physical servers always online, to enable the VMWare High Availability feature to eventually start a virtual machines on the other server in case it crashes on the first physical server.

Cautious estimations show that energy savings of about EUR 200.- per year and virtualized system can be achieved when using the Distributed Power Management feature, going with additional savings for air conditioning. Furthermore considerable savings can be expected due to simplified administration e.g. patch management, rolling out of new services, testing, template management etc, as well as much less planned and unplanned downtimes of the systems.

In the future we can expect tools for calculating the effects of system changes, like storage upgrades, memory upgrades, additional physical servers etc and of significant workload changes on the overall energy consumption of the system, including the additional energy needed for cooling.

Recommendations for practise

The following aspects regarding power management should be considered in the procurement of new hardware:

- Information on power management features at the hardware component level should always be required as a part of the hardware documentation from manufacturers. Unless there are specific technical reasons it should be required that power management at component level (CPUs) is implemented properly (regarding BIOS, operating system and hardware component) and can be activated easily if desired.

- For volume servers with 3 and 4 CPUs the Energy Star label indicates that power management of equipment at CPU level is implemented and activated by default.

- Implementation if power management on virtualization level (VMWare Virtual Infrastructure, Distributed Resource Management, DPM feature) can be taken into consideration when using at least 3 physical servers in a virtual clustered environment due to the fact, that powering down one physical server leaves the cluster in a state, where the other features like VMWare High availability are still functional and all necessary services for bringing the 3rd server back online are running in a secured environment. As the necessary VMWare Distributed Power Management services depend on VMWare Virtual Infrastructure and require at least 3 physical servers plus storage system for reasonable operation of the cluster this configuration is suggested only for larger enterprises, which use at least 30 virtualized systems or more.

Due to fact that power management is still considered experimental on VMWare ESX3.5 server and requires some testing to find reasonable settings and due to pricing considerations previously mentioned, power man-
Energy efficient servers in Europe

...agement on virtualization level will presumably used barely by small- and medium enterprises.
8 Evaluation of leasing and guaranteed saving concepts

8.1 Guaranteed saving concepts

Within this work package the relevance, the possibilities and limits of energy service guarantees were analysed. The approach involves an assessment of energy savings (guarantees) in the view of providers of server consolidation and benefits of energy saving (guarantees) on the side of potential customers. The key issue in this context is a way to count energy savings.

During the last 18 months Sun and IBM as IT vendors did a huge investment in the marketing of Green IT and the investigation of possible projects at customer side. The representatives of both companies in the project consortium had many discussions with customers about Green IT, energy efficiency and the motivation of customers for consolidation and virtualization projects. Theoretical approaches (energy savings calculations on paper) and practical projects like the best practice cases within this project have been conducted.

The findings about a possible guaranteed savings concept are disillusioning. Guaranteed savings concepts do exist in the IT landscape but do not play a role for energy efficiency now. First of all, firms don't have the detailed electricity consumption data they need to implement energy efficiency initiatives. The primary barrier to CIO's to invest in energy efficiency is that they rarely have the data on energy cost and in most cases even don't pay the electricity bill.

Consolidation and virtualization projects are mostly driven by maintenance issues and costs or changing demands for IT services with necessary high investment for new hardware and even much higher investment for the migration. Energy savings alone don't constitute a business case to start a consolidation or virtualization, to overhaul an existing data centre or build a new data centre. As a practical example the best practice case with the City of Bad Soden am Taunus done by Sun may illustrate the problem. The consolidation and virtualization done by the city represents a typical environment and scenario being deployed in the SME business. With energy savings of more than 60% the cost savings for electricity of about 3,5 k€ per year would lead to a pay back time of more than 10 years. This would imply not to do this project just based on energy savings.

In fact Green IT and energy efficiency is very much an early stage market today; the demand
for Green IT is driven by a general perception of high energy costs and tightening regulations, but most by green ambitions and firms with data centres facing power restrictions.

Nevertheless it is necessary to incorporate best practice energy efficiency projects for IT into existing business plans for data centre improvements. Pure energy efficiency projects in most cases don't get past the CIO because of the very long pay back period. But energy costs may influence a pay back towards a positive decision for a project (which than leads to energy savings).

Faced with zero actionable information, guaranteed savings concepts would need to start with energy efficiency assessments and integrated monitoring of energy consumption across IT assets, data centre hardware and facility services. The costs for such engagement will easily exceed possible savings from energy efficiency very quickly – today! With emerging management and monitoring software for energy consumption in the data centre the pre-investments will be reduced in the next years. With infrastructure able to provide the energy usage information with little effort guaranteed energy savings concepts might evolve and be valuable to customers. But still: energy currently is too cheap in comparison to other costs of IT and IT projects.

Albeit guaranteed saving concepts seem not to be a valid approach for the time being. Currently a lot of other instruments to support energy efficiency in IT are in development on the market. There are different approaches to support projects on energy efficiency and to support the procurement of energy efficient servers:

- a rebate scheme of up to 1.000$ by PG&E in California for buying the energy efficient Sun server T2000
- CO2 certificates or rebates for energy costs coupled to energy efficient servers
- the German Kreditanstalt für Wiederaufbau (KfW) does support SME business with rebates and low interest loans when energy savings of at least 20% are reached by an investment
- the investment program of the German government for stimulation of the economy during the finance crisis links aid money for IT project to savings of energy of at least 25%
- the German Technical Supervisory Associations (TÜV) are in the process of developing “Green IT” certifications for data centres which in turn can be used by a company to position themselves better in the market.

8.2 Measurements of energy savings in the data centre

Still, the key issue in this context is a methodology to count energy savings for IT and especially in the data centre. Here we now find new programs and approaches to support data centers and IT managers in making energy efficient decisions for their server infrastructure. The Data Centre Specialist Group (DCSG) Carbon Trust Project (http://dcsg.bcs.org//content/view/45/59/) may serve as an example.

The initiative was started by the Market Transformation Programme (UK) and was transferred to the Carbon Trust (also part of UK Gov. Department of Environment). Carbon Trust have linked up with the British Computer Society to provide an on-line tool to assist managers in data centres in making energy efficient decisions for their server infrastructure. The tool is being developed to address the IT industry’s need to manage growing power consumption and increased carbon emissions. The E-Server project has links in this and the results of the E-Server project have been used to influence the initiative.

“The British Computer Society (BCS) and the Carbon Trust have teamed up with Romonet to produce an energy & cost simulation tool for data centres. The project is co-funded by the Carbon Trust (Networks Initiative) and
Evaluation of leasing and guaranteed saving concepts

Romonet and is based on the BCS Data Centre Specialist Group’s mathematical energy model for data centres and ICT infrastructure developed over a period of two years by leading experts in the field.

The software tool is capable of understanding the complex interaction and dependences within a data centre environment and simulating how user load on IT applications and systems results in power consumption within the data centre. The tool is also capable of modelling these complex environments against time allowing the user adjust many different variables and simulate the environment before spending a penny or making any real business or technical change. The simulator results give a clear impact analysis over time in terms of what the data centre running costs and energy consumed would be. It is hoped that with the emergence of such a tool businesses will be able to make more informed decisions when it comes to the real running cost (fiscal and environmental) of running an IT platform or service with the ever increasing cost of power as well as power linear IT devices needing to be carefully factored into any business decision.

The aim of the project from a carbon emissions perspective is to collect data from the tools use (voluntarily and anonymously if desired) and feed this back to the the Carbon Trust as a gross yearly CO2 saving figure.

The software itself which is being developed by Romonet will be released in April 2009 under an open source license allowing businesses to access it for free via a dedicated website. Businesses will be able to register on the site to gain access to the tool either via its web-interface or as a Java module for download and integration into other existing tools. (http://dcsg.bcs.org/content/view/45/59/)

8.3 Development of leasing concepts for energy efficient server technologies

Based on growing demand for Green IT projects - even if driven more by green ambitions or infrastructure insufficiencies than by financial aspects of energy savings – “Green” leasing concepts exist on the market. Based on classical concepts they do offer so called Energy Savings Performance Contracts (ESPCs) to accomplish energy projects for facilities without upfront capital costs. This concepts will be described here.

Nevertheless such concepts have not been carried forward to server technologies up to now. Discussions of Sun with different utilities and ESPC providers did not show significant interest of target groups and did not lead to further concrete activities. The situation for leasing concepts for energy efficient server technologies will be discussed in one of the following chapters.

8.3.1 Energy Contracting for Energy Savings

For energy efficiency projects, customers can turn to energy service companies (ESCOs) for Energy Savings Performance Contracts (ESPCs). ESPCs are contract vehicles that allow customers to implement energy projects without up-front funding.

The savings generated from the improvements pay for the project over a term of up to 25 years. Generally the duration of contracts exceeds 8 years with the longest periods being in the field of plant contracting with tendencies of up to 15 years of contract terms.
Energy efficient servers in Europe

Energy contracting alters the original service of energy supply into a more complex activity of optimising energy utilisation. The functional need changes for the client, as he does not simply purchase energy or raw materials but demands a form of energy use at minimum costs. The form of supply is left to the hands of an external player, the contractor.

In order to offer a specific form of energy use or saving measures, the product supplied by a contractor is not simply energy on its own but is comprised of all necessary components for a modernisation or redesign of energy systems. Planning, finance, technical realisation, and operation if demanded are offered for this purpose. From the client's perspective the take-over of such energy services by the contractor constitutes an outsourcing which can even reach as far as the outsourcing of the entire energy area.

Energy service companies (ESCOs) have completed projects in: building automation and energy management control; renewable energy systems, distributed power generation systems, energy/utility distribution systems, and energy-related process improvement. But for example for the German market, ESP contracting contributes only 10% to the overall contracting market ([Hirschl2000]).

8.3.2 Energy Contracting

The study on “Creating eco-efficient producer services” funded by the EU gives a comprehensive description of energy contracting (see [Hirschl2000]).

“The concept of energy contracting comprises different forms of energy services with main emphasis being energy supply by use of energy saving procedures and efficient technologies. With this focus given energy contracting can generally be described as an eco-efficient service, since the realisation of energy saving potentials also implies the saving of resources by using less primary energy sources and a decrease in emissions.

The main focus is on a contractual relation between an energy provider (contractor) and its customer (energy consumer). VDEW, the German Electricity Association, defines this service in the following way: "the term 'contracting' is usually used for projects, in which the goal of exploiting economically viable potentials of rational energy
consumption is pursued by forming individual contractual relations. Therefore contracting must be understood as a solution which is tailored on the specific energy needs of the customer” (VDEW97, p. 7).

8.3.3 Description of the Service “Energy-Contracting"

According to the above definition, energy contracting stands for a contractual relation between an energy service company and a client which is mainly aimed at the reduction of the client’s energy costs. Despite the main goal being set a single method or technical solution for achieving this goal is not predetermined. It can occur, however, that the client wishes specific technical solutions in the context of contracting arrangements. Or the service company may prefer to offer solutions which are within its existing portfolio of products. The latter case for example refers to the fact that some companies already have existing technical solutions or as in the case of large energy utilities already possess the necessary infrastructure such as long-distance heating networks, which largely influences the kind of technical solution chosen for the particular purpose.

The object of a contracting service can entail the whole range of energy consumption possibilities (heat, coldness, light, steam, compressed air, ventilation, motive power), that are no longer seen as isolated but are optimised as a whole. Different contractual models can have various degrees of complexity of the services offered. Different components can be consulting, planning, finance and the technical performance of the measures. Beyond, the operation, servicing and maintenance often play an important role. The degree of co-operation can range from just using single components to complete outsourcing, i.e. the external management of the whole area of energy supply.

8.3.4 Basic Contracting Models

With regard to the kind of energy contracting preferred, various models have been developed on the market over time. Two basic core models ARE:

- Performance-Contracting
- Plant- / Operation-Contracting

**Performance Contracting**

Within this realm the contractor takes measures intended to reduce the energy need and to enhance the efficiency of the customer's plants. It therefore consists of modernisation and optimisation methods, which are limited in scope when compared to plant-contracting. The measures are usually planned, financed and conducted by the contractor. The refinancing normally results from the guaranteed saving of energy costs. Performance-Contracting can be applied to fields such as lighting, plant and system controls or pumps. In the sphere of plant-contracting the redevelopment or even the new construction of an energy supply plant is prominent.
Fig. 8.2 Cash-Flow in contracting concepts

*Operation Contracting*

The contractor plans, finances and constructs the technical plant designed for this purpose, and operates and services it when needed. In the simplest case it can be refinanced by an agreed basis price, which takes capital input as well as variable costs into account; the basis price however can also be complemented by a kilo watt hour price taking explicitly the variable costs of the energy supplied into account. In this case, the agreed price over the course of the contractual arrangement is generally below or equal to the energy costs the customer had to pay before the measures were taken. With regard to so-called Operation-Contracting the plant becomes the property of the customer when the contractual relationship ends, so that he can completely internalise the energy cost savings."

With an Energy Saving Performance Contract, the utility or energy service provider typically arranges financing to cover the capital costs of the project. Then the utility is repaid over the contract term from the cost savings generated by the energy efficiency measures. With this arrangement, customers can implement energy improvements with no initial capital investment.

Fig.8.3  Energy and operational costs
ESPCs require a strict process and a comprehensive plan specifically catered to the strategic and technical business goals of the customer – while promoting energy conservation and addressing environmental concerns.

This involves:

- Performing a preliminary analysis to determine the organization’s current energy usage and identify areas to maximize energy savings.
- Producing a detailed energy analysis to determine the collection of improvement measures that will best satisfy the investment criteria and make the most positive impact on the facility.
- Delivering and installing new equipment and implementing the facility improvement measures.
- Conducting regular measurement and verification to ensure the energy savings are maximised.

<table>
<thead>
<tr>
<th>Building Owner / Operator</th>
</tr>
</thead>
<tbody>
<tr>
<td>Letter of understanding</td>
</tr>
<tr>
<td>Letter of intent</td>
</tr>
<tr>
<td>Contract closure</td>
</tr>
<tr>
<td>Change of usage, consumption, accounting</td>
</tr>
</tbody>
</table>

**Preliminary analysis**

- Preliminary study

**Detailed analysis**

- Detailed study
- Planning, installation, project management
- Energy saving guarantee measurement, verification, service

**Implementation**

**Guarantee phase**

Fig. 8.4 Process for energy contracting projects

### 8.3.5 Energy contracting for energy efficient servers

The concept of contracting is based on the saving of energy costs. Usually this goes hand in hand with a reduction of energy consumption and therefore a saving of primary energy sources. The scope of a contracting service largely depends on the specific needs of the customer and the technical possibilities available for the project.

A vital condition for contracting is a sufficiently high economic potential which can be exploited by modernisation or redevelopment of the electric power supply. Hence the (permanent) success of contracting is closely connected with innovations and developments of energy efficient technologies (see [Hirschl2000]).

Energy efficient servers potentially are a possible new focus area for ESPCs. The existing server infrastructure in the data centres often still originates from the Y2K and internet bubble period, i.e. the systems are 6-10 years old.

Such an infrastructure allows for big savings opportunities, when utilizing IT techniques like consolidation, virtualization and technology refresh, as it was shown in the best practice cases of the E-Server project.

Nevertheless there do not exist leasing or contracting offerings for IT hardware and energy efficient servers on the market. And the development of such concepts is not in the focus of energy providers. This is for different reasons, as there are general business barriers for energy contracting and specific
barriers regarding energy contracting for energy efficient servers.

### 8.3.6 General business barriers

Essential barriers to energy contracting predominantly arise from the complexity of the concept and lack of information, possible radical changes of organisational structures for instance due to new forms of co-operation and different required qualifications on the supply-side. From the Demand-Side perspective the lack of a favourable legislative framework and the outsourcing of activities might cause restraints.

The complexity of energy contracting is first of all stemming from its variety of concepts, models, and expressions. Even those involved regard the blurriness of expressions as a restraint to the diffusion of this service innovation. Beyond that, complexity arises from grouping different service components such as consulting, planning, finance, construction, operation, servicing etc., which can lead to qualificational or organisational conflicts on both the supply and the demand side.

The complexity of contracting implies that acquisition efforts and consulting services are regarded as disproportionately high by those involved. Therefore the large cost factor is supplemented by a high degree of uncertainty. Nevertheless, the high risk of investment due to the pre-financing of contracting measures as another problem.

From the customer's perspective finance also constitutes a problematic affair because especially in the industrial sector contract duration is usually longer than the amortization period of comparable investments.

Contracting overall suffers from its limited dissemination in the market, the high complexity of contracting arrangements, and also the associated acquisition efforts in an increasingly competitive business environment. From the demand-side perspective, the long duration of contract terms constitutes a restraint which is particularly due to the dynamic developments in the energy sector at the moment. Besides, there is sometimes traditional scepticism with regard to outsourcing of internal affairs at conservatively managed companies.

### 8.3.7 Business barriers for energy contracting for energy efficient servers

The general business barriers apply also for the field of IT and energy efficient servers:

- complexity of energy contracting
- grouping different service components such as consulting, planning, finance, IT management, business drivers
- acquisition efforts and consulting services are disproportionately high
- contract duration is usually longer (8+ years) than the normal amortisation period for IT hardware (3-5 years).

The analysis of the best practice cases from the E-Server project even shows: Energy savings alone in most cases don't constitute a business case to start a consolidation or virtualization, to overhaul an existing data centre or build a new Data Centre.

As a practical example the best practice case with the City of Bad Soden am Taunus done by Sun again illustrates the problem. The consolidation and virtualization done by the city represents a typical environment and scenario being deployed in the SME business. With savings of more than 60% of energy the costs the pay back time for the investments is 10 years. From an economic point of view this pay back time would imply not to do this project (just based on energy savings). The same observation can be done with other best practise cases for the BMU and Salvation Army done by IBM.

Another barrier for leasing and contracting concepts for energy efficient servers is the problem of measurements. As stated in the chapter above the methodology and practice to
measure energy and to count energy savings for IT and especially in the data centre is a key issue in this context which is still not resolved today.

Sun had some presentations and talks to different utilities and ESPC providers also in the US during the last months about contracting concepts for IT and did not find much interest. No concrete activities could be implemented. This may be due to the fact that utilities have little knowledge about IT but also due to the low contractual value of such deals (most consolidation projects do have IT investments below 1 Million EUR) which may not justify for an engagement.

Considering the low level of energy prices there is another economic restraint for energy services and efficient technologies. Many experts regard the currently low energy prices for large enterprises as dumping-prices mainly set by large energy utilities on a marginal cost basis often leading to the effect that new investment and services seem uneconomic.

Generally speaking, a rise in energy prices would have a positive impact on the diffusion of contracting services for facilities and IT because the demand for efficient saving technologies as well as professional service providers will then be likely to increase.
9 Conclusions and Outlook

9.1 International status of activities and market impact

The E-Server project has calculated the energy consumption of servers and data centres for EU27 and has also estimated trends for energy demand in the near future. Energy consumption in data centres is expected to double between 2007 and 2012 from 40TWh to 80TWh if business as usual is continued. On the other hand enormous energy saving potentials of at least 60% could be exploited by means of broad application and optimized management of efficient IT hardware and infrastructure. However this requires broad concerted action at international level for the development and implementation of effective instruments supporting energy efficient market development.

The IEE E-Server project significantly contributed to a first stage of initiatives supporting sustainable market development for central IT equipment. Among other impact the project supported the development of first energy efficiency criteria for servers, demonstrated the economic use of efficient technology in best practise and provided and disseminated guidelines for the procurement and management of efficient technology.

Despite the little attention energy efficiency of servers and data centres has received until 2007 several initiatives have been implemented in the meantime at EU and international level in parallel to the IEE E-Server project. Among others The Green Grid, the Energy Star programme and the Code of conduct for energy efficiency have provided first instruments to support the market development for energy efficient central IT-technology and infrastructure. The focus of these initiatives so far has been on the development of first energy management schemes for data centres, the development of first criteria for servers as well as on management and procurement guidelines.

Overall this pioneering work helped to take a first step in the implementation of energy efficient central IT technology. However extensive further concerted action is needed to take the second step from first best practise and supportive instruments to a broad implementation of efficient technology in the market.

So far activities have been focussed on part of the important technologies only and the market supporting instruments already developed could not be broadly implemented yet. The yet still limited market impact is not surprising as initiatives have only been started two years ago and thus are extremely young compared to action for other standard technologies (e.g. heating, lighting etc.).

Further comprehensive measures are essential to exploit the huge energy efficiency potentials and especially to avoid an enormous further
increase of energy consumption within the next few years. The energy demand of Central IT equipment still will increase significantly due to the fastly growing IT services in the private and public sector and increasing hardware performance and capacity. Thus the ecological impact of these technologies in terms of energy consumption will be increasing but should be kept as low as possible by means of preventive measures.

The focus of future work needs to address both technological areas not covered so far and an effective implementation of market supporting instruments.

9.2 Concerted action proposed for the next stage of sustainable market development

9.2.1 Coverage of the major types of central IT hardware and infrastructure equipment.

As indicated above the previous actions only could cover part of the relevant technology and also part of the preparation and implementation of market supporting measures. At the technological level action has been focussed on consolidation and virtualisation of servers as well as on infrastructure issues in large data centres. The following aspects have hardly been addressed so far and need specific attention:

- Energy efficiency of storage and network equipment
- Energy efficiency of cooling solutions for small to medium size companies
- Power management for servers and in particular also virtualised server systems

These areas offer significant additional potential for energy savings. Volume and transfer of data is still increasing rapidly due to new IT processes and services in the public and private service sector. A large part of the data currently is stored inefficiently in terms of energy demand and costs. High data transfer and permanent access to data also increase the impact of network equipment.

At the level of power management technology energy saving options are available for servers as well as for other types of equipment but so far are hardly used in practise. Virtualisation of servers offers new options for flexible power management including shut down of servers during periods of low activity.

Energy efficient cooling concepts so far mainly have been focussed on large data centres, efficient solutions for small to medium size IT systems still need to be demonstrated and implemented.

For all these types of equipment energy efficient technology still has to be promoted and implemented and specific power saving concepts should be exploited.

9.2.2 Further development and broad implementation of supporting instruments for energy efficient market development

Concerning the development and implementation of instruments to support the development of the market for energy efficient central IT equipment several gaps have to be closed and tools available to date need further refinement and broad implementation.

Energy efficiency declaration and labelling

A major prerequisite for all other market supporting instruments is the transparency of energy efficiency of the relevant IT equipment. Energy efficiency of the equipment needs to be transparent to be considered as a criterion in procurement and management. Thus an internationally standardised declaration of the energy consumption of the equipment is indispensable.

During 2007 and 2008 a first step was taken in the development of energy efficiency criteria for servers. Preliminary criteria for energy...
consumption and first benchmarks have been developed in the Energy Star programme (EPA2009) and by SPEC (SPECpower2008). However these criteria cover only part of the important products and technologies. For the future a more comprehensive set of criteria is needed also covering larger servers (>2 CPUs) and addressing power consumption at different load levels and for different workloads. Further work on these issues already has been announced by both SPEC and Energy Star.

Besides hardware based criteria also criteria addressing energy consumption in relation to service output and specific applications should be addressed in the future.

Apart from servers energy efficiency criteria and declaration are also needed for storage and network equipment as well as for cooling equipment for SMEs. A development of energy efficiency requirements for storage equipment has be announced for the Energy Star programme already.

**Mandatory requirements for equipment**

Apart from the largely voluntary Energy Star labelling scheme options to support the market with some mandatory requirements also should be evaluated. For this purpose the planned consideration of this type of equipment in the preparatory studies for the EUP directive seems appropriate.

While for several features of central IT equipment declaration or labelling seems the preferable approach, for some aspects like power supply efficiency, idle consumption and power management also mandatory requirements are feasible and may be appropriate.

**Education**

The previous pioneering initiatives mentioned above including the IEE E-Server project have already provided a number of useful tools and guidelines. However since action to support energy efficient market development has been taken only recently dissemination still has to be further pushed. Many IT and infrastructure managers as well as consultants today are simply not aware of the options for energy savings which can be achieved without negatively affecting the central IT specific goals. One of the approaches to transfer information and guidelines to this target group is education. Education and training for IT hardware and infrastructure managers as well as IT consultants can effectively transfer expertise on selecting, implementing and maintaining energy efficient solutions to the demand side target groups.

Best practise and guidelines for example from the E-Server project as well as from other initiatives may serve as valuable source for such education/training action.

**Public procurement**

Public procurement can be an effective driver of market development for energy efficient technology. In this sense public procurement is also often highlighted in EU policies and programmes as a major instrument to stimulate the market and to demonstrate best practise. While energy efficiency criteria for client side office equipment have been part of public procurement for a long time already (printers, monitors etc.) criteria for central IT equipment have not been implemented so far due to unavailability. However based on the new energy efficiency criteria currently becoming available, public procurement should also be used as a major driver for energy efficiency in this area.

**Certification of data centres**

For larger data centres certification schemes provide an incentive to implement energy management for central IT hardware and infrastructure. A few first certification concepts have recently been developed at national and international level. Examples are a concept developed by TÜV and the Code of conduct for energy efficiency of data centres developed by JRC ISPRA in co-operation with international experts. The concepts available today are not broadly accepted and used in the market. They partly lack a strong implementation and marketing platform.
It would be most effective to develop a standardised internationally accepted scheme which is based on a solid management platform and offers participants clear advantage also regarding external communication and demonstration of superior quality and efficiency of IT processes. Some of the already developed approaches would have the potential to be used as a basis for the development of a generally accepted and applied scheme. In this context also an implementation at the level of EN ISO should be evaluated as a possible option.
10 References


[DCSG1] Carbon Trust Project, http://dcsg.bcs.org/content/view/45/59/


References

[EL3500] User Manual Energy Logger 3500


[ES08] Energy Star, Draft 1 ENERGY STAR Computer Server Specification


Energy efficient servers in Europe


[Hu04] Energie Schweiz, Erfahrungen in der Anwendung von ACPI bei Windows-Servern, Schlussbericht im Auftrag des Bundesamtes für Energie, July 2004


[IBM07] End to end power, link to be published


[IBMb] Virtualization on the IBM System x3950 Server

[IBMc] Service Processors Supported in IBM System x Servers

[IBMf] IBM System p5 510 and 510Q Technical Overview and Introduction


[IDC05] WE installed base 2001-2005

[IDC07] EMEA Quarterly Server Forecast, Q1 2007


[IDC07c] IDC WE installed base 2001-2005


References


[LBL07], LBL, Enabling High Efficient Power Supplies for Servers http://hightech.lbl.gov/documents/PS/Intel_Enableing.ppt

[LBNL02] Lawrence Berkeley National Laboratory, Berkeley (CA), USA., Data Center Energy Benchmarking, Case Study, 2002


Energy efficient servers in Europe


[SKO+97] H. Sanchez, B. Kuttanna, T. Olson, M. Alexander, G. Gerosa, R. Philip, and


References

[SP08b] Performance Evaluation Corporation, SPECpower_ssj2008 Published Results

[SSI07] Power Supply Management Interface Design Guide, 
http://ssiforum.org/Power%20Supplies/Forms/AllItems.aspx


[VDEW97] Vereinigung Deutscher Elektrizitätswerke e.V. 1997: Contracting Fallbeispielsammlung; VDEW Publikation; VWEW-Verlag, Frankfurt a.M.

[VM06] VMware Inc. ESX Server Data Sheet, 2006


Energy efficient servers in Europe