Product-Neutral Performance Description x86 Servers

Guide

Version 1.1

BITKOM

BESCHAFFUNGSAMT
des Bundesministeriums des Innern

Imprint

Publisher:	BITKOM Federal Association of Information Technology, Telecommunications and New Media. Albrechtstraße 10 A 10117 Berlin-Mitte Tel.: 030.27576-0 Fax: 030.27576-400 bitkom@bitkom.org www.bitkom.org
	Procurement Office of the Federal Ministry of the Interior Sankt Augustiner Str. 86 53225 Bonn, Germany Phone: 0228.99610-0 Fax: 0228.9910610-0 itk-beschaffung@bescha.bund.de www.beschaffungsamt.de
Contact person:	Oliver Lowin, Tel.: 030.27576-133, o.lowin@bitkom.org Michael Unger, Tel.:0228.99610-2900, michael.unger@bescha.bund.de Carsten Kolbe, Tel.: 0228.99610-2930, carsten.kolbe@bescha.bund.de
Current version:	www.itk-beschaffung.de
Editorial assistant:	Juliane Kukla
Design/layout:	Design Bureau kokliko/Anna Müller-Rosenberger, Juliane Kukla (BITKOM)
Copyright:	BITKOM/Procurement Office of the Federal Ministry of the Interior 2010 Version 1.1 (date: March 2011)
Title page illustration:	© Fotolia.de

This publication constitutes general, non-binding information. Its contents reflect the view of BITKOM at the time of publication. Although this information was prepared with the greatest possible care, we make no claim to factual accuracy, completeness and/or timeliness; in particular, this publication cannot take the specific circumstances of each individual case into account. Its use is therefore the sole responsibility of the reader. All liability is excluded. All rights, including the right to duplicate excerpts from this publication, are reserved by BITKOM.

Product-Neutral Performance Description x86 Servers

Guide

Version 1.1

BITKOM

BESCHAFFUNGSAMT des Bundesministeriums des Innern

Table of contents

1 Introduction 3
2 Server classification 5
2.1 Server types 5
2.2 Weighting of the components depending on the server type 6
3 Sizing 8
4 Designs 9
4.1 Floor-standing systems 9
4.2 Rack systems 9
4.3 Blade systems 9
5 Processors and performance description 10
5.1 Server processors 10
5.2 Benchmark 10
6 Server memory 12
6.1 Hard disks 12
6.2 SSD 12
6.3 Controllers 13
7 Interfaces 14
8 Applications 15
9 BIOS, drivers, operating system 16
10 System management 17
11 Power supply units/power supply 18
12 Backup and restore or data backup 21
13 Nontechnical requirements 22
13.1 Choosing the right contract type 22
13.2 Overview and reference source of the EVB-IT 23
13.3 Establishment of the power connections 23
13.4 Support 23
13.5 Logistics 24
14 Valuation of offers 25
14.1 Measurement protocols 25
14.2 Evaluation process 25
GLOSSARY 26

1 Introduction

This guide is the result of a working group under the leadership of the Procurement Office of the Federal Ministry of the Interior and the Federal Association for Information Technology, Telecommunications and New Media (BITKOM). The purpose of this document is to provide the contracting authorities (at the federal, state, and local levels) with reliable and comprehensive assistance in formulating their tenders for the procurement of servers in a product-neutral manner, i.e. without the use of protected brand names or references to specific manufacturers, and with due consideration of the current technical requirements.

European and German law generally prohibits the use of brand names in the context of public tenders. This is a consequence of European law prohibiting discrimination within the framework of directive 93/36/EEC of 14 June 1993, implemented in § 8 VOL/A, which aims to ensure that certain manufacturers or suppliers are not excluded from the potential bidders due to discriminatory wording. An exception is allowed only in cases in which the description of the performance with sufficiently precise and generally intelligible words is not possible. In any case, the additional words "or equivalent" must always be added.

Precisely in the field of IT systems procurement, this is of course not an easy task. Describing the technical complexity of the matter, the rapid succession of product cycles, and especially the difficulty of precisely describing the desired performance of a system while including all technical requirements precisely, poses major challenges to public procurement agents. This explains why, precisely here, the performance description is/was often performed using the relevant product names. It is precisely in these areas that this guide is applicable, because it constitutes a concise tool to support the observance of the legal requirements and thus the assurance of fair competition. In addition, this guide also names and describes the current technical standards.

We consider x86-compatible, smaller and medium-sized individual systems (floor-standing systems, racks, and blades up to max. 2 sockets), which are used primarily as group or departmental servers – whereby such servers are generally designated "workgroup servers". Additional platforms with other processors, such as RISC, mainframe, embedded, vector or matrix processors, as well as cluster systems, are not part of this guide, since more stringent operating requirements must be considered in these cases, which must be defined in cooperation with a consulting company or a technical department.

This guide does not give any consideration to environmental aspects. A separate guide is being developed in this regard.

To ensure that the guide is always kept up to date, it will be updated at regular intervals. Here, new technical developments are considered and the suggested benchmark values are adapted to the current technology. The latest version of this guide can be found at www.itk-beschaffung.de.

This document could only be created with the intensive collaboration of the participants in the "product-neutral server tender" project group. Particular thanks are due to:

- Wolfgang Dorst, Oracle Deutschland B.V. & Co. KG
- Peter Dümig, DELL GmbH
- Marco Junk, BITKOM
- Carsten Kolbe, Procurement Office of the BMI
- Stefan Kreger, Fujitsu Technology Solutions GmbH
- Volker Kwasnicki, IBM Deutschland GmbH
- Uli Norf, Intel GmbH
- Oliver Schubert, Hewlett-Packard GmbH
- Eric Stolle, AMD
- Jörg Walther, Intel GmbH
- Bernhard Wolz, Federal Employment Agency
- Gerold Wurthmann, Intel GmbH
- Thomas Zapala, Federal Office for Information Management and Information Technology of the Bundeswehr

2 Server classification

2.1 Server types

The x86 servers used by companies and government are primarily used for the following applications:

File/print servers

This server type is used for central storage of user data and for central control of group or department printers.

Mail servers

Mail servers are the central server instances with which the electronic mail traffic is controlled. Commercial mail server programs thus allow for configuration freedom regarding whether mails will also be stored on the workplace computers of the users, or whether the mails will be stored centrally on the mail server. Regardless of this configuration, it may be necessary to design a greater hard disk capacity. If, for legal reasons, it is necessary to store or archive e-mails for a longer period of time, then the advice of a professional consulting firm should be used, so as to include the legally necessary prerequisites in the tender.

Virtualization servers

These servers are used if different, generally older servers are replaced by one or fewer new systems, thus mapping the old systems to the new hardware platform in the form of "virtual machines". This allows existing applications to continue to run under their own operating system (or different versions of an operating system). Here, it is recommended to use a consulting firm that supports the virtualization server in the configuration.

Web servers

These servers are used to prepare HTML pages for an Internet presence or even for internal Web pages.

Database servers

These servers use various different database programs for the storage and processing of many different types of data.

Application servers

This server type has been included for the sake of completeness. Certain applications use this server type for the calculation of data, whereby the actual data is read from another database server. As an example here, we discuss the application server of the SAP standard software. Whereas the application data is submitted externally via a Web server, the application server performs the actual data processing and the actual data is in the background on a database server (3-tier architecture). Such a multilayer architecture is only useful for larger applications (for example, for financial authorities or in SAP installations).

Terminal servers

In a terminal server, the application is run on a central computer; the graphical output is provided via a network connection. Several workstation computers can access the terminal server in parallel and use the application simultaneously. The data created is stored centrally.

2.2 Weighting of the components depending on the server type

Depending on the applications, different preferences must be observed in the hardware. The following matrix shows in general terms what aspect should be given particular value in the hardware used. Definitions:

o == of subordinate importance
+ == of low importance
++ == of medium importance
+++ == of high importance

Example:

	CPU	RAM	Hard disks	Network
File/print servers	o	+	++	+
Mail servers	+	+	++	o
Virtualization servers	++	+++	++	++
Web servers	+	+	o	+
Database servers	++	++	+++	+
Application servers	++	++	o	+
Terminal servers	++	++	+	+

If a mail server is to be procured, then for the tender attention should be given to ensuring that the hard disk subsystem of the server used is large enough to store the e-mails of all end users. It must then also be checked if external equipment is useful for the hard disk subsystem. In addition, attention should be given to ensuring that the hard disk subsystem is as fail-safe as possible so as to be secured against data loss.

In so doing, the aspect of data security must be considered at all times, because for a mail server the shutdown of the system with subsequent data loss must be ruled out under all circumstances.

If there is any lack of clarity regarding which parameters are specifically important for the actual case, a sizing should be performed by an independent body to obtain an estimate of the configuration of the server to be procured.

Note:

When determining the hardware, it should be noted that this results in dependencies relating to the software licensing conditions. In some circumstances, the software license costs can clearly exceed the hardware costs (e.g. license costs per CPU core). See Chapter 8 of this guide for details.

3 Sizing

The most important question to be answered before a procurement is the question of the required server configuration depending on the concrete task, e.g. the number of users, the software used, etc. This is designated by the term “sizing” (sizing = layout of CPU/memory/HD/IO). Software and hardware manufacturers make concrete recommendations in this regard or make special planning tools available.

In the context of this document, due to the variety of possible parameters, no concrete suggestions can be made by means of which the procurement offers can determine the sizing of a server.

Example:

If an authority wants to procure a mail server, the following questions for example must be answered in advance:

- How many users should the system be able to serve?
- Should all e-mails be stored permanently?
- What average e-mail traffic is to be assumed?
- Are the users' mailboxes restricted in terms of their size?
- etc.

This and additional parameters can be determined in collaboration with the technical departments. If any lack of clarity exists regarding the sizing of the server to be tendered for, then a consulting firm specializing in this field should be consulted.

4 Designs

In the server market, three housing designs can currently be found:

4.1 Floor-standing systems

This equipment design is generally used outside of data centers or when rack integration is not possible.
A typical characteristic of floor-standing systems is their expendability in relation to hard disks and IO slots.
Floor-standing systems are also sometimes available in rack-mountable designs.

Floor-standing systems should be equipped with a lockable cover to prevent data theft and accidental switching off.

4.2 Rack systems

Rack systems are characterized by the fact that they can be integrated with a few height units (HUs) into 19" cabinets ("racks"). The internal expandability of the hard disks and IO slots may be limited due to the small design size.

Depending on the manufacturer, the cabinets are offered in different heights and with a different number of height units. To ensure trouble-free integration into the data processing landscapes, the integration dimensions of the cabinet should correspond to one of the following standards: EIA 310D, IEC 60297 or DIN 41949.

Cabinets for rack systems should basically be equipped with a lockable front and rear door and latchable side walls and covers.

4.3 Blade systems

Blade systems are an even more compact design type compared to rack systems. A blade server, server blade or blade for short is a component which, together with similar blades, is inserted into manufacturer-specific blade server chassis and uses the integrated infrastructure e.g. power supply units, fans, and backplanes together with the other built-in blade servers. There are various expansion options with IO modules (LAN and SAN), storage blades and management modules, which can be equipped according to the specific requirement.

The advantage of blades is their compact design, scalability and flexibility, as well as more simple cabling with low cable costs. Depending on the manufacturer, there are blade server chassis with different HUs and expansion options.

5 Processors and performance description

5.1 Server processors

The table of server types in Chapter 2.2 contains a column "CPU". This describes, with an evaluation from "o" to "+++", whether a server type requires a small or large amount of processing power.

- Server types such as file servers, which require only very little processing power, are given fewer CPU cores and lower clock frequencies.
- However, if they require high computing performance, for example in the case of virtualization or database servers, then more CPU cores and/or higher clock frequencies are needed.

For physical reasons, the technical development proceeds towards a higher number of CPU cores per processor instead of higher clock frequencies. Therefore, in the future, in addition to the currently available 2, 4, 6, 8 and 12-core processors, those with even more CPU cores are anticipated.

A server processor is a processor specially designed for the needs of a server, and differs in many respects from processors used in desktop PCs or notebooks. These differences include durability, reliability, performance and specific features.

For the performance of servers and server processors, various factors are decisive:

- Number of (processor) sockets
- Number of CPU cores
- Clock frequency
- Memory connection
- Size of the cache memory

5.2 Benchmark

Since public tenders must be formulated in a product-neutral way, i.e. without naming specific manufacturers or referring to specific products from a manufacturer, a server tender must also contain no specifications of processor manufacturers or references to specific models of a processor.

For servers, a variety of application fields exist, the most common of which are summarized in Chapter 2, Server Classification, with their associated hardware equipment preferences. To describe the performance of servers, application benchmarks must be used. The benchmarks known from the desktop PC or notebook fields (e.g. BAPCo Sysmark) are not well-suited to servers.

The advantage of benchmarks resides in the fact that they offer a specific, comparable and reproducible method for the performance measurement of a system. However, almost all application cases for servers differ to such an extent that standardized benchmarks can only very seldom produce reliable statements about the performance of a server in concrete application cases. Therefore, prior to procurement, specific, i.e. situation-related application benchmarks should be performed.

This may be problematic in the case of small procurement volumes and individual procurements for procurement officers and for suppliers, since it is associated with significant costs. Smaller companies might also not be in a position to run the benchmarks because the human and financial resources of such companies are generally very limited.

In that case, they could access standardized and generally recognized benchmarks for the respective application type. These application benchmarks are based on stringent test methods that are developed either by independent industry consortia or software manufacturers, and are recognized and supported by the server manufacturers. An application benchmark is a program or a number of programs that measure the overall performance of a system. The best-known consortia in the benchmarks field are the SPEC (Standard Performance Evaluation Cooperation) and the TPC (Transaction Processing Performance Council). Other benchmarks are defined by software manufacturers.

For all published benchmarks, the user must ensure that they have only been performed for specific configurations that do not always necessarily match the server to be procured. If benchmark values are used to compare systems, then it is therefore always recommended to take a look at the so-called "full disclaimer" on the website of the benchmark provider. The respective precise configuration of the measured system is developed from this. In addition, benchmarks often simulate a specific application constellation and are also independent of the operating system and application software used. Therefore, the results cannot necessarily always be generalized.

Benchmarks known from the desktop PC or notebook field (e.g. BAPCo Sysmark) are not suited to server systems. However, to ensure product neutrality and nevertheless a certain comparability of the offers even in an x86 server tender, it is recommended to have a neutral performance classification in the form of a benchmark. A recommendation for server procurement is made by the SPEC INT RATE BASE; however, it only shows the performance of the system with regard to the CPU/main memory. The data can be found on the website of SPEC under
http://www.spec.org/cpu2006/results/rint2006.html.

However, it must be noted that although the SPEC INT RATE BASE is a widely used benchmark for evaluating the performance of servers, due to the different server types and the variety of different use environments named in Chapter 2.2, it cannot deliver any reliable statements regarding the overall performance of the system in these different environments.

For all published benchmarks, the user must ensure that they have only been performed for specific configurations that do not always necessarily match the server to be procured.
Therefore, in the context of this guide, no concrete recommendations can be given with regard to benchmark values.

Due to the associated costs, the authors of this document recommend orienting yourself towards published benchmark values in the procurement of servers, and refraining from verifying benchmarks such as SPEC or TPC yourself.

Additional benchmarks are explained in the glossary.

6 Server memory

In principle, memory media specially released for servers should be used exclusively.

Technologies:
Today and in the foreseeable future, 2 memory technologies are used. They and their characteristics as well as their advantages and drawbacks are briefly described here, so that you can arrive at the optimal choice in procurement. The two technologies are called hard disks and solid state drives (SSDs).

6.1 Hard disks

The following general features apply to hard disks:

- mature technology, long-term market presence
- susceptible to strong shocks in operation

Hard disks exist with S-ATA and SAS interfaces. The SAS interface is specially developed for servers and permits a higher throughput as well as a shorter access time than the S-ATA interface. Therefore, SAS hard disks are used everywhere where high performance (I/O per second) is required. SAS hard disks also offer higher fail-safety. S-ATA hard disks on the other hand offer lower performance and fail-safety and are primarily suitable for sequential data access (e.g. image processing and archived data).

For server hard disks, there are various criteria that must be differentiated:

- Type of hard disk: S-ATA or SAS
- Revolutions/minute (RPM) of the hard disk: 7,200 (S-ATA) or 10,000 (SAS) or 15,000 (SAS)
- Data transfer rate: 1, 5, 3 or 6 GBit/s

Capacity and form factor (2.5-inch or 3.5-inch) are in constant technological change in combination with the above criteria, which is why no recommendations are issued in this regard. The general trend is towards smaller form factors.

6.2 SSD

The SSD (solid-state disk) is a very new technology which is nevertheless rapidly increasing its market share. The advantages of SSDs compared to hard disks are:

- very fast access times
- very high I/O performance
- very low energy consumption
- no moving parts, and therefore relatively immune to shocks
- wider temperature range than hard disks

In SSDs, we differentiate between so-called MLCs (multilevel cells) and so-called SLCs (single-level cells).

Note:
Compared to conventional hard disks, SSDs are more expensive in terms of storage capacity and still at an early stage of technological development. A general recommendation for SSDs can therefore not be made, at least at this point in time. However, in comparison to hard disks, they offer a multiple of the I/O performance:

- SSD: Several 1,000 I/Os per second
- SAS: approx. 160-220 I/Os per second
- S-ATA: approx. 90-100 I/Os per second

A high I/O performance is advantageous for database systems, for example.

SSDs are still not yet available in the same high storage capacities as hard disks. Therefore, hard disks are still predominantly used in the server environment at this time. Their combination with SSDs is possible.

6.3 Controllers

In servers, a so-called RAID controller is often used. RAID (redundant arrays of inexpensive disks) is the combination of several hard disks/SSDs into one unit in order to achieve greater speed/security/availability.

In storage controllers, we must differentiate between directly connected storage media and access controllers for central storage systems.

For directly connectable storage media, the controller has the task of administering the RAID functionality, as described in the annex. These adapters often offer caches to accelerate access or to increase the write speed, which have a battery to protect against data loss in the event of power outages. As the interface, SAS or SATA are generally used.

A special role for the local storage controller is played by PCI adapters, which themselves provide SSD-type storage for very high I/O loads locally.

The connection to central storage is made via so-called hostbus adapters for the Fibre Channel or SAS interfaces, or even via high-performance network adapters for the iSCSI protocol.

Securing the data via a hardware RAID can alternatively also be performed via a software-based RAID. The software RAID functionality is generally a component of the operating system. The possible RAID functions are dependent on the software used.

Overview of the interface recommendation according to the server types:

Server type	Interface
File/print server	S-ATA/SAS
Mail server	SAS (SATA)
Virtualization server	SAS
Web server	S-ATA (SAS)
Database server	SAS (PCI)
Application server	SAS (SATA)
Terminal server	S-ATA (SAS)

7 Interfaces

Server systems use a variety of interfaces to communicate with the outside world. The different interfaces permit high flexibility in relation to the upgradeability and application areas of the systems. They are defined and standardized by recognized committees so as to ensure the correct operation of components from different manufacturers. An overview of the current interfaces in the market can be seen in the annex.

Since the requirement for interfaces depends on the concrete use scenario, in particular the integration into the existing IT landscape, no recommendation can be issued at this point.

8 Applications

Software – Dependencies for server selection

In addition to the influence of the various hardware components already named in Chapter 2.2 ("Weighting of the components according to the server type") on the respective tasks, the software aspect should also implicitly be kept in view with scalability and licensing.

In recent years, performance increases in servers have been achieved not through higher clock frequencies as they were previously, but often through a higher number of CPU cores working in parallel. On the one hand, this led to changed licensing models from many software publishers as well as to the need for precise consideration of the scalability of the applications to be used. With regard to scalability, we must answer the question of whether performance increases are achieved via multiple cores or exclusively via higher clock frequencies. With regard to the licensing models, it must be considered that, for many applications, the licensing costs are charged on the basis of physical sockets for specific CPU types (e.g. operating systems) or even based on existing cores in the entire system (e.g. databases or middleware).

Clarification with the technical department with regard to the requirement of the software and its licensing can contribute to considerable cost savings.

9 BIOS, drivers, operating system

BIOS

With the BIOS (Basic Input Output System), in the so-called POST (power-on self-test), all system components are checked for their correct functionality.

In the BIOS setup, system functions and hardware configurations of the system, such as security and energy saving functions, server management, boot sequence, etc., are set.

This information from the BIOS manufacturer should be able to be queried, since the setting options are dependent on the manufacturer's hardware configuration.

Drivers

- All system drivers for the most important operating system versions should be up to date and permit conflict-free operation. Driver updates for individual standard components should not lead to system conflicts.

- These updates should be available offline or online according to the built-in components.

Operating system

We will not go into detail here on the further functionalities of industry-standard server operating systems. The servers should basically be registered in the hardware compatibility list of the respective operating system manufacturer.

10 System management

Server management software is today an important constituent of the secure operation of a server. Server management tools are therefore a component of a server system. This allows monitoring of the servers in operation and/or independently of the operating state.

Administrators or service technicians are allowed access to the server and can perform comprehensive controls even at decentralized locations. Routine tasks and maintenance activities for server problems can thus be performed efficiently. Additional important functions include deployment and the ability to integrate larger installations in the data center in enterprise management systems.

They can be operated via a web-based user interface; for certain components, an additional command line interface may also be useful. This server management must ensure through appropriate access controls that unauthorized access to management functions is not possible.

The use of a server management system reduces administration costs and increases server availability. Service costs are also reduced by preventive fault detection and various integrated diagnostic functions.

11 Power supply units/power supply

General

Servers are in many ways exposed systems whose availability should be subject to a high level of attention. Therefore, in addition to fans, hard disks and I/O adapters, power supply units are also components that must often be redundantly designed in servers.

The second criterion after redundancy is also their hot-pluggability, which permits the replacement of faulty components without operational downtime.

For large and important systems, i.e. systems in central functions on which many users work in parallel, such as database, virtualization and file servers, redundant and hot-pluggable components are the de-facto standard. For small and cost-effective entry-level systems, which generally have only one CPU socket, manufacturers generally do not implement redundancy and hot-pluggability of power supply units and fans due to cost constraints.

Power supply units

Redundancy in floor-standing and rack servers means that generally two, and less frequently even more, power supply units are required for each system. In the case of blades, power supply is the task of a blade chassis. Here, power supply units generally supply considerably more servers, known as blades, so that in sum the systems can generally be operated more energy-efficiently. For blades, redundancy for the majority of use scenarios should be a prerequisite.

For redundant power supply units, it should be ensured that the systems offer separate power connections and thus redundancy can also be ensured with regard to the energy inputs. If all power supply units are supplied via only one common energy input, the failure of the latter invariably leads to operational downtime.

In practice, an uninterruptible power supply is often found upstream of at least one energy input, which protects servers against power outages, but also against recurrently occurring voltage fluctuations or spikes in the power network.

In servers, power supply units are generally designed for the maximum extent of the overall system, so no explicit requirements in terms of wattage can be set here.

The efficiency of the power supply units is often indicated by manufacturers in different load levels. Here the highest possible efficiency should be ensured. In the mean time, values clearly reaching > 80% have become the state of the art.

To calculate the power consumption of specific server configurations, most manufacturers include on their websites corresponding calculation programs that can be used for sizing the climate control system. For climate control, the Btu/hour is often indicated as the unit (British thermal unit/hour), whereby 1 Btu/hour corresponds to approx. 0.293 watts.

Via the system management, many manufacturers also offer functions that can limit the power consumption of the systems to values clearly below the power values marked on the power supply units. This is primarily aimed at achieving better utilization of existing rack and/or climate control capacities.

Fans

The task of fans in servers is to safely convey the heat generated by components out of the server through the medium of air. Air can only absorb a specific quantity of heat in relation to a corresponding air volume, initial temperature and air pressure. Modern servers therefore evaluate the parameters of temperature and air pressure and thus calculate the required air quantity and the resulting RPM required for the fans depending on the design of the server and fans. Since fans are moving mechanical components, they are subject to wear and should therefore be redundantly designed and monitored. Indicating noise levels for fans in servers is consciously avoided, since servers should be operated in separate rooms. The smaller the design size of a server, the smaller the diameter of the corresponding fans. They must compensate for the required air volumes through having a higher RPM. This is inevitably associated with a higher noise level.

The primary heat sources in servers are CPU(s), power supply units, main memory, hard disks, and I/O adapters. Therefore, fans in servers are fitted with attention to ensuring that sufficient air flows can cool these components.

For servers in the middle and upper power classes, fans are generally redundant and hot-pluggable so as to ensure continuous operation. In low-cost entry-level systems, redundancy and hot-pluggability are rarely found for cost reasons.

Power supply

For rack-optimized servers and blades, the power supply is often provided via power strips located in the rack or also via PDUs (Power Distribution Units). In practice, two separate power inputs are generally provided per rack to ensure corresponding redundancy. These units are often supplied either via power supplies located in the rack or also external uninterruptible power supplies.

These power strips generally offer two types of power connectors for servers. The most frequently used connector for servers in Germany is the IEC 320-C13, also known as the cold device connector. For heavier power consumption, typically above 2000 W, the IEC 320-C19, a rectangular connector with 3 horizontal contacts in a triangular arrangement, is generally used.

The small power strips or PDUs with IEC 320-C13 connectors are themselves supplied either by safety plugs or the IEC 320-C19, e.g. via UPS. Larger power distribution systems generally require single-phase connectors for 32 or 63 amperes. However, the most frequent case in Germany is a 3-phase 16 or 32 ampere connector, also called a "power current" connector, which also provides the highest power.

For planning purposes, it should be noted that the electrical connectors just mentioned must always be designed by a qualified electrician!

A corresponding number of these connectors supply racks in the range greatly exceeding 10 KW with corresponding climate control.

Uninterruptible Power Supplies – UPS

Uninterruptible power supplies are available from approx. 300 VA up to the range of multiple 100 KVA. Those for rack integration, depending on the total power itself, generally offer several IEC 320-C13 connectors and, from approx. 3000 VA, you often find one or more IEC 320-C19 connectors that can be used to connect additional power strips or PDUs. The UPS must report a power outage to the connected servers so that they can be organized in a timely manner and automatically powered down so as to assure the consistency of the data to be processed.

For installations of individual servers, signaling is generally performed serially via an RS232 or USB interface. In rack installations, Ethernet is generally used as the signaling medium and SNMP as the protocol. For this, a corresponding interface card is required for the UPS. Here we must always note that the network components between the UPS and the server must also be backed up by the UPS.

To allow the servers to power down in an orderly way, UPSs must provide a specific minimum bridging time. The times required for the systems depend on the type of applications and should be the subject of inquiries to the technical department. Manufacturers' websites offer corresponding tables for determining the required number of additional batteries.

For planning purposes, it should be noted that the electrical connectors of UPSs above 3000 VA must always be designed by a qualified electrician!

12 Backup and restore or data backup

Data backup designates the partial or complete copying of the data existing in a computer system to another storage medium or another computer system.

The data saved on the storage medium is called the backup copy. The retrieval of the original data from a backup copy is known as a data restore.

The purpose of data backup is to protect the data against hardware loss, for example theft, fire, overvoltages or software defects, viruses, worms and also often user errors. Important in this regard are the physical and spatial separation of the backup from the system to be backed up in order to ensure that the data is protected.

In data backup, we differentiate between archiving, i.e. the long-term unmodifiable storage of the data for review purposes, and simple data and system backups. The principles of archiving and verifiability of digital datasets have been statutorily summarized in Germany since January 2002 for companies in the principles for data access and verifiability of digital documents (GDPdU in German), published by the Federal Ministry of Finance. Archiving will not be discussed further here.

For small and individual systems, tape drives are often used for data backup. In current systems, the serial SCSI interface – or SAS for short – is generally used for this.

If it is only a question of restoring the operational readiness of the system, this is often done using low-cost tape drives based on DDS (Digital Data Store). DDS has been widely available in the market for several years.

Larger data quantities generally require faster drives for backups, which can also store higher data quantities on each tape. The state of the art here is represented by LTO (Line Tape Option). LTO has also been widely available in the market for several years.

To back up networks and larger installations, separate servers are generally used. These servers are generally connected to larger tape autochangers (one drive with multiple tape cassettes – for larger volumes), libraries (tape robots with multiple tape drives, a significantly higher number of tape cassettes – for larger volumes and higher speed) or even larger separate storage systems based on hard disks (online backup – faster restore, but also higher power consumption and no ability to lock up the backup in a safe). These units are often also connected via SAS or even Fibre Channel.

Corresponding backup software is always required for the backup. In this regard, we refer explicitly to the fact that, between the operating system of the server, the backup software and the respectively used tape units, a high level of dependencies exists. We therefore recommend using only mutually certified systems for data backup.

13 Nontechnical requirements

13.1 Choosing the right contract type

13.1.1 Simple server procurement

If only the server itself is procured as standard hardware, then the EVB-IT purchase is the applicable contract.

13.1.2 Other services

In addition to simple hardware procurement, additional services are also often tendered. In the context of server procurement, these are regularly the following services:

- Preinstallations at the contractor's premises
- Minor hardware or software modification services
- On-site installation
- Establishment of the power connections
- Integration into the LAN and SAN infrastructure
- Preconfiguration with operating system and supplied software

According to the nature and scope of the "other services", either the EVB-IT purchase, the EVB-IT system supply contract or the EVB-IT system contract should be used.

13.1.2.1 Preinstallations

Preinstallations performed by the manufacturer/provider are service components of the purchase contract. In this case, the EVB-IT purchase is also applicable.

13.1.2.2 Minor hardware or software modification services

A purchase contract also exists if, in addition to the actual delivery and setup, modification services to hardware or software are also to be provided. This results from § 651 of the Civil Code, which prescribes the validity of the purchase contract law for these cases (with some regulations of work contract law).

13.1.2.3 Integration into the system environment

If the provider, in addition to the simple delivery of standard hardware and/or software to a small extent owes additional services such as integration and installation into the system environment or other minor services (e.g. training courses, introductory support), then the EVB-IT system supply contract should be used.

Since these additional services are non-essential, incidental services to the supply of standard hardware and/or standard software, this contract is also designed as a purchase contract. Accordingly, no approval of the services is required by the client as stated under work contract law. To protect the client, however, an approval-like "arrangement of operational readiness" of the system by the contractor is prescribed. The contractor must demonstrate the "executability of the system" and – if agreed – specific functionalities.

13.1.2.4 Complex server procurements

For those IT procurements in which the "other services" of the contractor are not only restricted to minor integration, installation, or modification services at the client's premises, the EVB-IT system contract is to be used. According to its predominantly work contract-related share of services, it is in fact a work contract, for which reason it also requires an approval.

13.2 Overview and reference source of the EVB-IT

An overview and a decision support for the integration of BVB or EVB-IT contract types into IT procurement contracts can be found on the Federal Government's website for IT officers at: http://www.cio.bund.de/cae/servlet/contentblob/1095360/publicationFile/88813/anlage1_entscheidungshilfe_pdf_download.pdf

The various EVB-IT contract types can be found on the Federal Government's website for IT officers at: http://www.cio.bund.de/cln_164/DE/IT-Angebot/IT-Beschaffung/EVB-IT_BVB/Aktuelle_EVB-IT/aktuelle_evb_it_node.html

There you can also find additional information about legal questions relating to procurement.

13.3 Establishment of the power connections

From the contractor's viewpoint, the premises-side power connection for servers and/or racks is the responsibility of the client. Here standard cables and fuses must be installed. A contractor may inform the client about the requirements (connection values and power consumption values), but the required electrical installation must be provided by the client. Only the operator or its building equipment department knows the precise infrastructure.

On the part of the contractor, all requirements are provided on the server or rack side. For the commissioning of a rack/UPS, the building services should then perform the connection with the power cable of the racks/UPS into the corresponding CEE plug socket or a fixed connection; for an individual server, this is not necessarily required.

The client can also commission these required services with a server/rack order, but then it is necessary to regulate the question of liability contractually.

13.4 Support

When necessary, the corresponding support should be agreed with the specification of the reaction times/restore times.

Commercial offers are subdivided according to:

- Term of the contract
- Reaction times (time between fault reporting and first reaction by support)
- Restore times (time between fault reporting and restoration of server functionality)
- Categorization into error classes (severe, critical and noncritical errors)
- Determination of service handover points
- Spare parts logistics
- Additional technical services (operations support, etc.) according to expense (hourly rates, travel costs)

The requirements may if necessary be:

- 3, 4 or 5-year on-site service
- On-site service with a reaction time of x hours
- On-site service with a restore time of x hours
- SLAs with guaranteed uptimes of 99.x %
- Hotline availability x hours y days a week
- German language hotline
- Spare parts delivery without replacement by service technicians (for hot-pluggable components)
- Spare parts storage by the customer

In the context of procurements for high-availability or safety-relevant solutions, individual agreements can be reached. Here, an estimation of the need for the requirements and the associated costs must be performed.

13.5 Logistics

The following logistical services may be agreed if necessary:

- Specification of the maximum delivery time
- Delivery free domicile
- Delivery free place of use
- Delivery abroad
- Delivery to different locations
- Installation of the servers
- Establishment of the power connections
- Integration into existing rack systems
- Connection to the LAN and SAN infrastructure
- Customer-specific server preconfiguration
- Handling of asset management
- Documentation of the configuration

For deliveries of racks, a site inspection or information on the part of the client regarding the transport route from the delivery vehicle to the setup site is required (e.g. paved paths, door heights and widths, service weight of elevators, number of steps, etc.).

14 Valuation of offers

The previous sections described the criteria for product-neutral tenders. On the basis of this performance description, the offers are drawn up by the tenderers.

These offers are checked and valued by the contracting authority. The contracting authority is obligated to award the contract to the most economical offer. In relation to the evaluation matrix, the "Document for the Tendering and Evaluation of IT Services" (UfAB) in the current version V offers comprehensive support (www.cio.bund.de/cln_093/DE/IT-Angebot/IT-Beschaffung/UfAB/ufab_node.html).

14.1 Measurement protocols

For many evaluation criteria, an evaluation on the basis of individual declarations by the tenderers is adequately possible and generally sufficient. This applies in particular to benchmarks.

We also urgently advise against performing benchmarks, because these tests can be very complex and may lead to extremely different results due to the variety of different configuration options and system settings. Therefore, for benchmark values, it is recommended that the published benchmark results on the respective websites be accessed.

However, there are performance requirements, the fulfillment and thus the evaluation of which can be proven more permanently by measurements on the performance object actually offered. These evaluation criteria include for example power consumption.

In the event of emergencies, the information from the data sheet can be used. Normally, however, only the maximum values are indicated here.

Noise emissions may be neglected at this point, since servers are generally installed in data centers or rooms for technical facilities and not in office environments.

14.2 Evaluation process

To determine the most economically favorable offer, the offers are expertly evaluated with regard to performance and price using the previously created evaluation matrix. The offers are thus checked and given point scores on each evaluation criterion. The calculated point score is multiplied by the set weighting points. The results are added and result in the corresponding performance points for each offer.

After determining the performance points, the price-performance ratio is determined, which provides the decisive reference point for determining the most economically favorable offer. To determine the key figure for the price-performance ratio (Z), the total number of performance points (L) is divided by the total price (P).

$$Z = \frac{L}{P}$$

For more complex tenders (functional performance description, numerous B-criteria), the extended guideline value method is recommended. For offers whose key figure for the price-performance ratio lies within a definable range of variation (5 – 10%) beneath the key figure of the leading offer, this allows the determination of a decision-making criterion (performance, price or another important criterion) which then determines the most economically favorable from among all of the preselected offers.

GLOSSARY

Benchmarks

SPEC Benchmarks

The SPEC Consortium has published a complete series of benchmarks. The descriptions of the benchmarks are available on the website http://www.spec.org. Running benchmarks independently by the procuring authority or the procuring company is generally unrealistic, since the benchmark software from the SPEC Consortium must be purchased and the correct setup and running of a benchmark may require several days.

SPECint

The SPECint value aims to provide information about the computing power of the central processing unit and the server with whole numbers. To this end, the SPECint benchmark software includes 12 different applications. These applications are modified so that they run completely in the main memory, i.e. they make no hard disk accesses. The programs are taken from various subsets of IT – among others, the speed of a special version of the PERL programming language is tested in three different applications; and here is a benchmark which compresses files by means of the bzip algorithm (when the files are in the main memory). Even games (Go), video compression (here too, the data is in the main memory and not on the hard disk), quantum computing simulation and protein sequence analysis are included in the program.

For running the SPECint benchmarks, there are four alternatives:

- SPECint_base

Here the algorithms are newly compiled for the target machine. Here no type of optimization options may be indicated. The result is an average of all the partial results of the twelve algorithms.

- SPECint_peak

Here the algorithms are newly compiled for the target machine. Here too, no type of optimization options may be indicated. The result is an average of all the partial results of the twelve algorithms. In addition, the tester can generate a higher benchmark value by skillfully selecting the compiler options.

- SPECint_rate_base

Here the algorithms are newly compiled for the target machine. Here no type of optimization options may be indicated. During the benchmark run, multiple programs run in parallel in the individual test programs. The result is an average of all the partial results. This run measures how well the server performs when multiple processors or even multicore processors are included in one system.

- SPECint_rate_peak

Here the algorithms are newly compiled for the target machine. Here, it is permitted to indicate optimization options. During the benchmark run, multiple programs run in parallel in the individual test programs. The result is an average of all the partial results. This run measures how well the server performs when multiple processors or even multicore processors are included in one system.

In addition, the tester can generate a better benchmark value by skillfully selecting the compiler options.

It should be noted that the SPECint benchmark values primarily evaluate the combination of the CPU and the main memory. The benchmark suite prescribes that all test data runs in the memory; in this way, the hard disk subsystems and I/O systems are no longer considered. Thus the SPECint benchmark is only conditionally suited to drawing conclusions on real speeds, since application data is provided either from the network or from other servers, or must be loaded locally from hard disks – and this leads to latency, which is not taken into consideration in the SPECint benchmark.

SPECfp

As for the SPECint benchmark, the SPECfp benchmark is intended to make a statement about the processing of floating-point numbers. Analogously to SPECint, not all subsystems are considered. In this benchmark, the procurement officer should consider that today only around 15% of all applications even use floating-point numbers at all, because they can often result in rounding errors, which must not be permitted today in the financial world, for example.

Similar to the SPEC INT, it also includes 4 categories.

SPECjbb2005

The SPECjbb2005 benchmark attempts to measure the speed of a server when executing a Java program. As a test, the benchmark simulates a warehouse in which multiple customers place orders and can make requests. The core of the algorithm is a Java port of the TPC-C benchmark (see below) for Java. In this benchmark, it is notable that the speed depends not only on the server, but quite significantly also on the Java virtual machine (JVM) used and the number of JVM instances. Since the SPECjbb2005 also stores data primarily in the main memory, the speed of hard disk subsystems can only be measured with provisos.

SPECpower_ssj2008

This benchmark attempts to compare the power consumptions of servers to each other. To do so, it launches a Java application on the system that simulates a transaction-oriented warehouse. The application itself runs completely in the main memory, so that the hard disk subsystems are not tested. The Java application is allowed to run for a longer period of time, the number of transactions is then divided by the energy required in each runtime and a value for the "transactions per watt" is then obtained. In this benchmark, the procurement officer must note that, when performing the benchmarks, manufacturers almost always use configurations that are optimized (energy-efficient) for this test. Realistic configurations may deliver significantly different values.

TPC

The Transaction Performance Processing Council provides benchmarks for the testing of databases. Currently, three test suites are being promoted:

TPC-C

This benchmark simulates a warehouse that possesses multiple sites that supply customers with different goods in different districts. In addition to the actual servers, the hard disk subsystem is also included in this benchmark. It should be noted that the data model of this benchmark is extremely simple. The benchmark itself originates from 1992. It has been revised several times, but the underlying data model is still very simple – there are a total of 9 tables and 5 database queries. Since this benchmark actually reads and writes back data from hard disks, it allows a lot of tuning to be performed with large hard disk systems. For standard hard disk systems, today's manufacturers use thousands of disks to achieve the highest benchmark value possible. Here, decision-makers in benchmarks should always take a look at the disclaimers for the benchmark results, because a total system that costs 30 million euros to procure to reach a high TPC-C value is now no longer advantageous.

TPC-E

TPC-E was conceived by the TPC-C consortium as the successor to TPC-C, but has so far not been able to supplant this benchmark.

TPC-H

The TPC-H benchmark simulates a datawarehouse application. It generates results for datawarehouses of different sizes.

The TPC benchmarks consider that the values are very strongly dependent on the hard disk subsystems used. The database programs used also plays a central role. Here it is assumed that future developments in the field of hard disks will lead to new results.

SAP SD benchmark

SAP is today one of the most used standard software programs in business and increasingly also in the public sector. The SAP company itself has developed a benchmark with which system and database manufacturers can be compared to one another. In the SD benchmark, the Sales & Distribution module of the SAP software is tested. It determines how many “standard” SD users the module can use on one server. In this benchmark, it must be noted that only one individual module of the SAP suite is considered. The procurement officer should also consider that, in the introduction of SAP in a company, modifications are often made to the software, to adapt it to the company's specific circumstances. These very modifications may have a positive or even negative effect on performance. Ultimately, a SAP project should not be implemented without the support of a technical department or a consulting firm.

In the actual benchmark, it should be noted today that the benchmark formerly required a response time of 2 second; in 2008, this value was reduced to 1 second, so in comparing older systems to newer systems, the response time should be considered. It must also be considered in the comparison that older SAP software worked only with ASCII characters, but today UNICODE is used, which places a significantly higher resource requirement on the underlying hardware.

VMMark

VMMark was developed by the company VMWare to measure the suitability of x86 server platforms for virtualization. To do so, it executes multiple virtual machines on the test platform. On each of these virtual machines, a separate benchmark runs for a conventional application, as also takes places in a data center: a database, a web server, a mail server, a Java application server and a file server. These virtual instances are combined together with an additional virtual standby server in so-called “tiles”. Within each virtual machine, a separate benchmark is the executed, for example the SPECjbb2005 benchmark is run on Java application servers. For each tile, the VMMark benchmark then determines a consolidated performance value. By indicating the benchmark value, an aggregate value for the individual tile values is then indicated. There is therefore always one number and one “tile” value indicated. Using the tiles, the user can first obtain a guideline value regarding how well the platform is suited to virtualization. (Many times also correspond to many virtual machines.) Based on the total benchmark value per tile, it can also be read out again how well an individual tile is performing. When it comes to many virtual machines, you should look for many tiles, and when it comes to performance in the virtual machines, you should consider the tile/performance ratio.

■ CPU

Socket

The term socket is understood to mean a slot for a processor (CPU) in the server. There are therefore terms like dual-socket machine for servers with two processors and single-socket servers with only one processor (e.g. for simple file servers or print servers). Other synonymous terms are one, two, four our eight-way systems.

Important note on four-socket or multisocket systems:

Servers with four or more sockets or processors are suited to special, high-performance applications. In this case, their conceptualization, design and setup must be performed by specialized experts. Therefore, such systems are also not covered by this guide.

Cores

Each current processor contains two or more cores that work in parallel. The more cores a processor has, the more tasks it can perform simultaneously. In contrast to applications for desktops or notebooks, server applications generally make more efficient use of the supply of cores.

Clock cycles, clock frequency

Each processor is driven by a pulse generator, called the clock. All current processors adapt their working frequency in operation dynamically in stages to the performance requirement so as to save energy when the load is lower. The indicated clock frequency of a processor is generally its maximum frequency.

Power consumption

Manufacturers indicate a TDP (thermal design power) value for a processors. This expresses the power loss in watts of the processor when maximum heat is emitted at the highest clock frequency and load.

For the sake of clarity, the TDP values are divided into stages, called power bands. Up to four stages are differentiated:

- High performance (particularly powerful, therefore increased energy requirement)
- Standard
- Energy-efficient
- Highly energy-efficient

The naming of the energy efficiency stages (power bands) is manufacturer-dependent.

Within a power band, processors with different maximum frequencies are offered. To reach the different stages, different techniques are used. Examples of this are the reduction of the clock frequency, switching off of features, reduction of operating voltage or selection during production.

Processors with particularly high performance or particularly low energy consumption are generally more expensive than those in the standard range.

Memory connection and equipment

The server table in Chapter 2 "Weighting of components by server type" on page 6 shows an indication of the working memory that a server type requires.

Computer manufacturers always offer typical configurations in this regard. There are not only differences in the amount of memory but also in the memory technologies and their connection to the processors.

In memory technology for servers, we differentiate between DDR2 and the newer DDR3 memory.

In the memory connection, modern processors now have so-called integrated memory controllers, but there

are still offerings in which the memory connection is provided via the so-called northbridge.

For all processors, it applies that for an optimal performance, all available controller channels per processor (memory channels) should basically be equipped with the most uniform possible distribution of the memory components via the available slots for each memory channel.

Memory controllers

Here we will only discuss the most frequently used hardware RAID solutions:

- RAID 0: Striping – Acceleration without redundancy: The simple combination of e.g. two hard disks together. Advantage: Through parallel writing to two hard disks, a high performance gain can be achieved. Disadvantage: If one hard disk fails, the data of both hard disks is lost.

- RAID 1: Mirroring: This is the storage of identical data on two hard disks. Advantage: If only one hard disk fails, then the data is generally still available on the other hard disk. Disadvantage: No speed advantage is achieved. Note: RAID 1 is no replacement for regular backups!

- RAID 5: Performance + parity: Like all other rarely encountered RAID types, this is a combination of the above-mentioned versions. RAID 5 is the RAID version that is generally used in servers. At least three hard disks are required (or four, with the hot spare disk). Here the disks are combined as a unit as for RAID 1, but parity information is also stored. With this information, it is possible in the event of the failure of one disk to reconstruct its information after replacing the disk, or to restore it to an available fifth backup disk (the so-called hot spare disk).

- RAID 6: Like RAID 5 with double parity, but as a result somewhat lower write performance.

- RAID 10: Combination of RAID 1 and RAID 0 for particularly write-intensive applications.

Interfaces

VGA video

Server systems generally possess a video or VGA connector on the back of the housing. The VGA connector is an analog image transfer standard for plugs and cable connections between the graphics card of the server and a console. The VGA connection is implemented via a 15-pin D-Sub mini connection.

PS/2

The PS/2 interface was a widely used serial interface for connecting input devices such as mice and keyboards. The PS/2 interface uses a 6-pin mini DIN plug. The USB interface has largely supplanted the PS/2 interface for connecting mice and keyboards.

USB

The Universal Serial Bus (USB) is a serial bus system for connecting external devices. The USB connector is in very widespread use and a variety of products on the market support this standard. The USB ports in the server environment are connected via a flat USB plug of type A “DIN IEC 61076-3-107”. The communication between the individual devices occurs via a point-to-point connection. An individual USB port can be expanded with a USB hub to connect additional devices. Current server systems offer USB interfaces both on the front and on the rear and if necessary internally.

KVM

Multiple computer systems are often connected via a KVM infrastructure to a central console. KVM is the abbreviated collective term for keyboard, video and mouse. A possible VGA connection on the front can be used if necessary for connection to a console. Integrated remote management solutions optionally permit a purely IP-based KVM solution.

Serial

In the 1980s, the serial interface was developed for data exchange between computers and peripherals. For connecting the serial interface, a 9-pin Sub-D plug (RS-232) or an Ethernet cable with RJ-45 connector is used. The data transmission is performed serially. The serial interface is increasingly losing importance and has been largely replaced by USB. However, in the server environment, this interface is still in use.

PCI

PCI-Express is an expansion standard for connecting peripheral devices to the chipset of server processors. PCI-Express is also known by the abbreviated terms PCIe and PCI-E. Through a serial point-to-point connection, PCI-Express achieves a significantly higher throughput rate than its predecessor, PCI-X, which still has a parallel bus architecture. PCI-Express is capable of full duplex and works to increase the data throughput with multiple data lines. Each data line can transmit 250 MB/s (PCIe generation 1) or 500 MB/s (PCIe generation 2) in each direction.

In the server environment, PCIe slots with 4.8 or 16 data lines are offered. A second-generation PCIe slot with 4 data lines is abbreviated to PCIe Gen2 x4. The server can be expanded via PCIe slots with corresponding PCIe cards. PCIe cards are available in various designs. We differentiate between cards with full height and full length and those with a low profile. Depending on the design and technology, servers offer different PCIe slots, e.g. PCIe Gen1 x8 full height/full length or PCIe Gen2 x4 low profile. It must be checked in the individual case whether the desired PCIe expansion cards can be integrated into the server.

Ethernet

Ethernet is a LAN standard for data transfer between two terminal devices. Ethernet was first developed in the 1970s and further developed and standardized by the IEEE in the 1980s. 10 Mb/s Ethernet and 100 Mb/s Fast Ethernet have today been superseded by 1 Gb/s Ethernet (1 GbE). The increasing power density of servers and the better utilization of hardware through virtualization require a network connection with high data throughput. In this environment, 10 Gb/s Ethernet is already sometimes used to save on expensive cabling and energy costs.
In data centers, the 1 GbE server connection is standard. For this reason, servers generally have one or two RJ-45 GbE ports integrated. For an RJ-45 connector, the LAN connection is established using a copper cable. A 1 GbE connection with a corresponding transceiver (SFP) via fiber optic cable is the exception, for cost reasons. Additional 1 GbE and 10 GbE ports can be implemented via additional PCI cards. 10 GbE connections are generally established via fiber optic cable. Low-cost 10 GbE connections over copper (CX4 and SFP+ Copper) can currently only be used for short stretches (< 15m).
Network card drivers inside the operating system permit channel bundling of multiple GbE ports to increase data throughput. In this case, the channels can be operated actively/passively or actively/actively while at the same time offering fail-safety.

Storage (SAS/SATA, Fibre Channel, iSCSI)

There are many options for connecting storage to server systems. We primarily differentiate between local storage inside the server system and external storage outside of the server system. Access to local storage is generally performed via SAS/SATA. Access to external storage can be performed via SAS, FC or iSCSI. NAS solutions will not be considered in this guide.

SAS/SATA

SAS is the abbreviation for Serial Attached SCSI. SAS has replaced the parallel data transfer via a common SCSI bus. The SCSI protocol continues to be used, but the transmission occurs serially via a point-to-point connection between the SAS controller and SAS or SATA hard disks. In contrast to an SAS controller, a simple SATA controller can only address SATA hard disks. The SAS/SATA controller is generally already integrated into the system board of the server or can be retrofitted as a PCI card.

For the connection of an external SAS/SATA hard disk unit without RAID functionality, an SAS controller with external ports is required. In this case, the RAID groups are formed via the SAS controller of the server and an only be used directly by this server. If multiple servers must be able to access a SAS/SATA storage solution together, then the storage solution must possess one or two storage controllers with RAID functionality. For this, the server requires an SAS hostbus adapter with external ports. The RAID groups are formed via the storage controllers of the external SAS/SATA storage solution. Only a small number of servers can be connected to the storage solution, because each server is connected directly to an SAS connection of the storage controller. The distance between the servers and the SAS/SATA storage solution may amount to no more than a few meters.

Fibre Channel

Fibre Channel (FC) is an optimized protocol for the serial high-speed transmission of large quantities of data. Today, transfer rates of up to 8 Gb/s are possible. Fibre Channel can be transmitted by both copper and fiber optic lines. Copper lines are only used inside the FC storage solution. Depending on the storage solution, Fibre Channel hard disks or SAS/SATA hard disks are used in the storage backend. Servers, FC storage solutions and tape libraries are connected to each other with fiber optic cables via a Fibre Channel network (storage area network) and permit the shared use of all resources. The server must be expanded with a Fibre Channel hostbus adapter (PCI card). In the Fibre Channel environment, the availability of data is of the highest importance. For this reason, the data of the Fibre Channel storage is generally accessible via at least two redundant paths. All components must be redundantly designed so that the failure of one of the components does not have any effect on the accessibility of the data. The Fibre Channel network is highly scalable and supports a high number of servers. The distance between the server and FC storage may amount to several kilometers.

iSCSI protocol

iSCSI (internet SCSI) is an option for connecting external storage solutions. It permits the transmission of SCSI packets via an Ethernet infrastructure. The use of iSCSI solutions is associated with low costs at a data transfer rate of 1 Gb/s and builds on the existing network knowledge of the IT department. A data transfer rate of 10 Gb/s is possible, but is associated with significantly higher costs. iSCSI has a comparably high overhead compared to Fibre Channel and the TCP/IP protocol is not optimized for block-based data access. It is therefore useful to keep the iSCSI network separated from the rest of the LAN. The network topology of iSCSI permits access from many servers to a common iSCSI storage. With the aid of an iSCSI driver (iSCSI initiator), the operating system can access the iSCSI network via a standard network card. The additional load on the servers is very low for new, high-performance processors. iSCSI offload and iSCSI boot functionality can be achieved via an iSCSI hostbus adapter and a multifunction network card. The low data throughput of iSCSI is problematic. To solve this problem, the data should be divided across multiple volumes and made accessible via different iSCSI ports to the iSCSI storage solution. The distance between the server and the iSCSI storage may amount to several kilometers, as long as the latencies are low.

The Bundesverband Informationswirtschaft, Telekommunikation und neue Medien e.V. (Federal Association for Information Technology, Telecommunications and New Media) represents more than 1,350 companies, over 1,000 of which are direct members, with around 135 billion euros of turnover, and 700,000 employees. These include providers of software, IT services and telecommunication services, manufacturers of hardware and consumer electronics, and digital media companies. BITKOM particular is working to improve the political and regulatory framework, to modernize the education system and to introduce innovation-oriented economic policies.

The Procurement Office of the Federal Ministry of the Interior purchases goods and services for 26 federal agencies, federally funded foundations, and international organizations. Our product portfolio ranges from A to Z, from alarm technology to tents, to helicopters and a great variety of services. In 2008, we awarded 1,073 contracts with a total value of € 1.0902 billion.

BITKOM

Federal Association
of Information Technology,
Telecommunications
and New Media.

Albrechtstraße 10 A
10117 Berlin-Mitte
Tel.: 030.27576-0
Fax: 030.27576-400
bitkom@bitkom.org
www.bitkom.org

BESCHAFFUNGSAMT
des Bundesministeriums des Innern

Procurement Office of the
Federal Ministry of the
Interior

Sankt Augustiner Str. 86
53225 Bonn, Germany
Tel.: 022899.610-0
Fax: 022899.10610-0
info@bescha.bund.de
www.beschaffungsamt.de