transtec IT Compendium

Edited by
Andrea Schneider / Dieter Weisshaar

Impressum

Author:
Björn Gruner

with the collaboration of:
Andreas Ackermann, Stefan Barth, Wolfgang Bauer,
Corinne Brucker, Dieter Bodemer, Franz Bochtler, Milka Kriegelstein,
Senol Meral, Marco Poggioli, Winfried Reiser, Marco Sanna,
Benjamin Schnaufer, Ute Textor, Guido de Vries

Design:
Björn Gruner

Layout and typesetting:
Grufi Design, Stuttgart

Printing and binding:
Universitätsdruckerei Heinrich Stürtz, Würzburg

transtec
THE EUROPEAN IT FACTORY

© transtec AG, Tübingen
Waldhörnlestrasse 18
72072 Tübingen
www.transtec.de
transtec@transtec.de
Tel.: 0049 (0)7071/703-0

1st Edition 2003

The work and parts thereof are protected by copyright. Any use other than in instances permitted by law requires the prior, written consent of transtec AG. The rights ensuing therefrom remain reserved, particularly those of translation, reprint, use in lectures or talks, extraction of text, graphics and tables, wireless broadcast or transmission through other means of data communication, microfilming or other forms of reproduction and storage in DP systems, even if use is made only in part.
All product names, commercial names, trade names, goods designations etc. are used without warranty that such are freely usable and they may be registered trademarks.
The information in this publication was writing with the greatest possible care. Neither transtec AG, nor the editors mentioned, authors, other collaborators or the editorial department accept the legal responsibility or any liability for any incorrect statements remaining or for any consequences thereof.

ISBN 3-00-011427-0
English Edition:
ISBN 3-00-011425-4
French Edition:
ISBN 3-00-011426-2

Information technology meanwhile faces us everywhere. Whether in our education, at work, while travelling or within our own four walls. IT assists us almost everywhere in our everyday life. Yet we soon reach our limits when it comes to understanding complicated concepts. This means that some things, perhaps even entire processes, may remain beyond our grasp. Lack of knowledge may hence detract from the objective benefits IT yields. IT experts also add to this confusion by the technical language they use - a confusion we would like to dispel.

There is therefore a very high demand for clearly structured and carefully edited information. This updated IT compendium, published in book form, is intended to redress this shortcoming and to contribute towards better understanding of IT. Decades of experience reaped by transtec AG's competent staff have gone into this reference work, making complex concepts and processes all the easier for you to understand.

transtec AG is one of Europe's principal server producers and IT system manufacturers. Our product strategy has led to the reputed IDC ranking of server producers putting us at 9th place in Europe and even 6th place in Germany.

Dieter Weisshaar, Chairman of the Management Board of transtec AG

When transtec AG, the IT specialists for hardware solutions, came up with the idea of writing a book on the current IT terms and detailed IT products, they were spot on. It represents the ongoing development of the yellow pages of the bi-annual product catalogue, which has formed a valued contribution to communication for many years. It is a source of information used by many, whether it be managing director, developer, system administrator, IT buyer or IT journalist.
These are people who are confronted daily with information technology and are always coming up against new questions. What are AIT drives? What does wireless transmission involve? What forms of computer architecture are there?

The IT compendium answers these and many other questions concisely, simply, intelligibly and comprehensively. It addresses those responsible for IT in the public and corporate sectors, students, scholars, and - of course - all those interested in information technology.

The new reference work is to be your constant companion whenever it comes to questions arising from the complex world of IT. It will answer your needs by providing encyclopaedic information while being easy to understand.

Andrea Schneider, Director of Marketing of transtec AG

Tübingen in May 2003

Table of contents

1. Computer Architectures

The following sections contain a brief overview of Intel-based computer systems as well as the most important RISC workstations available today from Sun and transtec.
The CPU is the central and most important component of a computer system, and it is often the most expensive part.
Currently, the most important subject for processors is the transition from 32 to 64-bit architectures. Most RISC architectures have already taken this step, for example SPARC. The Intel IA64 architecture went to 64 bit with the Itanium in the summer of 2001. The follow-up model Itanium 2 has been available since the end of 2002. AMD will be supplying the Athlon 64 (desktop version) and the Opteron (server version) in the 2nd quarter of 2003. An important factor here is not only the availability of the hardware, but also the right operating systems. Microsoft's Windows Server 2003, the successor of Windows 2000, will support 64 bits. Microsoft equipped the first systems with a special, 64-bit extended Windows 2000 version. Linux has responded with an architecture with a 64-bit support. However, only when there are applications available with 64-bit architecture, can the user take full advantage of the performance of the 64-bit support.

1.1 Intel-Based Computer Systems

1.1.1 Processors

PCs are mostly equipped with Intel processors, however, processors from competitors such as AMD are establishing a good reputation and indeed Intel now and then has had to relinquish their speed records to AMD.
Whilst the choice of motherboard was relatively straightforward during the era of the Pentium and Pentium Pro, it is becoming increasingly difficult to select the correct motherboard due to the different processor's designs and the wide variety of chip sets that are available. There are now numerous possibilities for processor slots and the life span for technologies is decreasing tremendously. For this reason, it is practically impossible to equip systems that are older than one year with the latest processors. Therefore, when planning, it should be assumed that only main memory and hard disks can be easily updated at a later time. In most cases, the performance of today's processors does not need to be improved by an update.

1.1.2 Computer Buses

The choice of a suitable bus system is fairly simple because there are only two major bus types available for PC systems: PCI bus and AGP bus. Other bus systems such as ISA, EISA or Microchannel are no longer being used. Nearly every system available today has PCI slots and

an AGP slot. The AGP bus supports only graphics cards.
The following section contains a brief survey of the bus systems currently used in PCs and servers.

PCI Bus

The first PCI bus version already operated at 32 bits with a clock rate of 33 MHz. The PCI bus can support a maximum transfer rate of up to 133 MB/s. The 64-bit version supports a maximum transfer rate of up to 267 MB/s and a crucial advantage of the bus is that it is not dependent on the processor type. Thus, the PCI bus can also be used with "non-Intel processors", e.g. in computers with SPARC architectures. The PCI bus is supported by the PCI Special Interest Group whose members include all major manufacturers of processors, circuit boards, hard disks, controllers and graphics cards.

PCI-X

Soon the performance capability of the PCI bus will not be sufficient to meet the increasing demands being placed on the performance of servers. For this reason, the PCI Steering Committee has created the PCI-X specification. PCI-X was originally developed by IBM, Compaq and Hewlett-Packard. PCI-X allows a slot to operate with 133 MHz and additional slots with 66 MHz or 100 MHz, thus achieving a throughput of up to 1 GB/s, which is double the throughput of the PCI. An advantage of the PCI-X is the reverse compatibility, which means the PCI-X cards can be used in PCI systems and vice versa. Another increase in the data throughput is planned with the PCI-X 2.0. The maximum transfer rate is to be increased to 2.1 or 4.2 GB/s using double data rate and quad data rate. You can find more information at: www.pcisig.com

PCI Express

In the future, PCI-X is to be replaced by PCI Express - the company that started under the name 3GIO (3rd Generation I/O). As a serial "rail bus", the PCI Express can be simultaneously scaled with up to 32 rails that can travel in both directions. With the planned transfer rate of 2.5 GHz, each line is to be equipped with 250 MB/s in both directions, which means a maximum of 2 x 8 GB/s at full capacity. PCI Express should be available as of 2004.

InfiniBand

Another high-performance interface for servers is the InfiniBand from the InfiniBand Trade Association. This is the result of the merger of the Future and NGIO groups. Unlike PCI-X, InfiniBand does not offer reverse compatibility, however, it does offer completely new functions. The basic concepts of the InfiniBand are extremely similar to those of fibre channel: Systems are connected to an I/O "fabric" using one or more Host Channel Adapters (HCA). With Target Channel Adapters (TCA), storage or network controllers are connected to this fabric. The

InfiniBand is addressed with IPv6 addresses, which means that practically an infinite amount of units can work together. The band width can be scaled up to 6 GB/s. Using copper cables, the cabling can reach up to 17 m, and using glass fibre, up to 100 m. This design allows considerably smaller systems to be built with InfiniBand, as bus systems are not needed in the server. You can find more information on the InfiniBand TA homepage at: www.infinibandta.org

AGP

The **A**ccelerated **G**raphics **P**ort is an individual graphics cards slot developed by Intel. Conventional graphics cards use the graphics memory to partially store textures. The graphics memory on the cards was developed to display increasing amounts of data at increasingly high resolutions. However, this memory is not used in standard applications. The AGP concept allows the graphics card to directly access texture data stored in the main memory. A graphics application can optimally select texture data according to size, colour depth, and resolution, and can transmit this information to the AGP card in the main memory. The main-board chip set and AGP exchange the data at 66 MHz with 32 bits, which corresponds to a transfer rate of 266 MB/s. Additionally, the AGP card supports the so-called x2 mode or x4 mode. In the x2 mode, data is transmitted on both edges of the clock signal so the data transfer rate increases to 533 MB/s, four times the data transfer of the PCI bus. A further increase is possible with 4x.

The AGP 4x also includes the additional strobe signals AD_STB0# and AD_STB1#, which form differential line pairs with the standard strobes. These four strobes are then able to work at 133.34 MHz. Even with the AGP 4x, the transfer can be successfully carried out with the falling and raising edges of the AD strobes. The strobe signals' 133.34 MHz produce four cycles per AGP clock, which effectively means 266.67 MHz and a theoretical band width of 1017 MB/s.

AGP is downwards compatible with PCI: PCI bus masters, for example Framegrabber cards, can write data onto the AGP graphics card with the PCI bus. The high bandwidth at which an AGP graphics card can access the system's RAM requires changing to SDRAM with 133 MHz or a Rambus module for RAM technology.

I2O

I2O (Intelligent I/O) is another product which was developed by Intel. I2O is based on the Intel i960RP I/O processor, which contains a complete subsystem on a single chip. Due to the on-chip PCI-to-PCI bridge, the i960RP can be connected directly to the PCI bus without any additional chips. The i960RP then frees up the CPU from processing interrupts and thus improves the performance of the entire system, although this requires software developed specially for this processor. Despite these advantages, I2O was not able to get a foothold on the market and is practically not used today.

PC Card Interface

The PC card interface is a standard for credit-card-sized expansion cards, which was established by the **P**ersonal **C**omputer **M**emory **C**ard **I**nterface **A**ssociation (**PCMCIA**) in September 1990. The first products that used this standard were strictly memory cards. The standard was gradually expanded so other hardware could be designed and used in accordance with the PC card standard, such as interface cards, modems and LAN adapters.

PC cards have standardised surface dimensions of 85.6 mm x 54 mm, and their socket strip has 68 pin holes arranged in two rows. They differ externally only in card thickness: Type I with a 3.3 mm thickness, mostly used for memory cards; type II with a 5 mm thickness for cards, and type III with a 10.5 mm thickness, for example in hard disks. A type III card usually fits into two type II slots on top of each other. The latest development of the PC card standard is the CardBus which has certain improvements such as reduced power consumption, better compatibility and higher performance. The PC CardBus specification describes a 32-bit bus similar to PCI with bus master capability and a frequency of 33 MHz with transfer rates of up to 132 MB/s. Thus, the performance of the CardBus is similar to that of the PCI bus, but for credit-card sized peripherals. This modern bus allows notebooks to use more sophisticated hardware with higher I/O performance (e.g. Fast Ethernet). The PC card standard is downward compatible with the PCMCIA standard, which means that older PCMCIA cards can be used in the new CardBus slot. However, older PCMCIA slots do not support new PC cards.

PC cards are primarily used with notebook and laptop computers. Since these types of computers generally have only a limited number of external interfaces and are normally not equipped with any internal slots for expansion cards, the PC card slots are a simple and versatile alternative for expansion.

1.1.3 Hard disks

Before purchasing a hard disk, the user must decide whether it should be an EIDE or SCSI. The size of a hard disk depends on the application and the operating system used. Nowadays, standard PC solutions have a storage capacity of approx. 40 GBytes. This is in part due to the rapidly falling price per megabyte of storage capacity.

1.1.4 Monitors and Graphics

When choosing a graphics card it is important to ensure that driver software is commercially available for the operating system or is provided with the operating system (e.g. Windows). The maximum resolution of the graphics card and the maximum frequency of the monitor can only be achieved when using the proper driver software. The quality of the graphics card is mainly determined by the graphics processor and the video memory installed on the card. Furthermore, it is important to select a suitable monitor which meets all the requirements

of an ergonomic workplace (high resolution, low electromagnetic radiation, high colour fidelity, etc.) and also supports the performance specifications of the graphics card. If a graphics card is run at a higher video signal frequency than supported by the monitor, it can be damaged. Since the monitor continues to meet future computer standards longer than any other component, do not be deterred from the better quality monitors just because of the higher prices.

1.1.5 Operating systems

Recommendations for operating systems will not be given here. The selection of the operating system usually depends on the application programmes to be used. After all, this decision is best made by the users themselves.
An important consideration for operating systems is the size of the main memory needed to run them properly. It is possible that the amount of memory actually needed considerably exceeds the recommendations of the operating system manufacturer.
For some systems, the recommended RAM is 16 MB. However, this is not sufficient for applications today. If Windows 98 or Windows ME is to be used, then the minimum amount of RAM should be 32 MB. If software packages such as Microsoft Office are to be used, then a minimum of 64 MB should be available. Otherwise the processor will spend too much time swapping data back and forth between RAM and the hard disk. For applications such as Windows NT or Windows 2000, 128 MB is needed to reach an acceptable processing speed. With Windows XP, it should be 256 MB. For Unix operating systems the requirements depend on the Unix variant used. However, at least 64 MB should be available when using a graphical interface.

1.1.6 Standard Applications

Standard applications include word-processing, database, spreadsheet and graphics presentation programmes. Before selecting the components of a computer system, it should be determined if the computer should run all day. If so, the use of a system board with a power-saving function should be considered. The purchase of a tape or CD burner should also be considered if database or graphic presentation packages are used because, usually, the amount of generated data is too large to run backups on diskettes. DVD-ROM drives perform very well when working with graphical applications, and sound cards are useful to complement multimedia applications. If presentations are to be given using a PC during lectures or training sessions, the purchase of an LCD projector would be advisable.

1.1.7 Special Applications

Special applications such as multimedia, computer-aided design (CAD), graphics processing or desktop publishing (DTP) require more sophisticated computer systems than standard applications. Especially in the area of processors and RAM, high performance is needed. DVD drives have become a necessity in this area. As an extra option, a suitable backup medium should be present.

1.1.8 Network Applications

On a network server, the graphics output plays a lesser role. Therefore, a simple VGA monitor is sufficient. When using a server it is more important that the processor, the bus system and the hard disk controller run at the same speed, so that optimal use is made of the storage medium. It is also very important to use a high-performance network card. Some server solutions permit non-dedicated server operation. This means that the server can also be used as a workstation. However, this type of operation is not recommended, as it severely limits the performance of the server functions.

Although system boards with power management functions can also be used in servers, this is not recommended because the response times to the connected stations can take several seconds following operation in stand-by mode (with minimal power consumption).
Network servers should be equipped with a suitable backup medium which can either be incorporated into the server or into a workstation. However, this measure does not provide overall protection against data loss. An optimal solution is the use of a RAID system. In any case, the server system should be provided with a continuous, uninterruptible power supply (UPS), otherwise, data security cannot be guaranteed even when using a RAID system.

With a continuously growing number of connected workstations, even the most powerful network server reaches its performance limits.
In high-load networks, it should therefore be determined, whether or not the load on a server can be reduced by the use of several servers, or by a division of the network into smaller sub-networks linked with bridges and routers.

The configuration of a workstation computer depends on the tasks it has to perform. A workstation computer does not necessarily need its own hard disk, since it is possible to boot the computer directly from the server's hard disk if the network card has a bootROM included.

1.2 RISC Computers from Sun and transtec

1.2.1 Processors

Sun Microsystems is introducing its third generation of UltraSPARC microprocessors, the UltraSPARC III processor, which was developed for the fast-growing Internet and its corresponding data load. It sets new standards for scalability, performance and data throughput and is the heart of the new generation of products from Sun and transtec. Now, Sun Microsystems can offer three types of processors: **Ultra IIe, UltraSPARC III** and the new **UltraSPARC IIIi**, which are available in different clock speeds.
Sun's UltraSPARC processors are available in 360, 400, 450, 480 MHz and have an L2 cache with up to 8 MB. These processors are used in all UltraSPARC models and in the Enterprise server, up to 106 CPUs can be installed (Sun Fire 15K). The models with Ultra IIe are only available in single-processor versions with 500 MHz. The new UltraSPARC III processor has clock speeds of 600, 750, 900 and 1050 MHz. Sun's roadmap envisages clock speeds of up to 2.1 GHz. transtec offers a wide range of alternatives, which are either identical to the original Sun model or differ from it only slightly.

1.2.2 Computer Buses

Sun systems use two different bus systems for expansion cards:

There are two versions of the **SBus** available: A 32-bit version with a transfer rate of 40 MB/s and a 64-bit version with a transfer rate of 100 MB/s. The system automatically recognises the type of card installed and selects the appropriate speed accordingly. The second bus system is the **PCI bus**, which can be found in all of the current models, for example Ultra60, Ultra450, and in the new Ultra III systems, such as "Blade 1000" and "SunFire 280R". The PCI models support the PCI bus specification in a 64-bit, 66 MHz configuration. They can theoretically reach a throughput rate of 528 MB/s. The availability of expansion cards is currently still very limited because most card manufacturers only include drivers for the Windows operating systems with the cards. Solaris requires special drivers. The PCI bus has completely replaced the SBus in all computer systems.
The UPA bus serves as a processor bus and is used in all UltraSPARC systems. On some computer models there are up to 64 UPA slots available.
The Ultra Port Architecture (UPA) uses an internal Crossbar Switch, which links and coordinates the CPU, the memory, and the I/O components of the system. With its throughput of up to 2.4 GB per second, the UPA shows an excellent performance. The UPA bus uses the 64-bit UltraSPARC processor with integrated 64-bit processing very effectively.

1.2.3 Peripheral Buses

The new models from Sun are equipped with FireWire (IEEE 1394) and USB. The **U**niversal **S**erial **B**us is primarily used to connect the keyboard and mouse. In the future, a fibre-channel arbitrated loop will be used for internal hard disks and external storage arrays. An Ultra SCSI interface is available for external peripherals. Network connection is established using a 100 Base-T interface. For all high-end graphics applications, the CREATOR Frame Buffer, which is plugged into a UPA slot, provides high performance. The CREATOR card combines both the large linear memory of a frame buffer and the direct access to the 128-bit UPA system architecture with the powerful Visual Instruction Set (VIS) of the UltraSPARC CPU.

1.2.4 Operating systems

Solaris, which is known for its maturity and stability, is the main operating system used on SPARC computers. The current version, Solaris 8, is fully 64-bit compliant and meets the most stringent security demands as well as providing optimum performance for Java applications. Further information on Solaris can be found in the chapter Operating Systems.
The Linux operating system is also available, but is not yet widely used.

2. Operating Systems

The following sections provide a brief overview of the most common Microsoft and Unix operating systems.

2.1 Windows

2.1.1 Windows 2000

Microsoft Windows 2000, referred up until now as Windows NT 5.0, has been expanded by the addition of several new features and functions as compared with Windows NT. These concern the areas of administration, scalability and expandability, as well as storage and hardware management. Microsoft offers Windows 2000 in four versions:
Windows 2000 Professional corresponds to the Windows NT Workstation and supports up to 4 GB main memory and two processors.
Windows 2000 Server is the successor of the Windows NT Server and offers hardware support for max. 4 GB main memory and four processors. The Windows Terminal services, which replace the Windows NT 4.0 Terminal Server edition, are already contained in this server version.
The Windows NT Enterprise Edition will continue as **Windows 2000 Advanced Server**. Up to 8 GB main memory and 8 processors are supported. In addition to the Windows 2000 Server functions, the IP load balancing (with up to 32 servers) and failover clustering for two servers are also available.
The **Windows 2000 Datacenter Server** represents the top end, supporting up to 32 processors and 64 GB main memory. It offers the following additional functions over Windows 2000 Advanced Server: Failover clustering for 4 servers and process control for work load management. Another important feature is the support of virtual servers. Thus, several instances of the operating system can be run on multi-processor servers, for example 2 virtual Servers each with 4 processors can be set up on an 8-processor server.

Installation of Windows 2000: Windows 2000 is installed in the computer without an operating system using a bootable CD. Plug and Play is a new feature of Microsoft Windows 2000, which simplifies the installation process. Another improvement made to Windows 2000 is that it must be rebooted very seldom in comparison to Windows NT. The USB support has also been implemented in Windows 2000. Unlike the Windows NT server, when installing Windows 2000, it is not necessary to determine if the Windows 2000 server should be used as the domain controller, or not. With the help of the Configure Your Server wizard, the service for the Active Directory (directory service especially for user administration) can be installed at a later stage.

Repair mechanisms: Windows 2000 is equipped with an improved, protected boot mode. An additional, improved repair mechanism has been implemented in the command line.

Administration: Microsoft Windows NT 2000 implements Active Directory as a central platform that simplifies access to and management of network and system resources. Unlike in the User Manager for Windows NT, users can be organised, configured and managed hierarchically in containers in the Active Directory. With Windows 2000, the user administration is not just more structured, there is no longer a limit of approx. 20-40,000 users per domain as there is under NT. Further features include centralised configuration management and the configurable and expandable Microsoft Management Console (MMC).
IntelliMirror technology allows Windows 2000 workplaces to be centrally configured. With the help of the Active Directory, the configuration presets for users or groups can be centrally filed. This means that the user will have exactly the same configurations at all Windows 2000 workplaces and the user's software will be automatically installed at the respective workplaces. The configurations can also be set so that the user can not change them.

Scalability and Extendability: Windows 2000 supports up to 64 GB of physical memory. With the Microsoft Cluster Server, two or more servers can operate in an interlocked way. Thus, the devices can monitor each other so as to maintain operation without interruption if one of the servers fails. During normal operation, the servers can share the workload, thereby increasing productivity.

Storage Management: NTFS now also implements quotas to define the maximum amount of disk space available to the users. The NTFS expansion EFS (Encryption File System) allows the encryption of sensitive data on file level or directory level.
With the DFS file system, the distributed structures of files and data on Windows 2000/NT, NetWare and Unix servers can be summarised and presented in an organised manner. In this way, users can find files in the network much more easily.

Hardware Management: Plug and Play allows PC cards to be used with portable computers without any difficulty. In addition, the expansion of the Windows Driver Model (WDM) is intended to enable identical driver software to be used in Windows 98 and Windows 2000.

Security Functions: In order to increase the operational security, Windows 2000 prevents deleting critical operating system files. In addition, it only allows the installation of certified drivers.

Network Security: The Security Service Provider Interface (SSPI) is already in use in Microsoft Windows NT 4.0, for example in the NT LAN Manager and the Secure Sockets Layer (SSL).

In Windows 2000, the SSL will be expanded and Kerberos Authentication will be introduced in accordance with Kerberos V5. Furthermore, Windows 2000 supports Smart Cards, which increases the security in user log-ons and digital e-mail signatures.

2.1.2 Windows XP

Windows XP impresses with its significantly improved, clear design, permitting intuitive operation even for users with relatively little practice. In addition to the new design, a multitude of additional capabilities have been developed.
The **redesigned Start menu** groups the most frequently used applications for easy access. The five most-used programs are displayed first, and the default e-mail program and Web browser are always available. The **new taskbar grouping** keeps the taskbar clean and organized. Open files are now grouped according to application type.
In addition to the design, Windows XP has a series of further novel features enhancing **user friendliness**. Windows wakes from hibernation far faster, i.e. the batteries of the laptop no longer have to run for nothing, because all the applications are immediately ready for use again when called up. It is also possible for several persons to use one PC together or side by side. A separate account is created for each user. The users can be changed quickly and simply while the other person's applications keep running. This means, for example, that downloading a file from the Internet does not have to be discontinued when there is a change in user, because all the open applications continue to run in the background.

There has also been a great improvement in the **hardware support.** Devices such as digital cameras, scanners or DV cameras are automatically recognised and all the recording, processing and playback functions Windows XP provides can be universally used.
The performance has also been significantly enhanced. The first improvement under Windows XP becomes clear the moment the computer is started up. After several starts with the same software and hardware, Windows XP arranges the start files on the hard disk for rapid access. Together with the so-called **Prefetching** feature and an optimised network login, the system starts up to 34% faster than with earlier Windows versions. The same function also means programs start faster.

The **enhanced multitasking performance** is supported by measures including utilizing downtime for system activities, adapting the user interface to the computer's potential and efficient management of memory.

The **Standby and Hibernate mode** are of assistance for laptops. In Standby, the system powers down the monitor, hard disk and other devices, but continues to supply the main memory where the open documents and programs are stored. After the system has been

re-activated, work can be resumed where it left off, in less than 2 seconds with relatively new laptop models. If the laptop enters the Hibernate mode, the data from the main memory is saved to the hard disk in compressed form and the computer is shut down completely. Startup from this mode takes a little longer, but the desktop state is as it was before Hibernate was entered. A further improvement has been made to the power management functions introduced in Windows 2000. Hibernate and Standby now operate considerably faster and more reliably, saving Notebook batteries and increasing mobile operating time.

With **Remote-Desktop** Windows XP allows users remote access to all their applications, documents and network connections on their desktop computers. Through a secure, encrypted Internet or dial-up connection, users can link to the company PC and create a virtual session there with the customary Windows desktop on the remote computer. Since all the programs run on the company PC and only keyboard and mouse inputs and display outputs are transferred, the remote desktop is also ideally suited for modem or ISDN connections.

Wireless networks (such as 802.11b-WLANs) are supported by Windows XP Professional as standard. Windows XP Professional recognises and configures everything automatically as soon as a WLAN connection is established or cell-to-cell roaming takes place. Certificates that are distributed or stored on smart cards as well as 802.1X authentication ensure the utmost transfer security (both in cable Ethernet and in wireless networks).
Windows XP Professional recognises networks and their TCP/IP settings, so that the **network connectivity** of the PC is established automatically and is maintained during roaming. If the network does not provide for automatic configuration, Windows XP Professional allows working with alternative TCP/IP settings. In this way a notebook user can draw the TCP/IP configuration dynamically from a DHCP server in the company network and work with static TCP/IP addresses at home.

Windows XP is an uncompromising further development of the proven Windows 2000 technology and therefore operates more stably than its predecessors. The new Windows engine is based on the dependable 32-bit architecture of Windows 2000, featuring a fully protected memory model. During installation, the setup program of the operating system can incorporate driver, compatibility and security updates from the Internet if the user chooses to install them, even if such have become available after the Windows XP CD-ROM was issued. **Windows Update** keeps the system abreast of the latest changes. Service packs, troubleshooting and drivers for new hardware devices can be downloaded from Microsoft on the Internet easily and automatically if the user so wants.
A large number of programs run under Windows XP that failed to run under Windows 2000. If a program can be run only under Windows 95 or Windows 98, most likely it can now also be run under the new operating system. Should an older application nevertheless

not be supported by Windows XP, such can be performed in a special **compatibility mode** presenting the characteristics of earlier Windows versions.

Windows XP gives top priority to the **quality of device drivers** since it contributes significantly to system stability. When a new component is installed, an indication is given whether Microsoft has tested and certified the driver. If a non-certified driver has been installed and fails to operate correctly, the system can revert to the earlier operational driver within a matter of seconds. By means of the side-by-side DLL support, multiple versions of individual components can run side by side, so that every application uses the version best suited to it. If failure is experienced despite the significantly enhanced driver handling, the **System Restore** feature can restore the PC to a previous state and thus cancel all configuration changes made in the meantime, even after weeks have elapsed.

Great importance is attached to security in Windows XP. Sensitive files can be stored encrypted, passwords and user names in special areas are protected from unauthorised access. The **Encrypting File System (EFS)** permits files containing particularly sensitive information to be stored in encrypted form on the hard disk of the PC or on network servers. This dependably protects data including offline files and folders from unauthorised access. The new multi-user support of EFS allows multiple (authorised) users to encrypt and jointly inspect sensitive files. The **Internet Firewall (Internet Connection Firewall - ICF)** integrated in Windows XP automatically protects data from Internet attack. The function just has to be activated for the data communication or LAN connection wanted. Windows XP Professional uses 128-bit encryption for all encryption operations. In addition, there are many further security features in Windows XP, such as Kerberos V5 authentication support, Internet Protocol Security (IPSec), centrally defined security guidelines and a lot more.

A series of features in Windows XP simplifies installation and administration of the operating system and helps cut the costs for these tasks within the company. Just a selection of these features is given here.
For companies wanting to install XP Professional on several PCs, Microsoft has provided sophisticated, tried and tested mechanisms for **automatic installation**. These allow unattended Windows XP installations to be pre-configured, enabling the setup routine proper to run reliably and without user interaction. Operating system images can be created with ease, even for PCs with different hardware, and can be installed automatically across the network with the aid of the Windows 2000-compliant Remote Installation Service (RIS).
Of interest to international companies and educational establishments: Windows XP Professional supports the **Multilingual User Interface** (**MUI**): A PC can have a number of interfaces at a time (English, German, Spanish etc.), providing every user with the Windows desktop in the language he or she prefers.

Administrators can deploy **software restriction policies** to prevent unwanted applications from running on Windows XP Professional PCs, thus restricting workstations to the programs important for the company.
In addition to the Group Policy settings provided for Windows 2000, there are numerous new ones provided for Windows XP Professional for even more comprehensive, policy-based management through the Active Directory. With the new **Resultant Set of Policy (RSoP)** Windows XP Professional tool, administrators have a powerful tool to plan, monitor and troubleshoot the impact of various group policies on a specific user or computer. The location-related group policies prove to be exceedingly useful for users often en route with their notebook. If the user is at the office, the company group policies apply. When the user is away or at home, however, he can employ the Windows XP Professional functions useful for single PCs or in small LANs, without this requiring reconfiguration by an administrator or the user himself.
The **USMT (User State Migration Tool)** allows easy migration of a user's data and personal settings from the original system to Windows XP. Preferred desktop settings, folder options of the Windows Explorer, Internet configurations and favourites, e-mail access, messages and addresses, network drives and printers, specific folders etc. can be migrated with ease.

Windows XP offers universal **support** and end-to-end-integration **for all digital media**. Transferring and recording, display and playback, archiving and sending - Windows XP offers context-sensitive support through every step and automatically proposes all the tasks coming into question. Windows XP opens up audio, photo and video capabilities as never experienced before. The Media Player can create MP3 files and play DVD videos via the PC through separate plug-ins. Windows XP automatically recognises if a digital camera is connected to the PC or the scanner.
Windows XP identifies what files are in a folder and not only displays a thumbnail, but also proposes the most likely options for the respective **media type** in the new taskbar: Audio data, for example in the My Music folder, is played at a click of the mouse.
The **Remote Assistance** integrated in the Help and Support Centre of Windows XP enables colleagues or an in-house support centre to be called for help at any time. Instead of the problem being described to a specialist in words over the phone, the expert can show the solution in a clear and intelligible manner on the user's own screen by remote control. If the expert's aid is enlisted via the online contacts of the Windows Messenger or via e-mail, he receives the authorisation to view the user's screen on his own PC to help find or even implement the solution. The user can follow every step and retains full control at all times. The expert's Remote Assistance authorisation for the user's PC expires automatically after the session ends. This comprehensive and central support tool of Windows XP allows the number of calls for support addressed to the company's own support centre to be reduced and hence cuts the costs for support with hardware and software problems.

2.1.3 Windows Server 2003

The new Microsoft Windows Server 2003 will be available as from April 2003. As compared to the Windows 2000 Server, many features and functions were developed further or from scratch. The following Windows Server 2003 versions will be available:

Windows Server 2003, Standard Edition
Two-way symmetric multiprocessing (SMP)
4 GB RAM

Windows Server 2003, Enterprise Edition
8-way symmetric multiprocessing (SMP)
Clustering with up to 8 nodes
32 GB RAM in 32-bit versions and 64 GB RAM in 64-bit versions
Hot Add Memory
Non-Uniform Memory Access (NUMA)

Windows Server 2003, Datacenter Edition
32-way symmetric multiprocessing (SMP)
64 GB RAM in 32-bit versions and 128 GB RAM in 64-bit versions
Windows Sockets: Direct access for SANs
This version will be available only through the Windows Datacenter program, which offers a package of hardware, software and services

Windows Server 2003, Web Edition
2-way symmetric multiprocessing (SMP)
2 GB RAM
Windows Server 2003, Web Edition is tailored for use as a web server. Servers using this operating system can be members of an Active Directory domain, but not independently offer the Active Directory service. This version will be available only through special partners.

An overview of the principal features of Windows Server 2003:
XML Web services: The XML Web services provide reusable components built on industry standards that invoke capabilities from other applications independent of the way the applications were built, their operating system or platform, or the devices used to access them. The IIS 6.0 security settings are locked down during setup to ensure that only required services can be run. Using the IIS Security Lockdown wizard, server functionality is enabled or disabled based on the administrator's requirements.

Directory services: Active Directory security settings for users and network resources span from the core to the edge of the network, helping to make a secure end-to-end network. Active Directory is now faster and more robust, even over unreliable WAN connections, thanks

to more efficient synchronisation, replication, and credential caching in branch office domain controllers.

Update Management: The automatic update provides the ability to systematically download critical operating system updates, such as security fixes and other patches. Administrators can select when to install these critical operating system updates.

Internet firewall: The system's Internet firewall makes Internet connection more secure.

Server hardware support: Driver verifiers check new device drivers to help keep the server up and running.

Application verification: Applications running on Windows Server 2003 can be tested in advance with a view, for instance, to heap corruption and compatibility issues.

Server event tracking: Administrators can report an accurate record of uptime using the new server shutdown tracker. It writes Windows events for server shutdowns to a log file.

Configure Your Server wizard: The Configure Your Server wizard steps administrators through the process of setting up various server roles such as a file server, print server, or remote access server, ensuring that components are installed and configured correctly the first time.

Manage Your Server wizard: The Manage Your Server wizard provides an interface for ongoing management of the server, making it easy to perform such common tasks as adding new users and creating file shares.

Remote server administration: With Remote Desktop for Administration (formerly known as Terminal Services in Remote Administration mode), administrators can manage a computer from virtually any other computer on the network.

Shadow copy: This feature provides time-based network sharing. Administrators can save network folder contents and later determine the status of the folders as they existed at this time. End users can recover accidentally deleted files or folders on network shares without requiring system administrator intervention.

Terminal Server: When using Terminal Server, a user can access programs running on the server. For example, a user can access a virtual Windows XP Professional desktop and x86-based applications for Windows from hardware that cannot run the software locally. Terminal

Server provides this capability for both Windows and non-Windows-based client devices.

The additional functions of the Enterprise Edition include:
Cluster service: The cluster service for Windows Server 2003, Enterprise Edition and for Datacenter Edition supports up to eight-node clusters. This provides increased flexibility for adding and removing hardware in a geographically dispersed cluster environment, as well as providing improved scaling options. Server clusters can be deployed in a variety of different configurations:
- Single cluster configurations with dedicated storage
- Multiple clusters on a storage area network (SAN)
- Clusters spanning multiple sites (geographically dispersed clusters)

Metadirectory Services support: Microsoft Metadirectory Services (MMS) helps companies to integrate identity information from multiple directories, databases and files with Active Directory. MMS provides a unified view of identity information, enables the integration of business processes with MMS, and helps synchronise identity information across organisations.

Hot Add Memory: Hot Add Memory allows memory to be added to a running computer and made available to the operating system and applications as part of the normal memory pool. This does not require re-booting and involves no downtime. This feature currently will operate only on servers that have the respective hardware support.

Non-Uniform Memory Access (NUMA): System firmware can create a table called the Static Resource Affinity Table that describes the NUMA topology of the system. Windows Server 2003, Enterprise Edition uses this table to apply NUMA awareness to application processes, default affinity settings, thread scheduling, and memory management features. Additionally, the topology information is made available to applications using a set of NUMA APIs.

Terminal Services Session Directory: This is a load balancing feature that allows users to reconnect to a disconnected session. Session Directory is compatible with the Windows Server 2003 load balancing service and is supported by third-party external load balancer products.

The additional functions of the Datacenter Edition include:
Expanded physical memory space: On 32-bit Intel platforms, Windows Server 2003, Datacenter Edition supports Physical Address Extension (PAE), which extends system memory capability to 64 GB. On 64-bit Intel platforms, the memory support increases to a maximum of 16 terabytes.

Windows Sockets: Direct access for SANs: This feature enables Windows Sockets applications that use TCP/IP to obtain the performance benefits of storage area networks (SANs) without making application modifications. The fundamental component of this technology is a multi-layer Windows Socket service that emulates TCP/IP via native SAN services.

2.2 Unix Operating Systems

Unix is still the leading operating system in the workstation world. In fact, it is a family of operating systems because practically all workstations manufacturers supply their own version of Unix, which at least as far as the user interface is concerned, differs considerably from the other versions. However, there is a tendency to overcome this wide variance of interfaces, as several manufacturers have begun to port their system to alien architectures.
The Unix implementations can be categorised under two standards: Berkeley Unix (BSD) and AT&T's System V Release 4 (SVR4). At present, the SVR4 is ahead of its rival - new versions of Unix follow its standard. As a general rule, if a programme is written for one of these two standards, it can be ported to another system of the same standard without major difficulty. Different standards are also employed for the user interfaces (GUI - Graphical User Interface). However, the more recent ones all follow the X11 definition. For several years, the MOTIF definition - which is also based on that of X11 - has clearly been progressing. More and more Unix implementations are using this interface, while the use of the competitor interfaces, like OPENLOOK, have been on the decline.

2.2.1 Linux

Linux is a freely available multi-tasking and multi-user operating system. Linux was invented by Linus Torvalds and developed further by a number of other developers throughout the world. From the outset, Linux was placed under General Public License (GPL). The system can be distributed, used and expanded free of charge. In this way, developers have access to all the source codes, thus being able to integrate new functions easily or to find and eliminate programming bugs quickly. Thereby drivers for new adapters (SCSI controllers, graphics cards, etc.) can be integrated very efficiently.

Presently, Linux is successfully being used by several millions of users worldwide. The user groups vary from private users, training companies, universities, research centres right through to commercial users and companies, who consider Linux to be a real alternative to other operating systems.
The extensive network support of Linux, including different servers such as Appletalk, Netware or LAN Manager servers as well as the multitude of supported network protocols, makes Linux a secure and reliable network server system.

There are two different ways of obtaining Linux: All the necessary components can be downloaded free of charge from the Internet. This means that an individual operating system can be assembled almost free of charge. An alternative is to use a so-called Distribution, offered by various companies. They include a wide range of applications and installation programmes that significantly simplify the installation of Linux.
The distributions differ especially in terms of the included components, such as programming environments, network software and graphical user interfaces. We recommend distributions from SuSE or Red Hat, as both these Linux distributions are very sophisticated and include a wide range of documentation as well as a graphically supported installation. All transtec Linux systems are certified and offered with the current versions of SuSE and Red Hat.

In addition to their distributions for PCs and workstations, both SuSE and Red Hat offer special packages for server operation. In the case of SuSE that is the SuSE Linux Enterprise Server. Apart from the composition of the package specifically for server operation, it is distinguished from the "normal" distribution by the following points. For one thing, SuSE carries out extensive tests to ensure that the individual packages are compatible with one another and with important business applications. What is more, SuSE guarantees up to 2 years' support for the package, even after the respective version has been superseded. Equally important for the business environment is the provision of important patches and updates.
Furthermore, SuSE offers complete packages for special applications, such as e-mail servers or firewalls.
The development of these packages reflects the general trend: The use of Linux in commercial systems as well is on the increase, where it provides a cost-effective and dependable basis for many applications.

2.2.2 SunOS/Solaris

Solaris 8
The Solaris 8 version is the second generation in 64-bit technology. Solaris 8 is no longer restricted as are Windows or other environments. Practically an unlimited amount of Internet addresses (more than a 1 with 38 zeros), 18 Exabytes of main memory (that means 10 to the power of 18 Bytes) and more than a million simultaneous processes are supported. The Solaris 8 operating system also supports the latest technologies from the desktop to the Internet. This includes Java 2SE for the development of Web-centric software applications, the Java Media Framework for media streaming, X-Server video extensions, the synchronisation of PDAs (e.g. Palm Pilot) and Networked Real Time processing. Solaris 8 also offers additional mainframe abilities: Jobs, projects and accounting functions for balancing computer operations (ideal for service providers); automatic dynamical reconfiguration for improving uptime; and hot patching for dynamic alterations of operating systems without shutting down the system.

A Web-Install Tool makes the software installation very simple. Solaris applications which are already in use are compatible with the Solaris 8 software. Solaris 8 is essential for all systems with Ultra III or Ultra IIe processor.

Solaris 9

The Solaris 9 operating environment becomes a comprehensive service platform with a large number of software components integrated as standard. A Java 2 Platform, Enterprise Edition (J2EE) application server and a Lightweight Directory Access Protocol (LDAP) directory server provide the foundation for identity management and are just as much a part of the standard system as Solaris Volume Manager, Solaris 9 Resource Manager and a multitude of further products.

Solaris 9 helps lower the total cost of operations through the significantly reduced acquisition costs for key applications of the infrastructure. The integration with the **Sun Open Net Environment (Sun ONE)** brings new efficiencies and new ways of computing. Delivering services through the Sun ONE platform means integrating Web, application and directory servers, as well as file systems, volume management, security and provisioning.

With the **Sun ONE Directory Server** fully integrated and the **Sun ONE Application Server** included, Solaris 9 makes it easier to deploy scalable, reliable services on demand, with no incremental costs. Coming updates will include full integration of the Sun ONE Application Server, making Solaris 9 the only operating environment to include fully integrated J2EE application and LDAP directory servers in the core operating system, an approach inviting comparison with popular individual solutions.

Solaris 9 also helps lower the cost of operations with new **provisioning** and change management features support fast and smooth system expansions. What was a formerly 4-hour operation takes 20 minutes with features such as the **Solaris Flash Software**. Complete and accurate configuration control saves IT staff valuable time in routine tasks, bringing system management efficiencies. **Solaris 9 Resource Manager** brings integrated resource management that allows out-of-the-box server consolidation with no additional product purchases. **Solaris Volume Manager** contributes to high data availability and maximum reliability in working with Solaris 9.

Solaris 9 delivers the high level of security required for network services by security solutions supplied as standard and fully **integrated security services**. The new version of Solaris provides the decisive added security by a series of key security features. **SunScreen 3.2** ships with a firewall which need fear no comparison with separately available products. **Solaris Secure Shell** provides functions for secure, encrypted remote access. Other security features of Solaris 9 include Secure LDAP, Kerberos V5 server, and secure access via the Sun ONE Portal Server. These robust security features are all part of the core operating system to make all applications secure end-to-end at no incremental cost.

The demands of enterprise and high performance computing (HPC) mandate highly reliable

and scalable 64-bit environments. Solaris 9 and the UltraSPARC III Cu processors are able to make applications run faster without recompiling, recoding or redesigning. Some of the new features have delivered the following performance improvements:
Enhanced multithreading library improves **Online Analytical Processing (OLAP)** (Sun internal test with Oracle Express) by up to four times.
The UltraSPARC III processor **on-chip memory controller** allows data and code localization, optimized to increase system performance by 5-40% for high-end servers.
The UltraSPARC III processor has a **Dual Translation Lookaside Buffer (DTLB)** with support for large memory pages, increasing HPC performance almost three times (measurement according to SWIM benchmark).
A novel page colouring algorithm increases system performance by up to 10 percent for typical server loads.
Numerous features in the new Solaris are designed to **improve availability** and lower the total cost of operations. These include enhanced integration with the **Sun Cluster 3.0** software and configuration management features. In addition, Sun's RAS Profile technology and processes provide the best solutions to optimize customers' platforms. A reduced rate of errors, simplified problem resolution, reduced unplanned downtime, and increased productivity all lead to cost savings and increased efficiency.

2.3 Computer Viruses

A sad fact on nearly all computer platforms and operating systems are computer viruses. These are nowadays practically only spread via e-mails and the Internet.

Types of Viruses

A computer virus is an instruction sequence which, if executed, changes a memory area by copying itself there. This storage area can be an executable file, or a programme stored on floppy disk, hard disk, etc., or also in RAM.

File Viruses

File viruses attach themselves to selected programmes, employing various techniques to do so. They are spread whenever an already infected programme is called up. **Resident file viruses** represent an intensified version in as much as they take root in the memory after they are called up. Thus, each time a programme is executed, infection is repeated. This type of virus is harder to eliminate; in some cases, it can even survive a warm start of the computer. Most of the viruses known today belong to this category. So-called **stealth viruses** were developed to go undetected by virus scanners, checksum programmes, or protection programmes. Similar to the U.S. Stealth bomber, they possess a camouflage mechanism that enables them to conceal their existence. **Polymorph viruses** are the latest and most dangerous kind of file viruses.

This category consists of viruses, which alter their appearance for each new infection. Normally, they do this by re-encoding the virus code with a key which varies with each infection.

System viruses use particular programme components, like the master boot record, partition boot record, diskette boot sector, FAT, and the operating system's root directory as the infection medium. These system areas are connected to the respective data media, and infection occurs when the respective data media is used.

Boot sector viruses attack exclusively the boot sector on floppy disk or hard disks. The infection spreads slowly, since it can only take place during booting up from an infected floppy disk.

transtec recommends that systems should be continuously checked with the help of suitable virus software so as to prevent extensive damages.

3. Clusters

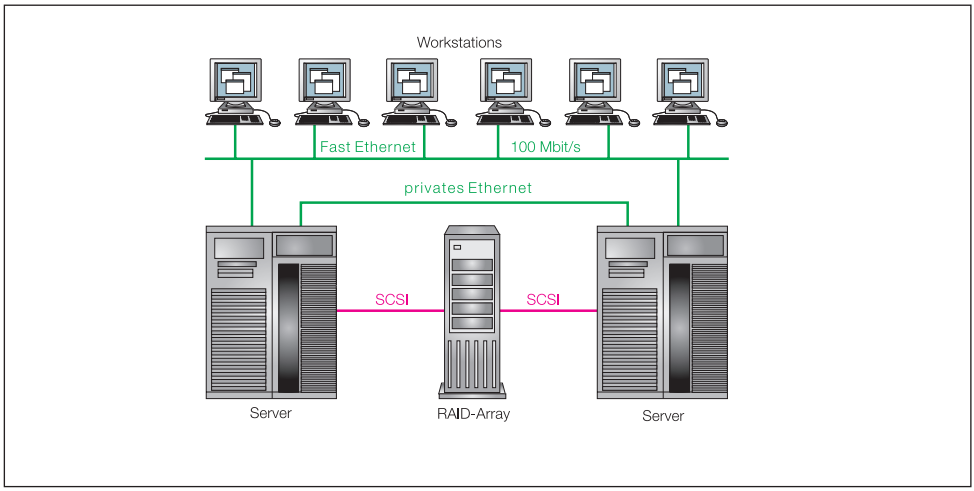
High-Availability Cluster

A **high-availability** cluster refers to the linkage of several systems to form a group which is managed and controlled by a cluster software. Physically, a cluster is a group of two or more independent servers that serve the same group of clients and can access the same data. Putting this into practice with the latest technologies means, in general, that servers are connected via the usual I/O buses and a standard network for client access. From a logical point of view, a cluster represents a single administrative unit, in which any server can provide any authorised client with any available service. The servers must have access to the same data and be equipped with a common security system. According to the latest technological standards, this means that the servers connected to a cluster generally have the same architecture and run with the same version of an identical operating system.
Although there are various possibilities for the set-up of clusters, they all provide the following three advantages:

- Fault-tolerant high availability for applications and data
- Scalability of hardware resources
- Simpler administration of large or fast-growing systems.

Higher Availability

In company-critical environments, the high availability of services (e.g. web servers, databases, or network file systems) is the basis for success of a company. Common reasons for problems with the availability of services are the different forms of system malfunctions. A malfunction can be due to the hardware, the software, the application or the system execution. Protection against these malfunctions can be provided using a cluster as an application server or a data server. Besides providing redundant computing performance, I/O and storage components, clusters also enable one server to function for another server without interruption, if the latter breaks down. In this way, clusters allow high availability of server services even when malfunctions occur.
The illustration depicts a simple cluster, which consists of two networked servers and a RAID system for data storage.



- When hardware or software malfunctions occur, or when a server fails, the other server can take over its functions
- When the network interface of a server fails so that the connection to the clients is interrupted, the clients still have the possibility of using the second server
- When an I/O bus or an I/O adapter fails, the data of the RAID system can be accessed via an alternative path
- In case of a disk failure, the data can still be accessed via the RAID system.

In all cases, both the servers and the clients must identify and deal with the malfunction. It is true that the cluster service performance is interrupted, but with clusters, this interruption normally lasts only a few seconds and not several minutes or hours as is the case with conventional restoration methods.

Scalability

A further advantage presented by some cluster architectures is the scalability, enabling application growth beyond the capacity of a single server. Many applications have multiple threads of relatively de-limited activities, which interact only on occasion. Applications with multiple threads can run as pseudo-parallel processes on a server with one processor or as true, parallel processes in symmetrical multiprocessor systems (SMP). In a cluster, groups of the application threads can be run on different servers, because the servers can access the same data. If an application becomes too large for one server, a second server can be installed to produce a cluster and thus increase the application capacity. Co-ordination is necessary to enable the servers to access the same data. This co-ordination can be realised through an application, a database manager or a distributed file system. In a cluster lacking in this

co-ordination, the simultaneous, direct access of any file by multiple servers is not possible.

However some applications can be scaled even with this limited access. The applications can be divided in such a way that the individual programmes use different data.

Easier Administration

The third advantage of clusters is their easier administration. Clusters simplify the complexity of system administration by enhancing the scope of the applications, data and user domains administered by a single system. Among other areas, system administration comprises the following:

- Operating system
- Middleware
- Application maintenance
- Administration of user accounts
- Configuration management
- Data backup

The complexity and cost of system administration depend on the size and especially the number of the included systems. For instance, running daily data backups is necessary for all servers that store important data, independent of the amount of data to be secured. In addition, user-account modifications must be updated on all servers accessed by the user. Clusters reduce the number of individual systems and thereby also the cost of system administration, by integrating a large number of applications, data and users into one computer system. One of the advantages is that a cluster system must contain only one set each of user accounts, data access authorisations, data backup rules, applications, data base managers, etc. Although the individual systems may vary due to the different cluster architectures used, it is normally more efficient to administer one cluster rather than the corresponding number of unconnected server systems.

The load distribution between the individual cluster systems is carried out automatically. Both Unix and OpenVMS clusters feature automatic load distribution and scalability, whereas Windows 2000 clusters provide greater availability only. With the Cluster server of the Windows 2000 Advanced server, two servers can be connected to one cluster, with the Windows 2000 Data Center server, four servers can be connected. The automatic IP load distribution, which is independent from the Cluster server, allows the distribution of Web applications on up to 32 systems.

High Performance Computing Cluster (HPCC)/Beowulf Cluster

The Beowulf cluster architecture in principle consists of various computing nodes and one or more access computers. The access computer is known as the front end or as server node. A user logs on, usually remotely, to this server node. Once they are in, they can then reserve and use any amount of computing nodes for his work (his computing task). To do this, the computing nodes do not need any input or output peripherals such as keyboard, mouse, or monitor. Access computers and computing nodes do not need any additional hardware; usually standard systems with x86 or Alpha processors are used. The systems can be connected using any network infrastructure. As a rule, a free operating system (Linux, FreeBSD etc.) is used.

A Beowulf cluster is a cost-effective solution for tasks which demand high computing performance. However, it is not an easy task to set up such a cluster, as there is no complete package available. The cluster must therefore be put together from different software components, depending on the requirements, and the applications must then be programmed in parallel. It is also possible though to use a cluster with non-parallel jobs as a computerfarm.

High-Availability Cluster with Linux FailSafe

The Linux FailSafe cluster software is a universal, freely configurable, and error-tolerant high-availability system, with which individual services can be redundantly set up. This software enables a service to be automatically or manually migrated to another node (server) in case of an error. An error is not necessarily due to defective hardware: Application errors, desolate processes, and critical system conditions are also recognised as such and handled accordingly. In most cases the individual nodes of the cluster must be able to access common data sections in order to take over the data in case of an error. A SCSI RAID or SAN architecture can be used here, which allows the servers to take over data sections in case of an error. If the data sections were successfully taken over, then the active cluster server can restart the services and make them available using the same IP. Thus, the downtime of a service can be calculated and even controlled (within certain limitations), since various criteria can be configured for the migration of services. With the corresponding redundant server systems, an availability of 99.99% can be achieved with this solution.

4. Storage buses

4.1 Introduction

Storage buses are understood to be peripheral buses permitting the connection of hard disks, tape drives and optical storage systems to the system bus (PCI, SBus and the like). Whereas there were initially a series of proprietary buses, which today are of no or only minor significance (e.g.: IPI, DSSI in the digital environment, SSA mainly from IBM), three standards have established themselves: ATA (IDE) as the cost-effective solution for internal storage systems, the SCSI bus primarily in the workstation sector and Fibre Channel, likewise using the SCSI protocol, for high performance. In addition, FireWire and USB (Universal Serial Bus), the general peripheral buses, are used primarily for workstations and notebooks, also for the connection of storage products by transfer to these buses with the aid of ATA adapters.

4.1.1 Parallel versus serial buses

Parallel buses initially promised to provide a high potential of additional bandwidth:
In comparison with serial connections, the throughput increases 8, 16 or even 32 times over. Thus the SCSI bus was initially defined with 8 bits, later with 16 bits. There were also non-realised plans with 32 bits. IDE was operated with 8-bit bandwidth from the outset.

Apart from acquiring new standards with higher performance by harnessing more intelligent protocols and enhanced transmission reliability, such were then secured by increasing the clock rate. However, this step-up in transmission speed is reaching its limits: The synchronisation of the parallel data signals is becoming ever more difficult. This is resulting in a dead end in the development of parallel buses, entailing that some present-day serial buses are faster than parallel ones. Future storage buses will therefore all be serial: ATA will become serial ATA (S-ATA) and SCSI will become serial SCSI.
Fibre channel was intended as a serial bus from the word go.

4.2 SCSI Interface

SCSI (Small Computer Systems Interface) is an I/O bus (in its original form with an 8-bit bandwidth, parallel, and later, serial defined I/O bus), which is widely used for the connection of mass storage drives of all types, and occasionally for scanners and other peripherals to various computer systems.
From the outset, the benefits of SCSI were a relatively high maximum transmission rate and flexible and easy configuration. SCSI devices incorporate a great deal of intelligence in the drives themselves, since SCSI is not actually a classical interface, but a peripheral bus, over

which the drives and the host adapter can communicate with each other. The first SCSI standard was passed in 1986 by ANSI. SCSI incorporates a three-layer system of rules, consisting of command, protocol and interface levels. The core of SCSI is formed by what are called SCSI Primary Commands (SPC). All devices must be able to process these commands. There are special commands based on these primary commands for individual groups of devices: Controller Commands (SCC) for host adapters, Block Commands (SBC) for hard disks, Stream Commands (SCC) for tape drives, Multimedia Commands (the so-called command level is made up of these commands). Below this command level is the protocol level which concerns the protocols for the various types of interface. These logical protocols control how a SCSI command appears on the respective interface, and how the individual devices communicate with one another. The classical SCSI interface (8-bit and 16-bit SCSI) uses the Interlocked Protocol (IPC). The new interfaces are based on their own protocols: The Serial Storage Protocol (SSP) for SSA as well as the Generic Packetized Protocol (GPP) to make access easier for suppliers of other interfaces.
Finally, the lowest level is the interface level in which the various physical transmission media are defined.The advantage of this modularity is manifested in the introduction of new serial technologies: Fibre channel and SSA can be integrated without difficulty, in what, without doubt, is the most widespread standard, namely SCSI. On the one hand, this saves expense and on the other, this speeds up the process of introducing new interfaces.
The extension of the original parallel SCSI (1 and 2) to serial transmission is called SCSI-3. SCSI-3 is downward compatible. When using the serial interfaces (fibre channel, SSA), substantially higher data transfer rates are achieved, and the division imposed with parallel SCSI between a faster data phase (20 MB/s with Ultra SCSI) and a slower command phase (5 MB/s) is no longer needed.
The following section describes in more detail the various protocols and interfaces.

Parallel SCSI interfaces

Parallel SCSI
- SCSI-1
- SCSI-2
- Ultra SCSI
- Ultra2 SCSI
- Ultra160 SCSI
- Ultra320 SCSI

Serial SCSI
- SSA
- Fibre Channel
- Fibre Channel Arbitrated Loop (FC-AL)

SCSI Interface	Data Bits	Max. Transferrate MB/s	Max. Number of Drives	Max. Cable Length (m)
Single-Ended	8	5	7	6
Single-Ended-Fast	8	10	7	3
Single-Ended-Ultra	8	20	7	1.5*
Single-Ended-Wide	16	20	15	3
Single-Ended-Wide-Ultra	16	40	15	1.5*
Differential	8	5	7	25
Differential-Fast	8	10	7	25
Differential-Ultra	8	20	7	12.5
Differential-Wide	16	20	15	25
Differential-Wide-Ultra	16	40	15	12.5
Ultra2 LVD	16	80	15	12
Ultra160	16	160	15	6
Ultra320	16	320	15	6

* 3 m for 4 drives

Parallel SCSI interfaces

4.2.1 Parallel SCSI

SCSI-1

The first accepted SCSI standard was SCSI-1. It supported primarily only hard disks and tape drives. Its other performance standards have rendered the SCSI-1 uncompetitive in today's market. The overhead for data transfer is approximately 90 percent. The transfer of data run in asynchronous mode reached rates of max. 1 MB/s, in synchronous mode 5 MB/s. Since 1990, SCSI-1 has no longer been used.

SCSI-2

The SCSI-2 standard was developed in 1986 and offers significant improvements. SCSI-2 features an expanded command set with improved support for hard disks, tape and MO drives, CD-ROMs, scanners and jukeboxes. In addition to the new standardised command set (Common Command Set), the SCSI-2 specification currently provides a higher maximum transfer rate of 10 MB/s (Fast SCSI, with the standard 8-bit bus), as well as the option of increasing the data path width to 16 bits (Wide SCSI). The combination of Fast SCSI and Wide SCSI provides data transfer rates of up to 20 MB/s on the bus. Not all devices that claim to support SCSI-2 automatically support the complete SCSI-2 specifications (including Fast SCSI and Wide SCSI) but generally, most of the newer drives use Fast SCSI. Some of the Wide SCSI-2 devices (16-bit data bus) have been misnamed SCSI-3 and non-Fast SCSI-2 devices have been called SCSI-1 standard, which can be misleading.

Ultra SCSI
The transfer clock rate was doubled again on the Ultra SCSI, a fully downward-compatible extension of the SCSI-2. By doubling the clock rate on the bus, the transfer rate is also doubled to 20 MB/s (8 bit) or 40 MB/s (16 bit). However, this only applies to the data transfer rate. Command transmission continues to take place at 5 MB/s. Plugs and cables were also kept for compatibility reasons. However, this increase in performance does have its price: The maximum permissible cable length is shortened to 1.5 metres in the case of single-ended and 12.5 metres in the case of differential cabling. An active termination is essential in order to ensure interference-free transmission.

To simplify the task of configuring the SCSI bus for the user, the Ultra SCSI specification contains the SCAM (SCSI Configuration Auto Matically) feature. In other words, the connected device configures itself. The implementation of SCAM in the devices, however, is left up to each manufacturer.

Ultra2 SCSI (LVD - Low Voltage Differential)
Products incorporating Ultra2 SCSI have been on the market since 1998. Once again the data transfer rate was doubled to 80 MB/s at 16 bit. The older 8-bit version will gradually fade away. Since further halving of the cable length for single-ended cabling would make the bus practically unusable, this standard was only defined as LVD (Low Voltage Differential) as a differential cabling which only works with one third of the signal voltage. The maximum cable length is still as much as 12 m.

To keep this standard compatible with the previous ones, Ultra2 drives are equipped with an autosensing interface, which enables operation with conventional single-ended controllers (but not differential), although only at Ultra SCSI speeds. However, they are not equipped with internal termination, which means that separate terminators must be used, even inside the computer.

Ultra160 SCSI
Ultra160 uses the same cabling as Ultra2 LVD SCSI. The transfer rate is doubled due to the data transfer on the increasing and decreasing flanks of the signal.

Ultra320 SCSI
The first products incorporating Ultra320 have been available since summer 2002. It is however debatable whether the bandwidth will be doubled again as serial standards such as fibre channel and serial ATA, later also serial SCSI, will increasingly threaten the survival of parallel SCSI. Even Ultra320 raises great technical difficulties for implementation in external storage solutions.

SCSI-3
SCSI-3 conforms to the standards of both parallel SCSI and those for serial interfaces described later.

SCSI transmission standards
Parallel SCSI interfaces are available in different versions, which have different methods of transferring data. SCSI signals are transmitted either on 8-bit (narrow) or 16-bit (Wide SCSI) buses. Up to seven drives can be attached to an 8-bit bus, and up to 15 drives to a 16-bit bus. Both bus widths are available for two cable types: Single-Ended (SE) or Differential (D) or Low Voltage Differential (LVD). SE SCSI signals are transmitted on a single wire, while D SCSI signals are transmitted on two twisted wires. The latter are therefore less susceptible to magnetic interference. The benefits of D SCSI include improved interference immunity and thus support of longer cable lengths.

The only drawback of D SCSI is that drives, host adapters and termination are more expensive than those for SE-SCSI.

When selecting a subsystem, it is important to take into account that the host adapter SCSI interface must match that of the drive. Technically, it is possible to connect 8-bit devices to a 16-bit bus; however, a number of special configuration rules must be followed.

Single-ended and differential or LVD SCSI cannot be operated on the bus simultaneously. Attempting to do so can result in damage to the drive and controller. Only LVD drives can adjust independently to single-ended.

Synchronous and Asynchronous Data Transfer
With parallel SCSI, data can be transferred in either asynchronous or the faster synchronous mode. When data is transferred asynchronously, each byte is sent and confirmed separately; whereas in synchronous mode, multiple bytes are sent and then confirmed together as a whole. Thus, the overhead is lower and the transfer rate higher in synchronous mode. Generally, all peripheral devices are able to work in asynchronous mode. Synchronous drives and controllers perform handshaking before data transfer, to determine whether or not the partner device is capable of transferring data synchronously. The best transfer mode is then automatically selected. Today's SCSI drives and host adapters generally support synchronous transfer mode.

Cable for Parallel SCSI
A few important factors should be considered when selecting the proper SCSI cable to ensure correct and error-free data transfer:

SCSI cables must comply with UL (Underwriter Laboratories) and CSA (Canadian Standard Association) standards. Individual cable strands should be made of copper braid, or better, tinned copper. They should be twisted in pairs, and then up to one meter of the entire cable bundle should be twisted once more. The entire cable should then be double shielded, which is usually done with a silver foil and an additional wire mesh over the cable bundle. If multiple devices are to be connected to the SCSI bus, the individual cables should be kept as short as possible, and optimally be all of the same length. This reduces the bus's susceptibility to interference.

Since Wide SCSI transmits over a 16-bit bus instead of an 8-bit bus, the normal 50-pin SCSI cable is not sufficient. Wide SCSI, therefore, uses 68-pin cables.
Due to the fact that the cables for single-ended and differential SCSI are identical, the same cables can be used for both types of interfaces.

Termination of the Subsystems
With growing data transfer rates, the demands on cable and SCSI bus termination are also increasing. Differential SCSI and active termination are becoming more and more important. In contrast to passive terminators, active terminators operate with an integrated voltage controller. They keep the terminator power line to exactly 2.85 V with active components. With passive termination, they are kept to about 3 V with a passive distributor. Depending on the cable type and length, the voltage in cables with passive termination can greatly vary.

The Forced Perfect Terminator (FPT) is a version of active bus termination mainly used in the IBM world (RS/6000). FPT terminators dynamically adjust SCSI bus impedance. With FPT, however, this type of termination must be used at both ends of the bus including the host adapter.
It is generally recommended to use active terminators, since they offer the SCSI bus better protection against electromagnetic interference compared with passive termination.

Connection of Subsystem
There are the following possibilities when connecting external SCSI subsystems:

1. Up to now no external subsystem has been connected to the SCSI port: In this case, the appropriate SCSI cable for the computer platform is required. Suitable cable exchange options for the corresponding computer platform are outlined in the price list section for all external subsystems. Thus instead of standard cables, cables with the plugs stated in the price list are available. The subsystem can be directly connected to each type of computer with the correct plug. Furthermore, a terminator is required. It is important that the maximum permissible cable length is not exceeded (always include internal cabling, that is, computers, disks etc.)

2. External subsystems are present and equipped with HD68 plugs: Using the supplied cable, the new subsystem can be added to or inserted at will between the existing subsystems.
3. External subsystems are present but are not equipped with HD68 plugs. In this case, there are two options:
- Connecting the new subsystem to the last existing box. This requires exchanging the standard cable for a cable fitted with the suitable plug. Furthermore, a corresponding terminator with plug is required.
- Connecting the new subsystem directly to the workstation. This requires exchanging the standard cable for a cable fitted with a plug suitable for the workstation. In addition, a further cable for connecting the new subsystem to the existing subsystems is required.

4.2.2 Serial SCSI

Since Ultra320 is presumably the last realisable parallel SCSI standard, work is currently being carried out on serial SCSI standards. They are intended to be more cost-effective than the fibre channel described below and hence will in future permit SCSI buses with higher transmission rates. First products are not expected before 2004 or 2005.

4.3 Fibre Channel

The name fibre channel is somewhat misleading, for the serial bus is specified not only for fibre-optic cables, but also for copper conductors. For both media, transfer rates of 12.5 MB/s to approx. 400 MB/s are intended. The physical transfer rates are a little higher with 132 Mbit/s to 2 Gbit/s. The resulting slightly higher gross transfer rate is used for an 8 to 10-bit code which allows simple error recognition. Various cable types can be mixed in a fibre channel system. The simplest version is a shielded TwistedPair cable. Large distances or high transfer rates require copper or fibre-optic cables. Thus, with one and the same interface, both low-end and low-cost systems, and high-end systems can be implemented. The fibre-optic connections are made via a Duplex SC or LC plug, the copper cables via a DSub 9-pin plug or via a 9-pin HSSDC (High Speed Serial DC) plug.

Fibre channel is the general term for a standard array that was developed and is being developed further by ANSI (American National Standards Institute) to establish new protocols for a flexible information transfer. This development began in 1988, as an extension to the standard Intelligent Peripheral Interface (IPI) Enhanced Physical and branched into several directions.

The principal goals of this development are:
- To support different types of physical interfaces
- To offer a means for networking these different interface types
- To enable high-speed transmission of large data volumes
- The logical protocol is transported by the physical interface; this permits the transport of various protocols via a common physical interface (optionally simultaneously)
- Relief for the increasing number of physical interfaces with limited circulation.

Fibre channel can be found in systems of varying sizes, but at the moment, it is first of all applied in mainframe computers and will later be applied in workstations. It will not be used in cheap PCs, though later it will be implemented in desktop systems of professional users that are linked to a network.

At present, system-specific fibre-optic interfaces can be found in some desktop systems and workstations. The use of very fast fibre channel interfaces will increase as soon as their speed and their function become familiar among these users. This includes small workgroups that have to be networked in order to obtain data transfer with high-speed performance.

Two basic peripheral protocols for system communication can be named: Channels and networks. Usually, the term channel refers to the peripheral I/O interface (with a host computer) that transports large amounts of data between host and peripheral system. The system processing input is kept as low as possible by using minimal or no software during data transfer on hardware level as soon as an I/O operation sets in. The term network, in contrast, refers to an I/O interface that usually implies numerous small transfers with larger system processing input which traces back to an information flow with software participation. As a rule, networks support a host-to-host communication.

Channels

As a rule, channels operate in a closed, structured and foreseeable environment in which all devices capable of communicating with a host are known in advance and every change also requires modifications in the host software or in the configuration tables. Most channels can manage these more complex knowledge levels.

The host system comprises all of the knowledge in the channels connected to this host. Occasionally, this is also described as a "master-slave environment". Peripherals such as magnetic tape drives and disk drives as well as printers are connected directly to the host system. Here, the host is the master, and the peripheral devices are the slaves.

Channels are used to transmit data. "Data" refers to files with information, which may contain thousands of bytes. An important requirement for data transmission is error-free delivery, in which transmission delay time is of secondary importance.

Networks

On the other hand, networks operate in an open, unstructured and basically unpredictable environment. Nearly every host or device can communicate with every other device at any time. This situation demands more intensive software support for checking access authorisation, setting up transmission sessions and routing transactions to the correct software service.

This unstructured environment, which assumes that the connected devices all have equal access, is called a "peer-to-peer environment". Several workstations and mainframe computers can be networked. In the process, each system is independent of the others; from time to time, they exchange information with the help of network protocols. A workstation and a mainframe have equal access in comparison to other systems of the same type. In this sense, this environment is comparable to the manner in which the telephone system operates, where all telephones have equal access. For this reason, analogies are often made to the telephone system.

Networks are not only used for error-free data transmission, but also for speech and for video transmission; here, timely delivery has top priority, with error-free delivery playing a secondary role. If, for instance, delivery of a video transmission is late, the data will be of no use; if, on the other hand, one or two pixels are lost en route, this will not be noticed as long as the image does not flicker.

Supported Protocols

The fibre channel seeks to combine the best aspects of each of these opposing communication processes in a new I/O interface, which meet the requirements of both channel and network users.

The fibre channel supports transmission of ATM (Asynchronous Transfer Mode), IEEE 802 and other network traffic. All users who are familiar with the Internet Protocol (IP), e-mail, file transfer, remote logons and other Internet services will see that fibre channel supports these protocols with higher speeds.

These are important aspects for connecting systems which work on a fibre channel basis to the most important global networks as well as to LANs which have already been companyinstalled. This includes the SONET-based system and LANs such as Ethernet.

A major contribution of the fibre channel is that both interface types, that is, channels and networks, can now share the same physical medium. Over the past two years, I/O channels have been extended to include network applications (e.g. with the help of SCSI, to network two workstations). In the same way, with the help of network data transfer protocols, networks can move data back and forth between systems and file servers (e.g. Network File System, or NFS).

With the fibre channel, it is now possible to use the same physical medium and physical transport protocol by means of a common hardware port in order to manage both channel and network activities. It is possible to send information to a network connected to the backplane of a workstation via fibre channel and at the same time to use the fibre channel for in-house communication with local peripherals (e.g. with hard disk and magnetic tape drives).

Fibre channel protocol features: Fibre channel does not have an instruction set such as SCSI and IPI, but rather provides a mechanism for placing other protocols on fibre channel. This is possible because fibre channel serves as a carrier for these instruction sets and in this way the recipient can distinguish between the two. This implies that various instruction sets for older I/O interfaces, for which software outlay was previously necessary, can be directly applied on fibre channel.

Separation of I/O operations from the physical I/O interface is a significant fibre channel feature, and it enables the simultaneous use of different instruction sets. The various instruction sets, such as SCSI, IPI-3, IP etc., are usually used on their own special interfaces. Fibre channel, however, defines a single, physical transmission mechanism for these instruction sets.

Fibre Channel

- Is not aware of the content or significance of the information which has just been transmitted
- Increases connectivity from dozens to hundreds or even thousands of devices
- Increases the maximum distance between devices
- Increases the transmission rate four or fivefold in comparison with the most widespread channels and a hundred fold in comparison with the most common networks.

The following sections describe how fibre channel permits network setup.

Networking Topologies
Fibre channel devices are also called nodes; each has at least one port to permit access to the outside world (in other words, to another node). The components, which link two or more ports with one another, are together described as the "topology". All fibre channel systems have just these two elements: Nodes with ports and topologies.

Each fibre channel port uses a wire pair - one wire to transmit information to the port, and one wire to transmit information leaving the port. Fibre channels have either electric wires or optical wave guides. This fibre pair is called a link and is a component of every topology. Data is always transmitted in units (so-called frames) by means of these links. The fibre channel standard defines three topologies; however, stress is placed above all on a topology which is also known as the fabric system. It is this topology which will be described first.

Fabric Topology
A fabric system permits dynamic connection between nodes by means of the ports which are connected to this system. Please keep in mind that this application of the term "Fabric" may also be used as a synonym for the terms "Switch" or "Router". Each port in a node, a so-called N port or NL port, is connected to the fabric system by means of a link. Each port in a fabric system is called an F port. With the help of the fabric system's services, every node is able to communicate with every other node connected to other F ports in the same fabric system. With this type of topology, all frame routing operations are carried out by the fabric system instead of the ports.
This peer-to-peer service is an essential component of the fibre channel design. A system designed for peer-to-peer services can be utilised in such a way that the host-type master-slave communication process is emulated. In this way, fibre channel is able to simultaneously support both channel and network protocols.

Like a Telephone System
The function of a fabric system can be compared to that of a telephone system - we dial a number, the telephone system finds the path to the requested phone number, the phone rings and the person being called answers. If a switch or link crashes, the telephone company routes the calls via other paths, which the caller rarely notices. Most of us are unaware of the intermediate links which the telephone company employs in order to make our simple phone call a success.

However, we do provide the telephone company with some information regarding our phone call. For example, the telephone number begins (in the U.S.A.) with the digit "1", followed by ten digits in the form of an area code (3), switches (3) and the subscriber number (4). If a "1" is not used at the beginning of the telephone number, the caller is making a local call, and

only seven digits will be used. This information assists the telephone company in setting up the connection. The telephone number corresponds to the fibre channel address ID. Part of the address ID is used in order to specify the fabric system domain which will be affected; the rest is used to determine the special port.

It should be remembered that the telephone system is not involved in the content of the discussion between the two participants in the phone call (or affected by it); it is only responsible for establishing the connection. In the same way, fibre channel is responsible for establishing the link, and the protocols which are placed on it (e.g. SCSI or IPI) carry the instructions. These protocols play a role similar to that of languages in telephone systems. Fibre channel and other protocols should be seen as integral components in the exchange of information.

The fabric system may consist of a single or several fabric elements. Just as with the telephone system, we don't know (or don't pay attention to the question of) how many switches we must go through as long as we are connected with the correct destination station.

A fabric system is also called a switched topology or crosspoint switching topology. Routing by means of various switches takes place: The fabric elements interpret the destination ID in the frame as soon as it arrives in each fabric element.

The fabric system can be physically implemented as a single fabric element with several F_Ports or it can be implemented as a series of several of these fabric elements which are linked with one another. Routing or switching of each connection is transparent for both N_Ports, which are connected to the fabric outside edge by means of F_Ports.

If the topology is disconnected from the nodes, as is the case with the telephone system and fibre channel, new technologies can be introduced for the router. New speeds and new functions can be implemented in the fabric system without losing all previous investment in existing nodes. Fibre channel permits a combination of accessory devices with varying speeds or capabilities.

Other Topologies

In addition to the fabric system topology, the fibre channel standard defines two other topologies. One is called "point-to-point topology" with just two connected ports. In this case, routing does not take place. The third topology is called "arbitrated loop". This is a simple and inexpensive topology for connecting several dozen NL_Ports on one loop. The ports in an arbitrated loop topology, which are called NL_Ports and FL_Ports, are slightly different from N_Ports and F_Ports. They contain all the functions of N_Ports and F_Ports

and are capable of working properly in a fabric system. An FL_Port is a port in a fabric system which processes the arbitrated loop protocol.

With arbitrated loop topology, each port sees each message (as in the token ring protocol) and passes over and ignores those messages which have no token acquisition protocol.

The telephone system analogy will be continued below in order for the function of a fabric system topology to be more easily comprehensible: You dial a friend's telephone number. In order to do so, you do not need to know the exact route the telephone system takes to reach your friend's house when you call. Routing is the responsibility of the telephone system. Fibre channel's fabric system has the same function: You enter a destination address, and the fabric system routes the data to the destination N_Port.

If you dial an incorrect telephone number, the telephone company will notify you that this number is unobtainable. Similarly, the fabric system rejects frames for invalid destinations.

Just as the telephone company can configure numerous routes between various points and also does this in order to provide reliable service, a fabric system can have numerous paths between fabric elements in order to handle traffic. This also facilitates provision of standby paths for the event that an element or link breaks down.

The fabric system and arbitrated loop topologies from fibre channel can be merged with each other in one system in order to add a variety of service grades and performance rates to the node. In addition, the fabric system can use other networks like SONET or ATM over SONET among fabric elements, in order to bridge distances between nodes that are too big to be mastered by the connection between N_Ports. These special connections can exist among fabric elements that are spread out over a larger geographical area and that are not directly connected to the node.

The capability to add other types of connections among fabric elements, called extension ports or E_Ports, means an increase in value of every disk drive that is connected to the fabric system. Particular attributes of fibre channel and fabric systems allow ports with varying speeds and media types a communication over short or long distances if a fabric system is available.

Within the fabric system itself, improvements of technology can be implemented without having to change the N_Ports in any way. The major proportion of the benefit of the new technology is indirectly passed on to the nodes due to higher speed, since speed, reliability, and the distance of communication has increased within the fabric system.

How many N_Ports can be implemented? The fabric system is solely limited by the numbers of the N_Ports that are displayed in the target address field in the header of the frame. This limitation amounts to slightly over 16 million ports that can be logged-on at the same time to one fabric system with a 24-bit address ID. This should cover all requirements for individual integrated systems for some time.

In the fabric system topology, the address ID is divided into three parts: Domain (8 bit), region (8 bit), and port (8 bit), which comes to 24 bit altogether. These components are comparable to a telephone number with area code, operator's and subscriber's number.

Some topics naturally belong together, which also holds true for fibre channel. Aspects that deal with establishing a reliable and testable connection for the fibre-optics have not much to

do with questions concerning problems that appeared because of a lost frame. These different fields of interests are called function layers according to the fibre channel standard. Five layers are defined by this standard; each is labelled FC-x.

FC-0
Defines the physical fibre channel proportions, including media type, connections as well as electrical and optical performance features necessary for the port connections. This layer is written in FC-PH norm.

FC-1
Defines the transfer protocol, including the 8-bit/10-bit coding, word transfer order, and error registration. This layer is written in FC-PH norm.

FC-2
Defines the signalling and framing protocol, including the frame layout, contents of the frame header, and application rules. Additionally, it defines particular protocol-independent frames and protocols like the user login. The FC-PH norm consists predominantly of writing this layer.

FC-3
Defines commonly used services that can be available at different ports in one node. For this layer there is no norm.

FC-4
Defines the mapping between the lower fibre channel layers and the command sets that use fibre channel. Here you can find SCSI, IPI-3, HIPPI, and SBCS. Each command set contains an extra norm, so that a third party is not occupied with unnecessary information foreign to the

system. If you work with SCSI-3 you are probably not interested in IPI-3 or HIPPI.

On layer FC-4, a node cannot take on all the different options that are acceptable by the norm. Each node can implement one or more services. On layer FC-0, solely speed and media type options (described below) can be implemented. Each individual port in a node can implement different combinations of speed and media. Layer FC-2s also supports numerous options from which an individual manufacturer must be chosen. Some industrial groups are already working on defining a profile that specifies the operating environment which is necessary for the application (e.g. SCSI, HIPPI, IPI-3 etc.). On layer FC-1 there are no options.

When developing a product or planning additional expansion options, FC-4 should be considered, so that a master port request list is available when the expansions are made. A system provider that integrated the IPI-3 protocol and plans to install the Internet Protocol (IP) later should carefully consider both FC-4s before he decides on a port design, since the requirements are different.

Fibre Channel Arbitrated Loop (FC-AL)
A slimmed-down version of the fibre channel is the fibre channel AL. AL means Arbitrated Loop and describes the topology of this type of fibre channel designed for local devices and disk arrays. A maximum of 127 loop ports (NL ports) are arranged as a ring. Data exchange is only possible as a point-to-point link. Each data packet reaches the device first via the read port, which checks whether it should process the information. If not, it resends it via the write port. In order to initiate this data transfer, the device must first obtain control of the bus. Parallel data exchange between several devices (as with the general fibre channel definition) is not possible.

In order to be able to handle disc arrays better, the FC-AL supports the backplane architecture, as well as the normal cable link. The hard disks are connected to the backplane via a 40-pin SCA (Single Connector Attachment) plug, which includes both data and power supply lines. If there is no drive in one of the bays, the backplane logic bypasses the empty slot, and the circuit remains closed. A further task of the backplane is the automatic configuration of the drive and also to support the hot-plug function, which is the exchanging of a drive during operation. The same principle is also used by fibre channel hubs. In a fibre channel loop, a device failure or cable defect will break the circuit, blocking the entire bus, so the hub bypasses any port that is not in use or blocked by interference. Consequently, the flow of data to other devices is not interrupted and the bus continues to operate normally.
FC-AL products have been available since Autumn 1996. Initially they were limited to high-end RAID systems, but due to increasing storage needs they are now also playing an increasing role in smaller servers.

4.4 SSA

Serial Storage Architecture (SSA) is a high-performance interface that combines I/O devices of all platforms.
This serial bus-like interface was developed by IBM based on IBM-9333 technology, and is predominantly used for connecting hard disks. With this, up to 128 devices can be connected together.

Like the fibre channel, SSA is a point-to-point connection, but two write and two read channels are available. On each channel (read/write), a maximum transfer rate of 20 MB/s is possible, which corresponds to a cumulated transfer rate of 80 MB. However, this rate can only be achieved if the read/write ratio is 1:1 and the host adapter accesses data located on at least four disks.
A standard TwistedPair cable is sufficient for distances of up to 20 m between two devices to be connected. Fibre-optic cables are used in lengths of up to 680 m. Only 6% of the data carried over SSA is used for control or routing functions, that is, 94% of the data is user data. However, only a few manufacturers support SSA (like IBM), and has meanwhile been almost entirely superseded by Fibre Channel Arbitrated Loop technology.

4.5 ATA (IDE) Interface

The Integrated Drive Electronics (IDE) interface is very widespread in the PC environment. In fact, it has now completely replaced the older ST506 interfaces. Up to two hard disks can be connected to each IDE interface. If two IDE hard disks are to be operated, one disk is configured as the master, and the other disk as the slave. Note that as standard, IDE hard disks are configured as a master drive and not as a slave.
The original IDE interface, also called AT Bus or ATA (AT Attachment), could theoretically transfer a maximum of 4.3 MB/s. However, in the real world, only 2 MB/s are reached.
It is used solely for the connection of hard disks and supports no other peripherals. A number of disk manufacturers have improved the IDE interface in order to meet the increased performance requirements.

Fast ATA and Enhanced IDE
The differences between these two developments lie primarily in the marketing strategy, while the implementations and functions are for the most part identical. Both Fast ATA as well as Enhanced IDE remain compatible with the old IDE adapters and hard disks. They use the same 40-pin cable.
For Enhanced IDE, there are ATAPI (AT Attachment Packed Interface) extensions, which - as for SCSI - allow peripherals like CD-ROM, tape drives, scanners etc. to be connected.

UltraATA
Much like Ultra SCSI, UltraATA increases the clock rate on the bus of the Fast ATA interface. Therefore with UltraATA a data transfer rate of 33 MB/s can be achieved on the bus.

UltraATA/66, UltraATA/100, UltraATA/133
After three further increases, the 133 MHz rate has been reached but the cabling requirement has become more stringent since ATA100 due to the increased clock rate: A special 80-pin cable is needed. However, because the connector remains unchanged, older disks and controllers can still be combined with the new standards, of course at the cost of reduced performance.

S-ATA
ATA technologies have gained a surprising lead over SCSI during the past few years (as a result of the cheaper ATA prices) and the new standard will probably contribute to strengthening this trend. Again, it is a serial technology, rather than a parallel, that is gaining in popularity: Serial ATA (S-ATA) provides a bus that offers a high growth potential for the transfer rate. It can also manage a spectrum of features from the SCSI world such as command queueing. This is why S-ATA will be used for applications with high I/O requirements - and thus offer high performance in RAID systems.
The first larger applications will be RAID systems, as initial problems are predicted with the launch of S-ATA in PCs: A simultaneous use of serial and parallel ATA is not worthwhile; on the other hand, there will be a long wait for S-ATA CD-ROM drives - due to the high costs.

4.6 USB Interface

USB (Universal Serial Bus) is an interface standard for connecting external devices (for example, printers, modems, keyboards, monitors, digital cameras, etc.) to PCs via a bus system. The advantage of a USB interface is the trouble-free integration of additional periphery. The USB controller recognises if additional periphery has been connected, and automatically installs the required drivers and prepares them for operation. Complicated new configurations of the PC systems are a thing of the past. Theoretically, up to 127 devices can be connected to the USB interface. The transition to USB version 2 is currently in course, with higher transfer rates and some extensions of the protocol.
Detailed information on USB can be found at www.usb.org.

4.7 FireWire 1394

FireWire (IEEE 1394) is a high-speed, serial data bus with maximal 800 Mbit/s which, originally used by Apple, is used in particular to connect multimedia devices such as camcorders and digital cameras to the computer. The standard was defined in 1995. The bus enables reserved bandwidths, which is particularly required in the area of video. 16 devices can be connected in series over a length of 72 m (4.5 m per link). However, there are also bridges which increase the number of devices drastically by cascading (a maximum of 1023 bridges are permitted). FireWire uses a 6-pin special cable that is made from two shielded TwistedPair wires and two additional wires for the power supply.
FireWire is currently of relatively great significance only in the Apple environment.

5. Hard Disks and RAID

5.1 Introduction

We are talking about so-called primary data memories here. This is a mass storage which makes data available by random access (direct access to information blocks). This category includes semiconductor memories and magnetic hard disk drives, e.g. hard disks and floppy disk drives.

In order to manage large data volumes exceeding the capacity of a hard disk and also to ensure increased fault tolerance, hard disks are combined to form RAIDs (Redundant Arrays of Independent Disks).

5.2 Hard Disk Drives

Hard disk drives (also known as Winchester disks or magnetic disks) are the mass storage devices used for most standard applications. They use a magnetic recording procedure and can be read from and written to as often as necessary. Hard disks are differentiated according to capacity, storage bus interface and form factor. Added to these are performance features such as rotational speed and access time.

5.2.1 Capacity

Mass storage capacity is measured in megabyte, with a differentiation between gross and net capacities. The gross capacity refers to the maximum theoretical capacity of the drive, calculated by multiplying the bit density of a track by the track's length, the number of tracks and the number of disk surfaces. The net capacity is the total amount of storage space that is actually available.

The figure listed by the manufacturer as the net capacity differs considerably from the amount of net capacity actually available to the user. This is partly because 10 to 20 percent of the capacity is used between the blocks for error correction and address information. Another reason is that manufacturers of hard disks and operating systems have not agreed on the definition of a Megabyte. Disk manufacturers calculate 1 MB = 1,000 x 1,000 Byte (= 1,000,000 Bytes), while most operating systems calculate 1 MB = 1,024 x 1,024 Byte (= 1,048,576 Bytes or characters). Furthermore, the file system requires some space for administration data (for example, SunOS reserves an additional 10% of a file system as auxiliary storage, and does not display this to the user for the df command). The net capacity is also dependent upon the formatting. It is not absolute, but can vary within certain parameters.

The required amount of mass storage depends on the application. Audio/Video (A/V) applications, graphic programmes and publishing programmes have to handle particularly high volumes of data. In addition, the operating system and application programmes require a lot of disk space. As a result, a disk that originally seemed quite large, can quickly fill up with applications and data. The disk space required is currently expected to double annually.
The hard disk capacity should be generously planned as more space is usually needed than originally assumed. One should also consider that the performance suffers when the capacity limits are challenged. The constant decrease in hard disk prices, however, is a reason for not purchasing much more than is initially needed. Purchasing a smaller disk with little reserve capacity and later upgrading to a larger disk could prove to be more economical than initially investing in a larger disk, but only if the time between the initial purchase and the upgrade is long enough. Thus a hard disk is recommended that is large enough to include the needs of the following year.

Some operating systems, however, impose limits on the size of disk that can be used.
It is important to note some limitations with HP9000 workstations: Only hard disks < 2,147 MB can be used as boot disks with the 7xx models under HP-UX 9.01 (with Patch PHKL_3325) and HP-UX 9.03 - 9.05 (or to be more precise, only file systems < $2^{31}$are recognised); hard disks of up to 4,294 MB are permissible as data disk (< 2^{32}). The HP9000/7xx models running HPUX 10.xx or above, and HP9000/8xx under HP-UX 9.xx, all possess a Logical Volume Manager (LVM). This allows boot disks of up to 4,294 MB, and data disks of unlimited size (or only limited by the number of possible file systems per SCSI ID). If a file system larger than 4,294 MB is to be addressed, the LVM simply divides it into several logical files, and thus evades the magic limit of 2^{32}.
For the HP9000/3xx and 4xx models under HP-UX, neither the boot disk nor the data disk may be larger than 2,147 MB. The same applies under DOMAIN-OS. Here, the hard disks must be additionally converted to SCSI-1.

Capacity restrictions also exist under DOS, Win3.1, Win3.11, and Windows95. Each disk can be configured with only one primary partition. In the system, as a whole, there can only be one active partition, the active partition usually being the boot partition. Unfortunately, the maximum permissible size of a partition is 2 GB, the maximum disk size is 8 GB. Please note, however, that the block size with a 2 GB partition is 32 KB, i.e. a 250-byte file still occupies a block of 32 KB. This can only be avoided by configuring a block size of 8 KB, which is automatically the case when using a 500 MB partition. The maximum size of the file system with Windows NT and Windows 2000 is 16 exabytes (16 million terabytes). The only limiting factor regarding disk size is the SCSI controller. A 20 GB partition can easily be configured.
With Sun Solaris1.x (not Solaris2.x), the file system is limited to 2 GB. Since a maximum of

seven data partitions can be created per disk, the disk size is limited to 14 GB. AIX 3.x from IBM likewise restricts file systems to 2 GB. Since it is possible to partition the hard disk almost at will using the AIX Logical Volume Manager (LVM), disk size is practically unlimited. AIX 4.x supports file systems with a maximum size of 64 GB, enabling larger disks or RAID systems to be addressed as a single file system.

Some older operating systems (IBM AIX3.5, SGI IRIX5.x, Sun Solaris1.x) interrupt the formatting of hard disks larger than 8 GB with a timeout error. Before buying a new hard disk, users should ensure that in such cases a suitable patch is available.

5.2.2 Storage Bus Interfaces

Hard disks are currently produced with ATA, SCSI and fibre channel interface. While ATA or SCSI (rarely FC) is used internally in computers, only SCSI or FC are used externally.
More on the differences of the interfaces is to be found in the Storage Buses chapter. ATA hard disks are being increasingly used within RAID systems (see below).

5.2.3 Form Factor

The form factor refers to the diameter of a medium (the disk inside the drive box). Also inside the box are the read/write heads, the mechanics for moving the heads (the actuator), the spindle motor, and the logics (electronics).

Currently, hard disks are available in 2.5" and 3.5" form factors. Smaller disks with 1.8" and 1.3" form factors are also available. Hard disks are offered as full height (82 mm), half height (41 mm) and low profile (25.4 mm = 1") drives, and in 19 mm and 12.5 mm heights.
The 3.5" versions are probably the most popular and have a net capacity of approx. 300 GB. The larger 5.25" drives will no longer be developed.

5.2.4 Performance Features

Apart from the processor speed and the size of the computer's main memory, the performance of the primary mass storage is the most important factor for the computer's response time. Data transferred from the hard disk must fit completely into the main memory, otherwise data must be reloaded. With ever increasing amounts of data that are transferred over the bus, the main memory should correspondingly be increased. The performance is primarily influenced by the time required to access the data and the data transfer rate. The access time is calculated from the time required for multiple sequential processes:

When a program wants to read or write data, the operating system generates a request which is passed to the hard disk controller. The time required for this process is called the operating system overhead. This request causes the drive to position the read/write heads at a specific point on the disk surface. This time is called seek time. The value given is always the mean seek time, i.e. the time required for the head mechanism to travel 1/3 of the maximum distance across the disk. The settling time specifies how much time the head mechanism requires in order to mechanically rest after a head movement or for oscillations to be extinguished. With modern drives, the seek time already includes the settling time. If the read/write head is positioned on the correct track, it must wait until the desired block is reached.
This wait time is called latency or rotational wait time. The time required for 1/2 of a revolution is always specified. This figure is exclusively dependent upon the disk's rotational speed. So,for example, 7200 rpm disks have a latency of 4.17 ms. At this point, data is actually transferred from the hard disk to the hard disk controller. At this stage the disk's (internal) transfer rate becomes important. It indicates the number of bytes transferred per second. Sometimes a frequency is given. Dividing the frequency by 8 will convert it into the transfer rate. The higher a disk's transfer rate, the less transfer time is required for the data transfer. Once the data has arrived at the hard disk controller, it is then sent via the SCSI bus to the SCSI host adapter. The transfer rates via the SCSI bus are up to 320 MB/s. For more details on SCSI buses, see the chapter on Storage Buses.

The shorter the total time, the better the system performance. This is especially true for operating systems that cannot process other tasks during the latency period, e.g. MS DOS. When comparing transfer rates of hard disks in tables, one should note that these figures are calculated under optimal conditions, i.e. the head is already positioned, a complete cylinder can be read, etc. Block size (i.e. how many bytes are grouped into a sector) also influences the transfer rate. Though 512 bytes/sector is most common, disks can be formatted with as many as 4,096 bytes/sector. Larger sectors leave more space for user data, because less space is required for storing formatting information. This consequently leads to a higher transfer rate. This is, however, not possible with all operating systems, nor is it appropriate for all applications. The smaller the read/write files are, the smaller the sector size should be (e.g. databases). The larger the files are, the more appropriate large sectors are (e.g. for graphics processing).
What is more, for SCSI drives the transfer rates are often indicated through the SCSI bus. But since SCSI drives have a built-in cache, this transfer rate only reflects the path from the disk cache to the host adapter cache. The actual transfer rate from the disk surface into the host adapter's cache is significantly smaller.

Most applications only read and write a few blocks, thus generating low-volume transfers per request. Therefore, the rate at which the hard disk transfers data to the host adapter is not of primary importance. The more segmented the data on a hard disk, the shorter the read access required. One should ensure, however, that the transfer rates of the components in the data path match each other; hard disk - hard disk controller - SCSI host controller - computer RAM.
Another means of increasing performance is by using controllers with cache memory (on the hard disk as well as on the SCSI host adapter).
A multi-segment cache is especially advantageous. The cache is dynamically administered and divided into several areas. Data is stored in the cache, depending on which area of the disk it came from. Accessing data in the cache eliminates delays caused by mechanical components.

These controllers process write requests like non-caching controllers: Each write block is transferred to the connected disk. Additionally, the controller notes the block's address and contents in its internal memory, the cache. The advantage lies in the reduced number of hard disk access. Caching hard disk controllers are recommended for the following situations: Computers with minimal or barely sufficient RAM, with hard disk drives which process a large number of read/write requests per unit of time, and disks using small numbers of blocks per request e.g. system disks, for databases, and in general for applications which often alternate reading and writing to the same files. Caching controllers can reduce the average seek time to under a millisecond.

The different types of cache are read, read-ahead and write-back cache. A read access to a read cache causes the cache memory to be searched for the requested block. If it is found, it is immediately transferred. There is no delay for disk positioning and rotation. If the block is not found in the cache, it is retrieved from the disk.

When a read-ahead cache reads a requested block, it also reads the subsequent few blocks and stores them all in cache memory. A write-back cache stores blocks that need to be written to the disk until there are no more competing read or write requests. This has the effect of speeding up write requests for the user. One danger however, is that if there should be a power failure, all of the data stored in cache would be lost. This is not an issue with a read cache, but it could destroy the file structure of a system using a write cache.
Thus, it is imperative that systems using a write cache also have an uninterruptible power supply (UPS).

Hit rates of over 90 percent can be achieved, depending on the size of the cache buffer, the cache logic (i.e. which blocks are overwritten when the cache is full) and the type of application. This provides an enormous improvement in throughput.

Many operating systems automatically reserve a part of the main memory for cache, such as Windows NT, NetWare, Unix and OS/2. More main memory can radically improve the total system performance.

5.3 Semiconductor Disk Storage (RAM Storage)

If disk drives cannot deliver the data for a given application quickly enough, a semiconductor disk storage (RAM disk or solid state disk) can be used. The RAM disk uses special software to allocate RAM from main memory, which is then organised as a virtual disk. A solid state disk is a hardware solution which consists of a bus interface and memory modules. Both types of memory provide disk-like access. As with a real disk, files can be stored and subsequently read.

Advantage: The seek time and rotational wait time do not apply; thus, data is transferred with the maximum possible speed.

Disadvantage: These types of storage are not suitable for permanent data storage (even with a battery backup). They are also considerably more expensive than disk drives of the same capacity. A battery backup is always required for this type of storage, as even the shortest power loss could result in total data loss.

5.4 Floppy Disk Drives

Disk technology is based on a removable, rotating magnetic disk with an access time of 150-250 msec and has the advantage of being widely used for the exchange with other systems. Its limited capacity of 360 KB to 2.88 MB is a disadvantage. It is no longer sufficient now that powerful PCs have become common not only in companies, but also for private use.
This disadvantage has been largely removed by the ZIP drive from IOMEGA. At 750 MB per storage medium, the capacity. This drive also features an attractive price/performance ratio.

Since the middle of 1999, a HiFD floppy disk drive has been available from Sony, which supports 200 MB of storage capacity. It is compatible with the drives that are currently used and can read the 1.44 MB disk.

However, the small distance that the magnetic head maintains from the storage disk can lead to a head crash. This type of contact usually results in the loss of the data stored at the location of the crash. A head crash can also occur in the removable disk drives described below.

5.5 Removable Disk Drives

Large amounts of data can be managed with removable disks, without losing time in transferring data from one medium to another. In addition, such subsystems are well suited for high-security situations, for example when data needs to be stored in a safe overnight. The use of removable disks also makes sense when the location where data is used is constantly changing, as in the case of a programmer, for example, who is active at several locations and always needs access to his project data. Since this method continues to store data on a disk drive, where the danger of data loss remains, an additional data backup on another medium should be provided for.
Whereas removable disk systems were produced as independent products in the past (Iomega, Syquest), today standard hard disks (often with 2.5" form factor) in appropriate canisters are used, since the transport safety has significantly improved through new technologies, particularly with the influence of notebooks.

5.6 RAID

CPU and system memory power roughly doubles every year. In comparison, it takes about two years to double the capacity of mass storage devices and seven years to reduce the average seek times by one half. Conventional disk technology evolves at a markedly slower pace than CPU technology.
Therefore, techniques are needed which enable the more slowly evolving disk technology to keep pace with the processing speed and large data amounts of current systems - by means of increased access speed and increased capacity.
Typical MTBF times (MTBF = Mean Time Between Failure) of drives today are over 500,000 hours. When using 100 disk drives with a single host, a failure is likely to occur once every six months. Even if the level of disk reliability were to continue to improve over the next few years, this situation would still be unacceptable. Therefore, systems that are fault tolerant are required.

RAID is the generic term used to describe technologies where several disk drives are grouped together into larger systems, often with a higher fault tolerance. RAID stands for Redundant Array of Independent Disks, where "inexpensive" is often read in place of "independent".
A storage solution is referred to as redundant = "fault-tolerant" if the failure of a single drive doesn't lead to the failure of the whole system or to a disruption of operation or even to a loss of data. RAID uses common processes such as mirroring (a system's single controller writes identical data to two disks simultaneously), duplexing (one system writes to two disks using two separate controllers), and striping (data blocks are distributed evenly across a number of drives grouped together as one logical unit).

Different levels of RAID are being discussed today. The term "level" is actually misleading because RAID level 1 does not actually build upon RAID level 0. "Type" would be a more appropriate term.
Before going much deeper into the different RAID levels, however, a few basic remarks on accessing disks are called for since the performance varies significantly with the application, when moving from a single-disk drive strategy to RAID storage techniques:
Primary factors include the proportion of read vs. write requests, as well as the average time required for individual data transfers.
In a single-disk drive system, a read request takes just as long to process as a write request (caches, read-ahead buffers, and other techniques which increase performance will be disregarded for the question under consideration). This is not the case with RAID, which also differs from level to level. There are also differences in the processing speed of short requests (typically one or only a few blocks) and long requests (usually several hundred blocks). The first type of request is typical in database applications, transaction processing systems and commercial multi-user applications, whereas the second is characteristic of graphic applications, super computing, and technical/scientific data recording. Each of the RAID concepts discussed below differs from the others in terms of processing speed for requests of different lengths.
RAID is not the answer to all disk throughput problems, nor is it faster than using conventional single-disk configurations. When using one of the RAID types described below, the optimisation of data distribution is an important factor. As a rule, a relatively small number of data is the most frequently accessed, while the majority of data is only rarely accessed. The best solution is a hierarchical data storage system incorporating semiconductor storage for continually used data (e.g. RAM disks); very fast mass storage devices for frequently accessed data (possibly mirrored); relatively fast, large mass storage devices for less frequently accessed data (e.g. RAID); and archiving storage systems for infrequently used data (e.g. optical disks with automatic changer).

5.6.1 RAID level

There are different ways of distributing data on several hard disks to achieve maximum data throughput and security. They are called RAID level. But the term is misleading: It is not the case that a higher RAID level signifies higher redundancy or performance. Every RAID level has its own advantages and disadvantages and the selection of an appropriate level depends on the application. Criteria for selection are: Write speed, read speed and the degree of redundancy (and hence costs).

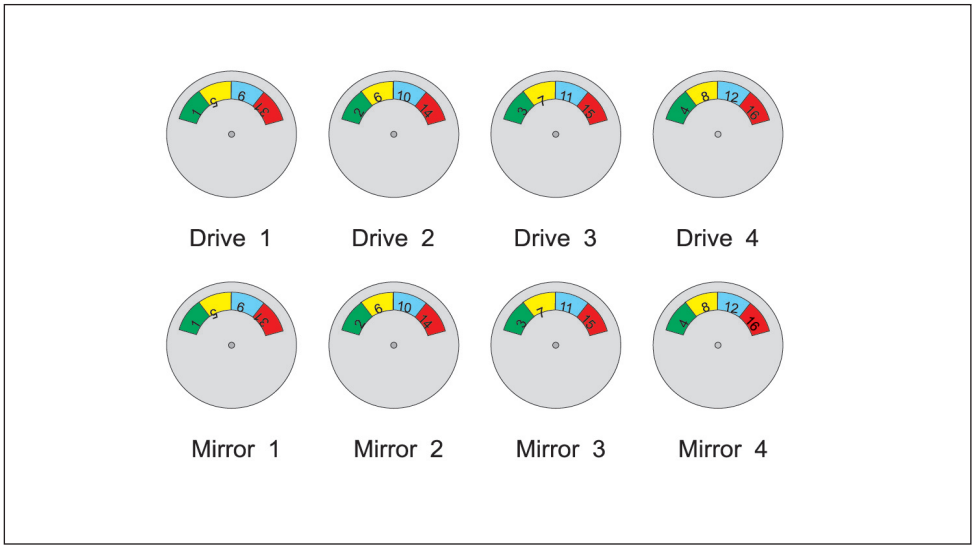
5.6.1.1 RAID Level 0

This type of RAID stands for striping, a process in which several smaller drives are grouped together to form one larger logical drive. The striping factor indicates the size of the blocks written to the individual drives. Since data is transferred to each of the drives at approximately the same time in a RAID 0 set, the throughput can be increased considerably, independent of the length of read/write requests, if these blocks are very small (approximately one byte).
The disadvantage of this technology lies in its performance for short write requests (e.g. single blocks): Additional capacity is consumed by the added room for block headers, if the physical length of the blocks on the disk is reduced (assuming 512-byte blocks and four drives, for example, then only 128-byte blocks will result). Such a capacity loss can be avoided by restricting the length of the physical blocks to the standard length. This means, however, that prior to writing a block, all of the disks, which contain parts of the physical block to be modified, must be read. It is only after this partial information has been substituted that all affected blocks can be written back to their respective disks.
Using a low striping factor with RAID 0 provides the advantage of high transfer rates for reading and writing long requests. For short requests, however, this combination performs rather poorly.
When compared with multiple smaller drives, which can process short requests in parallel, rather than with a (correspondingly larger) single drive, the inferior performance of RAID 0 becomes even more evident.

If a high striping factor (several blocks) is selected, the resulting read/write performance for short transfers is comparable to that of individual disk drives. In addition, it is possible to process several short requests simultaneously on different disks.
In both cases, however, there is still the disadvantage that no part of the stored data can be accessed if a single drive fails. In the strict sense of the word, RAID 0 (striping) does not constitute RAID (no increased data security).

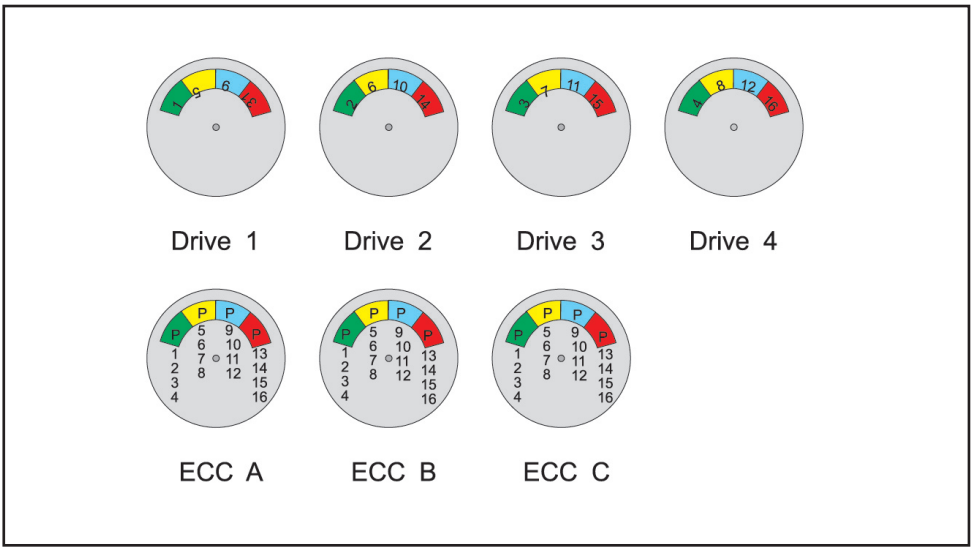
Host-based striping is available as a software solution for a number of operating systems and can even be integrated on the disk controller. When opting for the controller solution, however, the advantage of higher data transfer is not realised, as the controllers usually access disks via a single data channel. Thus, the requests to the individual disk drives are processed sequentially instead of in parallel.

5.6.1.2 RAID Level 1



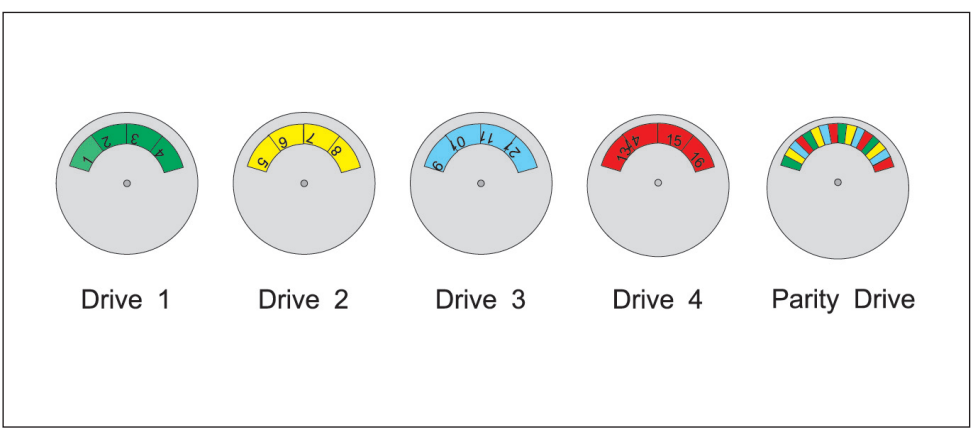
RAID 1 signifies mirroring which means that all the data on a disk is written to all other disks in a RAID set. Advantage: In case of disk failure the data remain available. The write performance is equal to that of an individual disk or RAID 0, assuming the additional processor load for the redundant write process is not calculated and the disks are distributed on channels.
In the best case, the read performance may be doubled since read requests can now be spread over two disk drives (or two sets) reading independently.
However, the costs for RAID 1 are high: Only the capacity of one hard disk is used for original data storage, the rest of the disks are taken up by redundant data. The implementation of RAID 1 is, however, straightforward.
Mirroring (of individual disks) is available either as a controller solution or as a software solution for various operating systems.

5.6.1.3 RAID Level 0+1



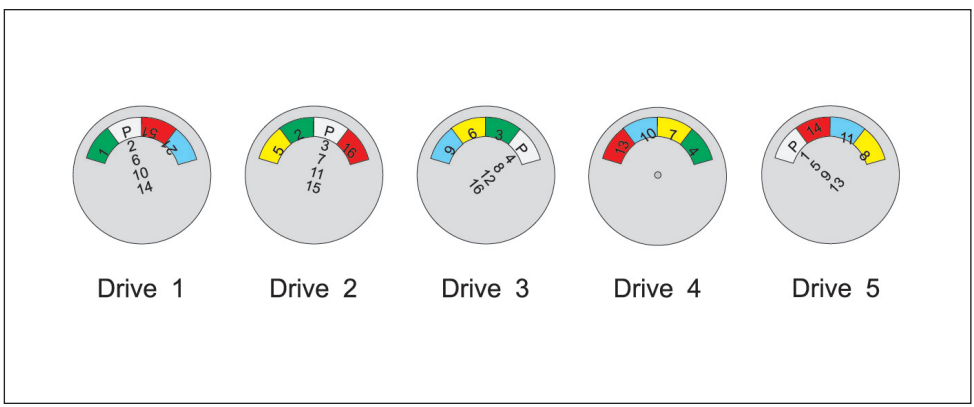
RAID 0+1, sometimes called RAID 10, combines mirroring and striping. Half of the total capacity is mirrored: The high security of mirroring is combined with the performance of RAID 0. In most cases, this is the fastest RAID solution, but the same argument applies here as in the case of RAID 1: Since twice as many disks are used, the costs are very high.

5.6.1.4 RAID Level 4



While fail-safe operation for RAID 1 and 0+1 requires double disk capacity - and is not even given in the case of RAID 0 - RAID 4 uses parity to backup the data. In addition to x hard disks, which are written to using the striping method, there is just one additional parity disk, whose nth bit contains the parity bit for the x nth bits of the remaining hard disks. If a hard disk fails, the content of the defective disk can be reconstructed bit by bit from this parity disk together with the data of the remaining disks. However, writing data to a disk always requires the parity disk to be changed - both blocks must first be read, and the new parity value is calculated from them and the new data. Therefore, performance is lower than in the case of RAID 0 or 1.

5.6.1.5 RAID Level 5



With RAID 4, the parity disk must be accessed for every write request, creating a bottle-neck at the parity disk. RAID 5 spreads parity and data blocks across all of the drives. creating a bottle-neck at the parity disk. RAID 5 spreads parity and data blocks across all of the drives. Thus each drive is the parity drive for a certain block segment. Read requests can be processed faster because they can be distributed over even more drives.
However, RAID 5 is also far inferior to a single drive in terms of access per Megabyte per time unit for short transfers.

5.6.1.6 Other RAID Levels

The aforementioned RAID levels were first described in a study by the University of California at Berkeley and have since become quasi standards. In addition, a number of manufacturer-specific RAID levels are available; these are normally only modifications of the previously described RAID levels.

So where is RAID implemented? There are two possibilities: Either in a controller between the computer and the drives (installed as a card in the computer or in a separate box) or in the host computer itself. The advantage of the first solution is that RAID processes do not place an additional load on the CPU. The advantage of the second option is that by establishing several different data paths, a higher transfer rate can be achieved, since not all of the data must pass through the bottleneck at the RAID controller to the host computer.
A number of manufacturers (of disk drives as well as controllers) provide RAID implementations, which differ in the following technical specifications:

In simple implementations, several drives are connected to a SCSI bus and are grouped together using RAID or a RAID-like configuration. Other implementations allocate one controller per disk drive, thereby increasing fail safety, as even in the event of a controller failure, data is still accessible.
Drives can be permanently integrated. In any case, however, it is better (even if more expensive) to install the drives in such a way that they can be exchanged while the system is up and running, so as not to interrupt data access (hot swap). When a drive is exchanged, good RAID systems allow the data for this drive to be reconstructed during normal operation.

There are also other factors which significantly influence the fail safety of RAID systems: If all the drives are attached to the same power supply, the probability of failure (but not necessarily the probability of data loss) is higher than with separate power supplies for each drive, or some other redundant power distribution system. To avoid a single point of failure, all cables, controllers, and host connectors should be configured redundantly as well.

Another word of caution: Beware of the misleading statistical premise that drive failures occur independently of each other. In the real world external influences (such as lightning or a power surge) may result in a much higher probability of failure for several drives at once; therefore, MTBF calculations for RAID systems are to some extent unrealistic.
Even if the time between failures of a RAID system is very high, backups must not be forgotten. RAID does not protect against accidental deletions or software errors that destroy data. Neither will worms, Trojan horses, viruses, or bombs cease to exist or develop respect for RAID.

5.6.2 Criteria of RAID Selection

RAID systems serve to improve data availability, not performance. They are generally inferior to single drives in terms of speed. However, the use of RAID levels 0 through 5 does make sense for storing large data sets which are accessed either rarely or only with long transfers. RAID offers a cost-effective solution for configuring fail-safe mass storage systems. For frequently accessed data, the use of (mirrored) single disks is recommended; for data that is constantly accessed however, the use of RAM disks is still recommended, which can be mirrored with standard Winchester disks, assuming that the number of write requests is not too high. Mirroring should generally be performed in the controller or host adapter due to a lower computer load. If you require a high transfer rate for long transfers, striping should allow several data paths to the computer and therefore be host-based.

In any event a RAID (and the computer) should be connected to a USP, because otherwise in case of power failure the actual aim, absolute data availability, is no longer achieved.

5.6.3 RAID Implementation

Hard disks can be connected to a RAID in three different ways: By RAID software, by dedicated RAID controllers incorporated in the computer, or as external RAID systems regarded by standard controllers as normal, even though uncommonly large, hard disks.

5.6.3.1 Software RAID

Windows, Solaris, Linux and many other operating systems already support grouping of multiple disk drives to form a RAID system as a standard feature. The disk drives simply need to be connected internally or externally with an SCSI host adapter. Configuration is then carried out through the operating system. This solution is both economical and simple to implement, but does have a few disadvantages as well. Especially when large amounts of data have to be transferred, and when multiple users on a network access the RAID simultaneously, there are consequences for the performance of the computer. In this case, the entire processing load must be borne by the CPU of the system (distribution of data on the individual disk drives and calculating of parity). Furthermore, the operating system is loaded from a boot disk drive, which cannot be redundant since the RAID configuration only becomes available after the operating system has been loaded. All configuration data for the RAID unit is located on the boot disk drive. Should the boot disk drive fail, the RAID system will no longer be operable. With the following hardware RAID solutions, this situation is counteracted with the help of a separate RAID controller.

5.6.3.2 Internal Hardware RAID with PCI RAID Controller

In this case, as with external hardware RAIDs, the RAID controller takes over the entire system load. As a result, the controllers work with consistent performance, regardless of the CPU load. The configuration data for the RAID system is found on every hard disk drive in the RAID system, thereby making them available even when one disk drive or even the whole controller fails. When the controller or a hard disk is exchanged, the RAID configuration data are read into it and a rebuild is initiated.

RAID controllers are equipped with internal connection sockets and an external connector for SCSI hard disks drives. In this way, the hard disk drives of the RAID system may be installed in the computer or connected externally.

5.6.3.3 External Hardware RAID

Hardware RAIDs (or SCSI-to-SCSI RAIDs) present the high-end RAID solution. Here, the controller and the hard disks are housed in a separate enclosure. The RAID system is connected to the computer's host adapter using an SCSI cable or a FC cable.

The controller and hard disks are connected directly to a backplane, allowing the shortest possible cable lengths and maximum data security as well. Additional data security may be achieved by installing additional redundant controllers. In the event of a defective controller, the system will automatically switch over to the second controller in the enclosure, without any loss of time or data. The defective controller can then be exchanged for a new one while the system is still up and running.

Like PCI RAID controllers, hardware RAIDs always function with consistent performance, regardless of the load on the system CPU. In this arrangement, data from the host adapter of the computer is transferred via the SCSI bus directly to the RAID controller. There, all data is read into the cache, after which the SCSI bus is no longer burdened. The hardware RAID controller then distributes data to the individual disk drives and computes parity, depending on the RAID level chosen. In the meantime, the CPU of the system is not hindered by the computing operations required for the RAID system. A correspondingly large cache, configured as a write-back cache, can drastically improve writing performance. Reading of data in this system happens in exactly the same way but in reverse order. Compared with PCI RAID controllers, hardware RAIDs can be used across platforms.

5.6.4 IDE RAID

An increasing proportion of external RAID systems are internally operated with IDE (ATA) hard disks. They are connected to the computer via SCSI or fibre channel.

In point of fact IDE RAIDs live up entirely to the original definition of RAID as "Redundant Array of Inexpensive Disks", with the emphasis on "inexpensive".
It later became customary to talk of "independent disks" instead of "inexpensive disks", because the RAID systems were initially very expensive.

Yet the concept behind RAID was: The use of inexpensive disks with low life expectancy is compensated for by the redundancy of the system.

It is true that nobody will use IDE RAIDs in situations demanding maximum performance, since for read and write of large files the data throughput is meanwhile considerable, but not the I/O throughput. IDE disks are too slow for databases. (This will not change in the first generations of serial ATA either, but only when SCSI characteristics such as command queuing are integrated in the standard).

In Storage Area Networks based on fibre channel as well, only part of the data capacity is used for databases. There is also a demand for file servers, e-mail servers, archives and the like in the storage network. Here the costs can be cut drastically if several servers share a common RAID, and that can by all means be an IDE RAID. The databases and other I/O-intensive applications then have room on the far more expensive fibre channel RAID.

A further, ideal application for IDE RAID with FC interface is suggested by the low price: Near-line backup becomes affordable on hard disks through the low price per GB and is then of course significantly faster than on tape or (magneto-)optical libraries. An argument which becomes important particularly in cases where backup windows become ever shorter.

6. Storage Networks

6.1 Introduction Storage Centralisation

While more and more mainframe solutions were replaced by local servers and workstations in recent years, today a centralisation of the servers and memories connected to them is taking place. The reason for this is the use of ever faster bus technologies, but also, and principally so, the increasing networking of data.

The centralisation of the memories naturally puts high demands on availability. Since system failure can bring an entire company to a standstill, it is vital that data paths and the memories themselves are designed to be redundant.
The availability of a system is measured in percent. 99.9% availability signifies an average failure per annum of 8.8 hours, 99.99% an average failure of 53 minutes and so on.

A significant advantage in setting up central storage networks is the possibility of using a storage system (e.g. RAID) by several servers at the same time. The servers share the capacity of the data memory and thereby reduce the costs both of acquisition and administration. This process is called "hardware consolidation".

The centralisation of storage solutions requires the use of suitable network technologies. Instead of direct attached storage (DAS), i.e. storage units directly connected to the individual server via SCSI or fibre channel, technologies are used also providing data over larger distances. On the one hand this can be the use of network attached storage (NAS). The file systems are provided directly via network to other servers or the clients.

On the other hand, for higher performance and larger data volumes, preference is often given to a technology which supplies the data to the computer at block level - hence like a local disk. This takes place in the storage area networks (SAN), which either use the fast fibre channel, or recently iSCSI, if the speed does not need to be all that high. Here the SCSI data are transported via the Ethernet, mostly a Gbit Ethernet. A driver installed on the computer makes the Ethernet card appear like an SCSI controller.

6.2 Storage Area Network (SAN)

LAN enables access to local systems which use communication protocols such as TCP/IP, NFS, HTTP. LANs are used to transfer data between various host systems and for user access to systems like workstations, terminals and other devices. LANs handle the flow of information between host system and user. On the other hand, SAN enables the sharing of the flow of

information with the entire environment, i.e. between the computers and the memory units and between the memory units themselves. Thus, all computers connected to the SAN have access to the memory units. Similar to LAN networks (Ethernet, FDDI, ATM, and Token Ring), SAN also has different topologies to build a storage area network.

Storage area networks are special memory networks which link the servers and memory systems via wideband networks. It does not really matter, at first, what type of operating system is installed on a server. This means that several different servers with different memory subsystems can access a network. The SCSI technology is the best known way in which to build a storage area network in its simplest form. Due to large limitations, such as the number of devices, distances, terminations etc., fibre-channel technology is normally used today as the communication medium for SANs. The characteristics of fibre channel make it very suitable as a medium. A SAN requires a guaranteed bandwidth (network speed) and low error rates.

Fibre channel offers the following advantages:
- 400 MB/s data throughput
- 2.0625 Gbit bandwidth
- 16 million devices can be addressed
- Serial connection 8/10-bit encoding
- Hardware-based connection protocol between the interfaces
- Flexible cable configuration, copper and fibre-optic, up to 10 km in length.

Currently, the transfer rates of fibre channels are being developed so that in the near future these rates will reach 4 Gbit and 10 Gbit. Further information can be found in the Fibre Channel chapter (4.3).

Since the entire exchange of information is handled in this memory network, the SAN can take the load off the LAN. When backup drives are connected to the SAN and not to the LAN, the load can be relieved even further with the corresponding software. This enables an easier administration of the systems and higher data transfer rates. A storage area network offers the ideal opportunity to centralise the entire data.

The easiest way to build a storage area network is using the point-to-point connection via fibre channel. It connects, for example, a RAID system with a computer. Only a host adapter (HBA), a cable, a fibre channel subsystem (RAID, etc.), and a loopback plug as termination (which closes the internal loop) are needed. Due to the simple cable connection, this is substantially easier than SCSI.

A larger system is built via switches in ring topology form. This ring topology is called FC-AL and connects several computers to several memory units. The FC-AL (Fibre-Channel Arbitrated Loop) enables a distribution of devices, lines, and user times. Fibre channel does not necessarily need a fibre-optic cable. A simple copper cable can be used. The only disadvantage is the reduction in distance that the copper can bridge (approx. 30 m).

SAN Software

Up to now only SAN hardware has been dealt with. To be able to share the data between common memory units, special SAN software is needed. This software uses the storage networks and enables file sharing on the subsystems. Therefore each system, regardless of the operating system, has access to each memory unit at any time and can share the data. Therefore it is referred to as a SAN operating system that is, however, set on the "normal" operating system. Companies such as Veritas, Mercury or SANergy offer SAN software which create and synchronise mirrors, and (in the case of errors) activate the mirrored memory unit, and this is all done within the memory network while the applications are online. The software enables an uninterrupted, dynamic change of the application volume. In addition, the software permits an assignment of memory units or of group memory units and a server. Some SAN software applications support clustering to guarantee improved availability of data within the memory network. The memory units are not assigned to an explicit server when clustering, except for when this has been configured so that if the server breaks down, the memory units can still be accessed or assigned to another server.

It is not absolutely necessary to have a software to build a working SAN network. In this case, only the relevant connections between the computers and the memory units must be made. Ultimately, the memory units are assigned to the computers via the operating system where the integrated standard tools of any operating system are by all means sufficient. Additional functions require additional software.

6.3 iSCSI

Internet-SCSI, in short iSCSI, is a standard which was accepted in autumn 2002, which like fibre channel transfers data at block level, hence making hard disks and other storage units appear as local SCSI drives. However, Ethernet is used as the transport medium, allowing on the one hand use of existing cabling, which saves time and expense, and on the other hand distances of any magnitude, since recourse to the wide area network is not excluded. Manufacturers such as Cisco and Adaptec are the forerunners in offering products incorporating this technology, which in the coming years will replace fibre channel wherever maximum performance is not the prime consideration.
On the computer side iSCSI is either installed via a driver, which translates the SCSI data to

TCP/IP and then transmits the data via the Ethernet. Or - what is better - a dedicated iSCSI card attends to this task itself, together with a TOE (TCP/IP Offload Engine), an Ethernet accelerator.
On the storage side, a dedicated and specialised server establishes the interface to the network, which is called an iSCSI router. In future these servers will probably be integrated directly into storage solutions, in the same way as print servers in printers and FC bridges in libraries.

6.4 Network Attached Storage (NAS)

The ever increasing demand for larger storage capacity in PC and workstation networks call for more and more file servers. Previously, this involved purchasing more PCs and workstations. Simpler and more cost-effective is the use of technologies known as network attached storage. Here, as with a communication or print server, a network device is constructed for a single purpose; in this case, for accessing storage media such as hard disk drives and CD/DVD-ROMs.
The benefits this technology offers are obvious:

Simplified installation: Due to the fact that the network storage server is actually a file server, management of this system is limited to setting network addresses and enabling storage for users. It is then immediately available to all authorised network users.
Simplified maintenance: Since the storage servers work with standardised protocols, there is no overhead when updating operating system versions.
Simplified upgrades and configuration changes: When using more than one storage server, storage space can be easily expanded and reconfigured.
Flexibility: Their variety of protocols makes network storage servers also ideal for mixed networks consisting, for example, of Windows, Unix and NetWare computers.

All data is available on all defined operating system environments, much as if they were held ready on a single server in the network. The data is accessed using the method in each case typical for the platform. From a user's point of view, all data is available as if it were stored on a file server under one's own operating system.

7. Magnetic Tape Storage

Magnetic tape storage systems are today used mainly for the data backup of servers. The low costs per GB for drive and tape media as well as the possibility of safely archiving the tapes over several years are decisive factors in this context. While, even in the age of DVD, optical storage systems are still too expensive to serve as an economical backup medium, data backup on Ultra ATA hard disks is on the upswing as a result of dropping prices.

Magnetic tape drives store data sequentially at block level. Accordingly their forte is writing large information blocks in one continuous data stream. On the other hand, magnetic tapes are rather unsuited for rapid access to specific files or information stored in scattered locations.

A more sophisticated backup strategy is therefore gaining ground. Whereas the full data backup at the weekend is still written to magnetic tapes and safely archived, data backups during the week are stored on IDE RAID systems. This enables rapid restoration of individual servers or files on a day-to-day basis, while the tapes still provide protection against total breakdown.

7.1 Linear Tape Recording Method

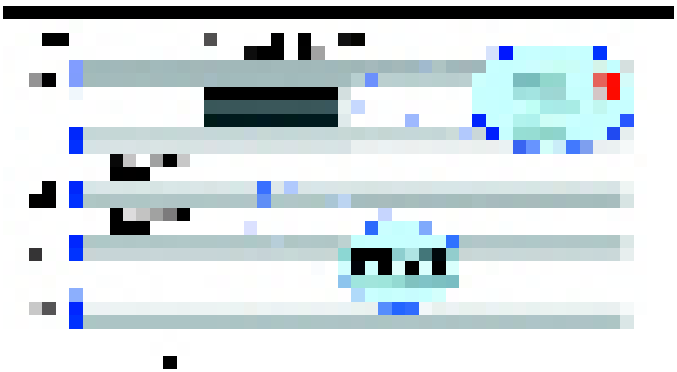
7.1.1 LTO Drives

The **LTO** (**L**inear **T**ape **O**pen) technology was developed by Seagate, HP and IBM. These manufacturers were also the first ones to offer the appropriate drives. Many other manufacturers have licensed this technology and also offer the appropriate drives. Due to the fact that a lot of manufacturers support this technology, LTO has quickly established itself on the market. The main part of the LTO recording format is made up of various linear channels which all run in the form of a serpent. The data is written on tracks which run the entire length of the tape.

Two standards have been developed: **Accelis** and Ultrium. However, as yet only Ultrium has gone into production. The Accelis format was defined for rapid access. The Accelis format uses two tape rolls and starts accessing data from the middle of the tape in order to keep access times down to a minimum.

This means that the path from the start to the end of the tape is exactly the same.
The advantage of this principle is clearly demonstrated when writing or reading parts of data, however, it is of no advantage when writing or reading the complete tape. A typical application of this format is with HSM (Hierarchical Storage Management), where it is often

necessary to read individual files from the tape. With the Accelis format, 256 tracks are divided between two data areas. The first generation will offer a capacity of 50 GB with data compression and a data transfer rate of a maximum of 40 MB/s.
The **Ultrium** format, in comparison, is the optimum medium for high capacity applications and exhibits outstanding operational security in an automated or stand-alone environment. The Ultrium medium provides only one tape roll and is ideal for backup, restoring and archiving programs.
The Ultrium format divides 384 tracks between four data areas. The servo track for head positioning is positioned between the different data areas. The tape is written in the form of a serpent with eight parallel tracks. The data is written from the inside to the outside, as the inner data area is physically better protected. With data compression, the first generation offers a capacity of 200 GB and a data transfer rate of max. 30 MB/s, while the second generation available since 2003 doubles capacity and speed.
Both types of cartridges have a free-from-contact cartridge memory (LTO-CM) with 4 KB, in which manufacturer and user information is saved. A back software support is necessary for reading or writing the user information.

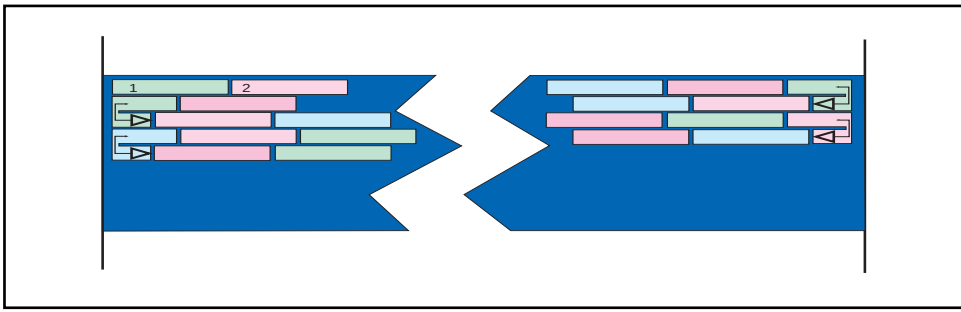


Track Recording with LTO Drives

7.1.2 DLT/SDLT Drives

The **DLT** (**D**igital **L**inear **T**ape) technology records data, like LTO, in a serpentine-like linear way, so that the tracks lie parallel to the tape edge. 1/2“ magnetic tapes are written in compact cassettes. DLT4000, DLT7000 and DLT8000, the last three drive generations, use the same DLTtapeIV tape medium and with compression achieve up to 40, 70 and 80 GB storage capacity.

Until now, data was written through two channels located on the read/write head, at a density of 62500 bits/inch, and the transfer rate was 3 MB/s. In this format, one data unit consisted of twenty 4 KB blocks, each individual data block equipped with various error detection procedures (parity, cyclic redundancy check (CRC) and error detection check (EDC)). The last four blocks of each unit consist of four ECC blocks (ECC, error correction check).
In the new storage method, **Super DLT**, the data is written by four channels located on the read/write head, at a density of 86000 bits/inch, which increases the transfer rate to 5 MB/s. In this storage method, a data unit consists of twenty-four 4 KB blocks, which are also equipped with error detection procedures. The last five blocks of each unit are ECC blocks.
After every write operation the DLT controller carries out a read test. The written data is compared with the contents of the buffer and corrected if an error is detected. These safety measures guarantee very high data integrity.



Track Recording with DLT Drives

An additional advantage of this technology is that wear and tear on the tapes and read heads is low. This is achieved through the use of stationary magnetic heads and a simple tape threading system. Therefore, DLT drives are capable of up to 500,000 tape winding cycles per cartridge. DLTtape IV tapes are capable of more than one million cycles, due to the use of the new technology described above.
In 2001, Quantum launched the first SuperDLT drives on the market. They are equipped with a new technology which can reach 220 GB at 22 MB/s (2:1 compression). They require the same size of tapes but with a better coating. The systems are read-compatible with DLT8000, DLT7000 and DLT4000, which is the most important sales argument for SuperDLT over LTO. SDLT320, the second generation of SuperDLT drives with 320 GB capacity at a transfer rate of 32 MB/s (2:1 compression), has been available since mid-2002.
To downwardly round off the DLT product lines, the year 2000 saw the introduction of the DLT VS format. The current generation VS-160, at full 2:1 compression, is capable of writing up to 160 GB of data at a throughput rate of 16 MB/s.

7.1.3 ADR Drives

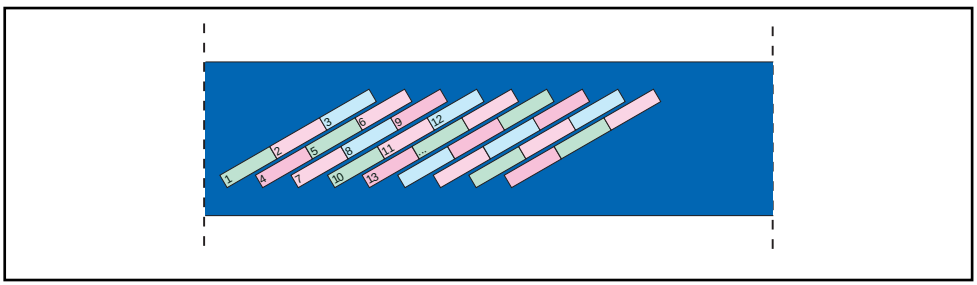
The **ADR** (**A**dvanced **D**igital **R**ecording) technology was developed by Philips and is used in the drives from OnStream. The data is recorded on eight parallel tracks. Servo information is stored between the data tracks which is used to accurately position the head.
This servo information is also able to detect faulty sections on the tape. ADR drives have variable data transfer rates ranging from 0.5 to 2 MB/s. The drive is able to adjust itself to the data rate of the system which therefore avoids time-consuming repositioning of the tape. ADR drives offer a capacity of up to 50 GB (compressed) per cartridge with a maximum data transfer rate of up to 4 MB/s.

7.2 Helical Scan Recording Method

7.2.1 Exabyte Drives

Very high data densities are achieved with magnetic tape devices from the audio and video sector, which use the helical scan method. Rotating heads record the tracks diagonally onto the tape. These tracks can feature an extremely high density. Devices that have been adapted to meet data storage requirements using this technology have already claimed a large share of the backup technology market.

Exabyte drives emerged from Video8 technology and were specially improved for use as backup devices. In the course of this, their recording technology was refined. The large amounts of data that can be stored on these cartridges (a maximum of 40 GB with the Exabyte Mammoth and a 170 m tape) help to reduce the price per GB for the medium, and the space requirement of a cartridge is minimal. Dispatch and storage are unproblematical. And using the latest generation of cartridges, even long term archiving is not a problem. The manufacturer guarantees a minimum archive life of 10 years.



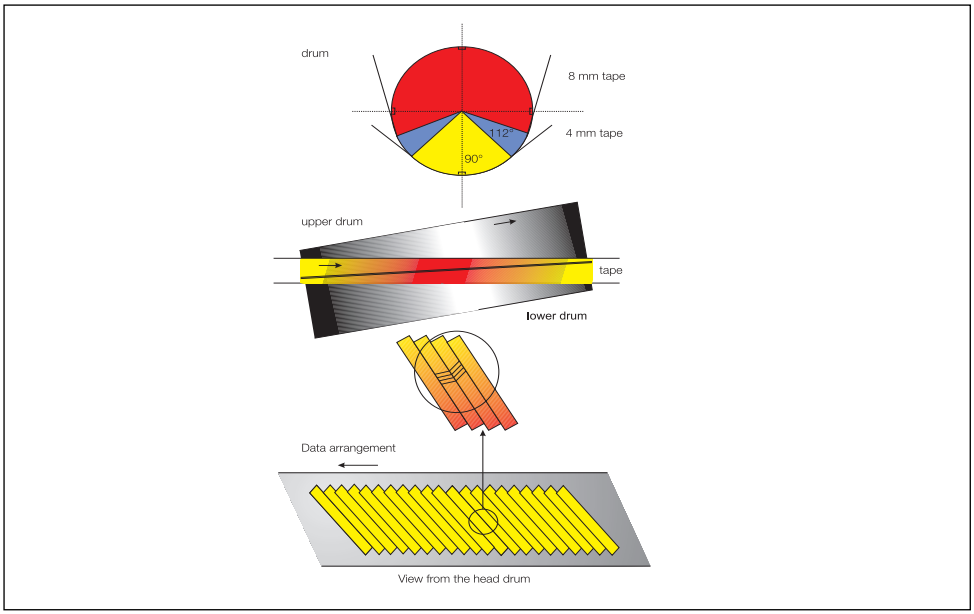
Track Recording with Helical Scan Drives (DAT, Exabyte, AIT)

7.2.2 DAT Drives

DAT (**D**igital **A**udio **T**ape) technology also emerged from digital recording methods in the audio field. It uses the helical scan method to record data on a magnetic tape. In comparison to the 8 mm Exabyte tapes, however, DAT tapes are only 4 mm wide and in the drive itself, the tape's coiling angle around the tape drum is only 90° while Exabyte technology uses a coiling angle of 112° which puts greater strain on the tape, as well as on the drive's head.
The **DDS** (**D**igital **D**ata **S**torage) format was developed by HP and Sony and has become the standard. The developments in DAT are the direct consequence of competition between manufacturers. DAT drives can store up to 8 GB (120 m tape) at transfer rates of up to 1 MB/s using the DDS-2 standard. The claim of 8 GB storage capacity typically represents a figure achievable with data compression (2:1). Theoretically, up to 16 GB of data can be stored. In practice, however, these theoretical maximum values prove to be over-optimistic.

Drives recording in DDS-3 format are capable of recording up to 24 GB on a 125 m tape at a data transfer rate of 2 MB/s. The DDS-4 standard ensures capacities of up to 40 GB (with compression and 150 m tape) with a data transfer rate of up to 6 MB/s. It also ensures a read and write compatibility with DDS-2 and DDS-3. DDS-3 drives offer special features that go a long way toward ensuring error-free data storage.
The first models of the latest 5th generation of DDS drives from Hewlett-Packard and Seagate are expected in May 2003. At the same speed as DDS-4, they offer an almost double native capacity of 36 GB. The matching tapes will be 175 metres long. The new DAT 72 drives can read and write old DDS-4 and DDS-3 media.
One significant source of errors is a dirty head drum. The installed **head cleaner** automatically sees to the cleaning of the heads at 15-minute intervals. That, combined with a cleaning tape, ensures years of trouble-free drive operation.
If an inserted tape is damaged, too old or no longer usable, however, or if the cleaning tape is defective, the network administrator is immediately notified. This feature is called **TapeAlert** and is a new firmware developed by HP, which carries out network-wide diagnostics of data media problems. In addition, TapeAlert constantly monitors the condition of the drive and the media during a backup.

With the aid of the **time-tracking** function, the read head is always aligned to the middle of the data track. This guarantees that even tapes that have been written to by the drives of other DAT drive manufacturers can be reliably read. A special feature of HP DAT drives is the **OBDR** (**O**ne **B**utton **D**isaster **R**ecovery). With the help of the appropriate software, this enables boot-capable cartridges to be generated, which can recover the system in emergency situations without needing additional floppies or CD-ROMS.



Helical Scan Data Recording

There are automatic cartridge changers for DAT drives as well as for all other tape drives. The simpler models only allow sequential changeover of the cartridges. The more expensive random access models usually require special software to operate the so-called tape libraries. The smallest tape libraries or changers hold six cartridges, the larger ones can hold more than 100 cartridges. These types of tape libraries make it possible to store data without operator intervention.

We recommend that only DAT cartridges manufactured explicitly for data storage be used. For example, these cartridges bear the DDS seal for DAT drives and the D-Eight seal for Exabyte drives. As with all magnetic tape drives, the read and write heads should be cleaned regularly. It is also important that only cleaning kits recommended by the manufacturer be used. Incompatible cleaning tapes can cause mechanical damage. Cleaning is recommended every 30 GB, or at least once a month.

7.2.3 AIT Drives

The **AIT** (**A**dvanced **I**ntelligent **T**ape) launched by Sony in 1997 is based on 8 mm helical scan technology, like Exabyte, but is not compatible with it, in part due to the use of a different cartridge format. The maximum capacity is currently approx. 260 GB (100 GB uncompressed), whereby the chances of achieving a good compression factor (approx. 2.6) have increased by the enhanced algorithm ALDC (Advanced Lossless Data Compression). In addition, just like DAT drives, this drive fits in a 3.5" slot. A new feature is the self-cleaning mechanism, which significantly reduces the necessity of using cleaning tapes in normal working environments and increases the service life of magnetic tape and drive.
In case of serious errors, the internal head cleaner is automatically activated, the AME media also making a decisive contribution to reducing the cleaning intervals. An interesting novelty is also the use of a memory chip **MIC** (**M**emory **i**n **C**assette) in the cartridges. The size is 16 kBit for 25/65 GB media and 64 kBit for all other ones. On this chip, backup software can store information, which for conventional technologies is only stored at the beginning of the tape. This makes it possible to halve the average seek time on a tape, which will then be 27 seconds. Today, all major manufacturers of backup software support the AIT drive with the features set out above.
The AIT-3 drive available since summer 2001 has put Sony into the performance class of LTO and SDLT drives. By using 230 m tapes, the drive achieves a data transfer rate of up to 31 MB/s (compressed) with a capacity of max. 260 GB. Sony will continue the AIT format at least until the year 2005 through three further generations reaching the TB range. At special feature is that the drives have read and write compatibility with all the predecessors.

7.3 Other Recording Methods

7.3.1 QIC Drives

With the **serpentine format recording method**, the 1/4" cartridges are written to at high speed in several cycles, each time on one track, parallel to the tape edge. This serpentine track format must also be traversed when reading data.
1/4" tape cartridge drives work with the **QIC** (**Q**uarter **I**nch **C**artridge) format as a standard and record in streaming mode. If the transfer rate of a system exceeds the recording rate of the streamer, the tape will be written to continually (streaming), without pauses and adjustments, and thus be used to its full potential. When the streamer's data buffer is empty it requests more data from the system.

In order to fully utilise the tape on systems with slower transfer rates, the tape is stopped and backed up to the end of the last recorded data. The next data block is directly appended.

Unlike helical scan technology, one advantage of the QIC technology is that the tape always remains within its cartridge, resulting in much longer life.

7.3.2 Mainframe Drives

1/2" magnetic tape drives are actually antiquated in today's world of high-performance data storage options. But they are still used since they are so widely spread and are mainly used for exchanging data with mainframes. Since more compact and less expensive media are currently available, they will certainly become less important in the future. 1/2" tapes are usually written to with 9 tracks.
On a large magnetic tape spool (with a tape length of 730 m), 23 MB can be stored using a write density of 800 bpi. This can be increased to almost 46 MB using a density of 1,600 bpi, to 92 MB using a density of 3,200 bpi, or 180 MB using a density of 6250 bpi. These are less than favourable ratings in comparison to storage media for magnetic tape cartridge systems that can record 10 GB and more on one cartridge. The disadvantages of the 1/2" drives are the long backup and restore times and the space required for both the magnetic tape drive and the tapes.

7.4 Software for Magnetic Tape Storage

7.4.1 Backup Software

Due to the unavoidable possibility of data loss, whether it is caused by human error, hardware failure, sabotage or viruses, backups cannot be neglected. Although it is possible to carry out a backup without supporting software, ever larger amounts of data and the use of libraries have made this an almost impractical solution. Appropriate backup software packages make daily backups easier, while offering many tools that also allow management of the data to be stored, as well as data that has already been stored (archiving and HSM). The necessary developments in the data media themselves also result in constant revisions and expansions of the various backup software packages.
Storage capacity requirements are growing faster than the capacities of individual media. Since purchasing more and more hard disks causes costs to skyrocket, there is an increasing need to utilise archiving, i.e. storage of data on more affordable but slower media. On the other hand this leads to high input in terms of administration.

7.4.2 HSM Software

HSM (**H**ierarchical **S**torage **M**anagement) software is able to perform this administration independently. Based on preset parameters like frequency of access, file size, file age etc., files are evaluated and if necessary transferred to ever more affordable, slower media, for example, from hard disk to MO automatic changers, from there to tape libraries, and then on to tapes kept externally. In the latter case, the system administrator might be asked, if necessary, to insert the appropriate tape.

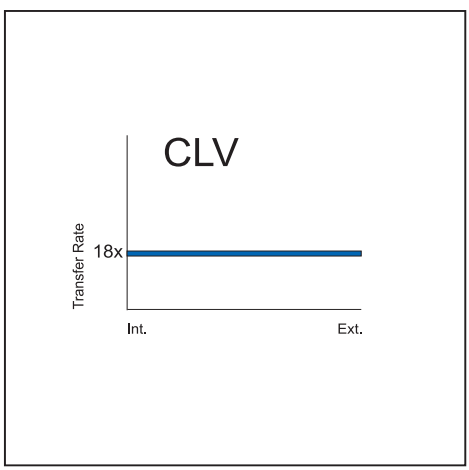
A virtual file system always shows all the files to the user, not only those actually contained on the hard disks, but also those that have been archived. When a file is opened, it is first copied back to the hard disk. The only thing the user notices is the correspondingly slower access time. In the worst-case scenario of the data being located on external tapes, waiting times when accessing a file might be as long as half an hour or longer. Should the slower access time intimidate some users, then the fact that 80% of all data is accessed less then once a month might serve as some consolation. Not even 5% of all data is accessed more often than once a week.

8. Optical Storage Media

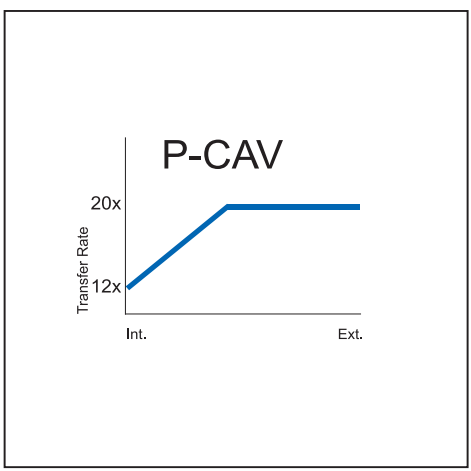
Optical disks have found general acceptance in the form of CD-ROM technology. New technologies, like CD-R and CD-RW, as well as the increase in capacity created by DVD, are causing continuing growth in this market. This area also encompasses magneto-optical technologies, where improved storage capacity is also being achieved.

8.1 CD-ROM Drives

CD-ROMs are removable optical disks only for reading data that has been recorded as part of the manufacturing process. With a storage capacity of approx. 650 MB, the CD-ROM is an ideal medium for mass distribution of data, e.g. for software distribution, since the reproduction costs are much less for greater quantities. The drawbacks of the CD-ROMs, i.e. the inability to be rewritten and the relatively slow access times (currently 85-120 msec, depending on the drive) are usually not important. The media itself is extremely inexpensive. Mass production of CD-ROMs is usually only worthwhile in batches of 100 and greater. The production of CD-ROMs is much like that of long-playing records, the information is pressed onto the media.



CLV Technology



CAV Technology

The reading speed of the CD-ROM is measured in reference to the transfer rate of an audio drive, which is 150 KB/s. A 14x drive, for example, can read at 2.05 MB/s. The present CD-ROM drives work at up to 52x speed and with a data transfer rate of 7.8 MB/s. Drives with read speeds higher than 14x are based on partial **CAV** technology (**C**onstant **A**ngular **V**elocity).
The drives are distinguished by a variable data transfer rate and a constant rpm in the inner zone of the CD. As the outermost track of a CD is read, the drive switches to the standard **CLV** (**C**onstant **L**inear **V**elocity) method where the disc rotation speed is adjusted prior to reading. This yields a constant data transfer rate.
Furthermore, in addition to the obligatory increase in speed, present-day CD-ROM drives offer a new feature: **Multiread**. For CD-ROM drives this signifies that they can also read rewriteable CD-RWs.
In addition to the single-disk drives, CD-ROM changers are also offered. The typical changer time per CD is around 7 to 10 seconds. CD-ROM changers are especially well suited to applications where large amounts of data that are seldom accessed need to be available for direct CD-ROM access by a computer (e.g. desktop publishing or multimedia applications). Changer systems available on the market make it possible for a single peripheral unit to manage quantities of data up to the terabyte range.

8.2 CD-R and CD-RW Drives

Using CD recorders (CD-R), data can be written onto blank CD-ROMs (write-once CDs) without great effort. Data is written onto the blank CD-ROM using a laser beam. In fact, it is becoming more economically feasible to produce single CD-ROMs. Thus, small quantities of product catalogues, software products or graphic collections can be created on CD-ROM. Writeable CDs are also suitable for long-term archiving. The data media are very stable, but the data cannot be changed once it has been written.
CD-ROMs written with such a CD recorder can be read on normal CD-ROM drives, just like those distributed over-the-counter in larger quantities. With automatic changers that contain a CD recorder there is a cost-effective way to archive and back up in small or medium-sized networks. If, however, production or copying of CDs is at the forefront, CD duplicating systems have lately been offered as stand-alone devices for duplicating CDs.

CD rewriteable (CD-RW) is a technology that allows a medium to be written to repeatedly, up to 3000 times. These drives can also write a single data session to CD-R media. CD-RW media can be read with commercially available CD-ROM drives having multiread capability (see table for compatibilities).

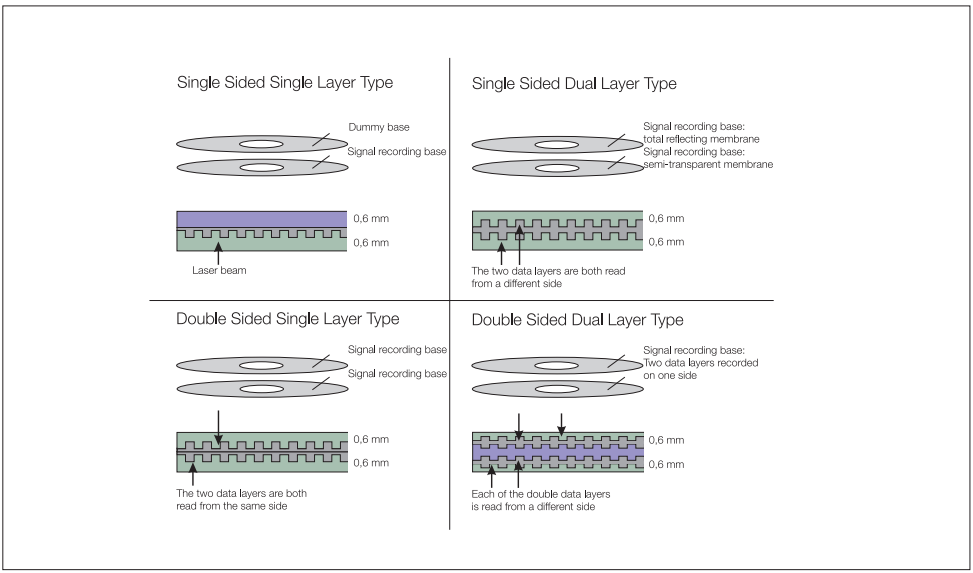
Media	CD-ROM	CDR	CDR W	DVD-ROM	DVD-R	DVD-RAM	DVD-RW
CD-ROM	R	R	R	R	R	R	R
CD-ROM	R	W/R	W/R	R		R	R
CD-RW	R	W/R	R	R		R	R
DVD-ROM				R	R	R	R
DVD-R				R	W/R	R	R
DVD-RAM						W/R	
DVD-RW							W/R

R = read compatible W = write compatible

8.3 DVD Drives

For a long time, the capacity of CD-ROMs was considered sufficient. With pressure from the video industry, a new optical technology emerged with **DVD** (**D**igital **V**ersatile **D**isk, originally: Digital Video Disk) which extended the media capacity to 5.2 GB. By using both sides of the medium and employing a second data layer, this capacity is to be quadrupled in the future. This capacity extension was achieved by reducing the minimum length for pits and the track spacing of the windings. Furthermore, the DVD drives can read all common CD formats. For some time now, DVD-R drives (recorders) have been available on the market. These, however, as yet only provide a storage capacity of 2.6 GB/side.

The unrivalled success of DVD was, however, halted due to the lack of a common standard. Toshiba, Hitachi, and Panasonic developed the DVD-RAM drive with a 2.6 GB capacity per side; it works with a caddy. Sony, Philips, and HP, on the other hand, have developed a DVD-RW drive, which offers a 3 GB capacity, works without a caddy, and can also be read in future DVD drives. Furthermore, the DVD-RW drives can also read the contents of CD-RWs, while DVD-RAM drives cannot.



Setup of DVD Media

8.4 Magneto-Optical Drives

Magneto-optical data storage, representing optical storage technology, has a number of advantages over competing technologies. Magneto-optical disk drives use a laser to heat the MO disk surface, so a magnetic pattern can be written to it, all without touching the disk or causing any wear on it.

The data media can be written to **double-sided**, and data is read through the reflected laser beam. The laser beam returns information that directly corresponds to the magnetisation of the material. Until now, writing to MO media was a very slow process. The reason was that any writing process needed to process each block three times: the first revolution would erase the block, the second would write the block, and the third revolution would verify it. Therefore, in comparison to a read request, a write request would take two additional disk revolutions.

By now, all commercially available magneto-optical drives are capable of LIMDOW (Laser Intensity Modulation Direct Overwrite), and can be switched to this mode. LIMDOW means, that it is now possible to directly overwrite old data with new data in a single pass, without the need for an erase cycle first. The claim of lower performance in comparison to conventional removable disk drives, for example, which reduced the magneto-optical technology's attractiveness until now, is no longer valid.

In summary, the advantages presented by these technologies are as follows: The data media are small, lightweight, and not very expensive. The MO drives can be used by a computer just like a normal disk drive. No special file system software is required.

Areas where it makes sense to use magneto-optical disk drives are, for example, those where the data medium has to be removed from the computer (for example, high-security applications) and the higher performance of Winchester disks in removable cartridges is not required (or you do not want to pay their higher price). In addition, MO disks provide quick access to backups without copying data onto other media, such as disks, before it can be accessed. MO changers make it possible to archive very large amounts of data. However, these changers will become less important in the future, as CD-RW and CD-R jukeboxes take their place.

WORM disks (**W**rite **O**nce **R**ead **M**ultiple) are optical disks which are written to only once but can be read any number of times. With the introduction of CD-R, however, WORM technology has almost become obsolete.

8.5 Outlook

CD-ROM technology has won a place for itself in the fields of both audio and computing. The development of CD-R (once writable) and CD-RW (rewriteable) will ensure that these technologies will be successful in securing a market share. The situation is somewhat different for the DVD (Digital Versatile Disk) as this technology may replace CD-ROM in some areas, since each disc's capacity is many times higher.
In parallel with the developments in DVD, there is also work on developing an MO video disk. This so-called MO-7 (ASMO) is expected to have a storage capacity of 6.1 GB. At the same time, drive manufacturers like Maxoptix have introduced new products to the market that double the current capacity to 5.2 GB. The optical data storage market is developing rapidly. The **OSTA** (**O**ptical **S**torage **T**echnology **A**ssociation) assists in gaining an overview of present and future technologies: www.osta.org

9. Main Memory

When judging the performance of a computer, not only is the processing power of a computer's CPU a decisive factor, but also the system's working memory. New technologies, faster chips and different hardware configurations make it increasingly difficult to choose the appropriate memory module for a specific computer. Some typical terms are explained in this chapter.

9.1 Memory Technologies

In principle, there are static, and non-volatile dynamic, memory technologies.

9.1.1 Non-Volatile Memories

What all non-volatile memories have in common is that they are able to retain their contents, i.e., the stored data, even if the power supply is switched off. Examples of memories in this category are EEPROM and Flash memories.

EEPROM, E2PROM: **E**lectrically **E**rasable **P**rogrammable **R**ead-**O**nly **M**emory
EEPROM is the term used to describe all non-volatile memories whose contents can be changed or deleted electrically rather than by other means (e.g., by exposure to ultra-violet light as with its predecessor EPROM). There are several architectures available, each with different electrical properties. For example, EEPROMs are included in various memory modules where the particular module properties (operating voltage, access times, bank set-up, error correction, etc., down to the manufacturer) are stored in coded format (e.g. SPDs on SDRAM DIMMs).

Flash Memory
A characteristic of flash memories is that individual bytes can be addressed and read out, whereas write and delete processes can only operate on blocks of addresses at a time. Read access times, currently about 100ns, are about double those of dynamic memories. The number of programming and delete cycles is limited to about 100,000. In general, data is retained for a guaranteed period of 10 years. SIMM, PC Card (PCMCIA), Compact Flash (CF) Card, Miniature Card (MC), and Solid State Floppy Disc Card (SSFDC) are among the various forms of flash memory available. Regardless of their exterior appearance, there are two main types of flash memory modules: Linear flash and ATA flash. Linear flash modules have a "linear" address space where any address can be directly accessed from outside. On the other hand, for the ATA flash cards address conversion takes place internally, so that this version is accessed similarly to a hard disk, a fact that can for instance simplify driver programming.

Flash memories can for example be used as mass or programme memories in notebooks, network routers, printers, PDAs, and digital cameras.

9.1.2 Dynamic Memories

DRAM: Dynamic **R**andom **A**ccess **M**emory
DRAM is a dynamic memory with a free-to-choose access. It can be accessed at any time on every data cell. An extreme example of a contrast to this would be, for example, tape drives. This type of memory is generally used as main memory.

A characteristic of dynamic memories is that information is stored in capacitors, which, like accumulators, can receive energy and retain it for a certain time. A capacitor is charged to store a logical one for instance, whereas it is discharged for a logical zero. The capacitors are arranged in a matrix of rows and columns.
In order to keep the chips physically small, and to reduce the number of connections and related costs, these modules are controlled electrically on two levels: The address of a data element is subdivided according to the matrix into a row address and a column address, each of which are transmitted one after the other via the same connections to the chip. The advantage of this technology is that it can be manufactured at a relatively low cost and with large densities. Its disadvantage is that the capacitor setup is not ideal. If it is not used for a long time it will eventually discharge, just like an accumulator. To prevent the loss of data, it must be refreshed at regular intervals. Further developments have been made to this basic architecture and they are briefly described below:

FPM: Fast **P**age **M**ode
The Fast Page Mode available with some memory modules is the result of an improvement of the "standard" DRAM memory. The technology-limited delay times are reduced for particular applications using a special addressing method. Normally within computer programmes, contiguous memory contents are processed by the processor. When a memory bank is accessed, first the row address is normally transferred, and then the column address. However, for consecutive memory addresses it is only the column address which changes since consecutive data are generally located in the same row (on the same "page"). Therefore, it is not necessary to re-transfer the unchanged row address. The Fast Page Mode makes use of this fact. Row and column addresses are only transferred for the first access. For every subsequent access only the column addresses need to be transferred, so that the cycle time needed before the data is available at the outputs of the memory bank, is reduced. Of course, in order to do so, this mode of access must be supported by the system used and its chip set.

EDO: Extended **D**ata **O**utput
With respect to FPM memories, memory modules with EDO represent a further development. In a similar way, memory access is accelerated by the use of particular addressing techniques. In the case of FPM memories, the electrical signal of the data circuits is deleted (not to be confused with the contents of the memory cell which are retained!), when new address information is referenced. Since a certain period of time is needed to process data, the memory must be "kept still", during a finite interval, so that the electrical signals on the data circuits can be received and further processed. With EDO memories, the output stage has been structured in such a way to retain referenced information even if a new address is accessed. Thus it is possible to simultaneously process the referenced data word and load the next address required in the memory module. This helps to further reduce the cycle times.

BEDO: Burst **E**xtended **D**ata **O**utput
A further improvement of the EDO memory modules is the BEDO which is the last asynchronous DRAM. As the BEDO was introduced at the same time as the SDRAM, it has never had the opportunity to succeed in the market. It stands out due to an additional Burst mode. After transmitting an address, the module outputs the contents of the cell similar to an EDO-RAM, however, it also attaches the following three cell contents with a clock cycle of one pulse per value.

SDRAM: Synchronous **D**ynamic **R**andom **A**ccess **M**emory
Like FPM and EDO, SDRAM technology is another development in existing memory architectures and of their access modes. But in contrast to FPM or EDO, SDRAM technology is not downward compatible, i.e., SDRAM memories can only be used in computer systems which explicitly support this technology. The development present in SDRAM is just the relocation of part of the memory controller to the memory chip itself. A step which is perhaps similar to the introduction of IDE hard discs (SCSI discs are perhaps an even better example), which in a similar way incorporate within their housing the controller adjusted to their particular needs. Just like FPM or EDO access methods, SDRAM technology offers a powerful means of accessing consecutive data in the address space. Just as for all types of DRAM, a typical SDRAM access is executed with the consecutive transmission of the row and column addresses. Unlike with previous technologies, a "command transmission" to the SDRAM is also performed during this addressing and, depending on the command, defined processes are set in operation on the memory chip. A typical command could be the following: Read out address X and the three subsequent addresses. In this case, the start address X is transmitted together with the command and, without having to do anything else, the same operation is executed on the contents of all four consecutive addresses. Since the required data must be valid at a definite time, SDRAM modules are provided with a clock pulse for synchronising all processes. SDRAM memories provide speed advantages when very large amounts of data

have to be transferred in blocks, e.g., with large graphics.

DDR SDRAM: Double **D**ata **R**ate **SDRAM**
The DDR SDRAM, normally known as SDRAM II, is a faster version of the SDRAM. It has been further developed so that it can read data on the increasing and decreasing flanks of the system clock, which therefore doubles the transfer rate of the memory. The maximum transfer rate of this technology is 1 GB/s.

RDRAM: Rambus **DRAM**
Rambus DRAM is based on one of the completely new technologies developed by the company Rambus Inc. The RDRAM is in the position to meet very high demands due to the overall reworking and redefinition of the inner structure e.g. conductor path lengths, pin capacities and voltage fluctuations. The memory is addressed on the increasing and decreasing flanks of the system clock. A one-channel Rambus memory with 1.6 GB/s offers three times the performance of a 64-bit 100MHz SDRAM module. The main advantage of the Rambus technology is that 2 or 4 Rambus channels can be used in parallel. Thus with 2 channels a bandwidth of 3.2 GB/s can be achieved, and with 4 channels a bandwidth of 6.4 GB/s is possible.

9.1.3 Static Memories

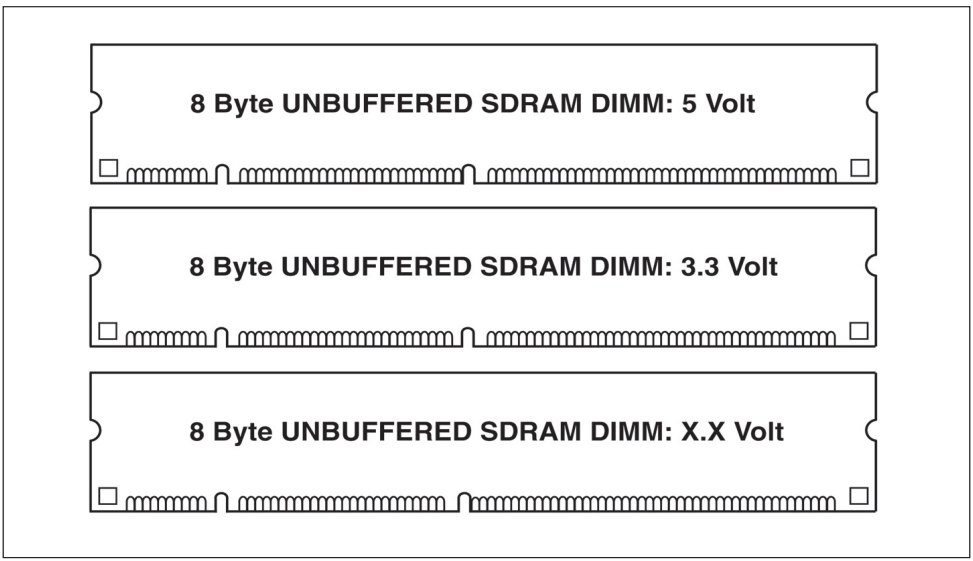
Static memories are used as cache memories due to their high speed. Unlike dynamic memories, they do not require that memory contents be refreshed. Static memory cells consist of a circuit composed of various modules. Depending on the data to be stored, the circuit adopts a specific state and then locks itself, so that its state can only change if induced from outside. Static memory only loses its data content if the power supply breaks down. Due to the more complex structure of their memory cell, the memory densities which can be realised in static memories are lower than for dynamic memories. Furthermore, for a comparable memory capacity, they are considerably more expensive than dynamic memories.

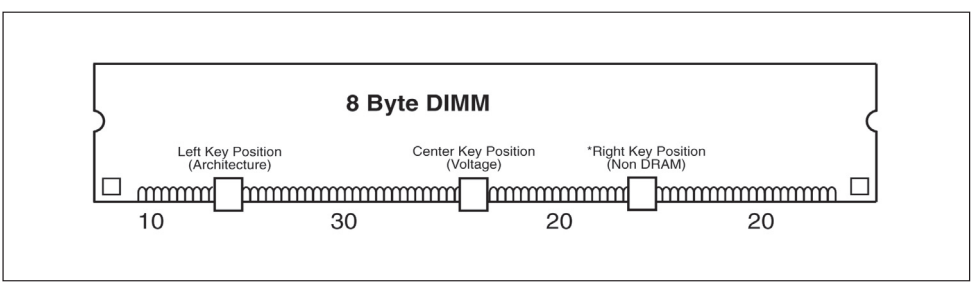
9.2 Modular Structural Forms

Why modules? Modern computer systems require memory sizes, defined by data width (i.e., the number of data bits used) and address space depth (i.e., the number of data words that can be accommodated in the memory), which cannot be used within a single memory chip. Therefore, modules are built from single memory chips, which can be electronically addressed almost as though they were large chips. The following widely used structural forms can be distinguished:

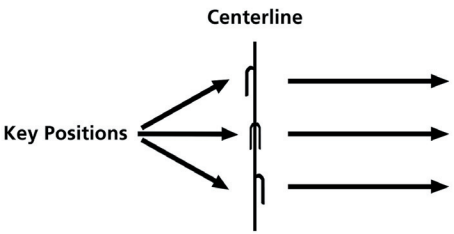
SIMM: Single **I**nline **M**emory **M**odule
SIMMs have contact areas on both sides of the board which are, however, interlinked. Therefore a 72-pin SIMM has 144 contact areas, but only 72 of them are electrically relevant.

DIMM: Dual **I**nline **M**emory **M**odule





A DIMM has contact areas on both sides of the board but in this case the opposite contacts are electrically isolated from each other. Thereby, more electrical connections are available to a given sized module (= twice as many as with a comparable SIMM), so that more data bits or a larger address space can be accessed. Furthermore, the number of connections for the operating voltage and control signals increases, providing advantages with regard to the design of the board and consequently the module's electrical properties. DIMM modules have already become standard components in the PC, server and workstation sector.



Left Key Position	Center Key	*Right Key
RFU	5V	No Key
**1st GENERATION 168 PIN DIMM	3.3V	No Key
***2nd GENERATION 168 PIN DIMM	X.XV	No Key

9.3 Terms relating to Memory

Cache

A cache is a relatively small high-speed memory, usually SRAM, which buffers frequently required data between the CPU and main memory. The CPU can access data in the cache much quicker than the data in the main memory. A small cache memory can therefore considerably improve the computer's performance without incurring a large amount of extra expense.
There are different levels of cache, according to the position of the memory in the data stream.

Level 0 Cache:
Decouples the data stream between the various processing units within the CPU. Its typical size is between 1 and 128 Bytes.

Level 1 Cache:
Decouples the data stream inside the CPU from the outside; its size varies from a few Bytes (128) to several KBs.

Level 2 Cache:
Decouples the data stream between the CPU and the memory bus connected to the main memory. Its size ranges from 256 KBs to several MBs.

Level 3 Cache:
Decouples the separate system bus from the memory bus with multi-processor systems.

Parity/Non-Parity
External influences (electromagnetic fields, voltage fluctuations, radioactive radiation - cosmic radiation, etc.) can cause individual data bits to "invert": Thus a logic zero becomes a one and vice-versa. The stored data is therefore altered. Non-parity modules only store data and do not offer any means of error detection. Parity modules, on the other hand, store data and checksum information. This is derived as follows: If the number of set bits is even, the parity bit is likewise set, whereas if the number is odd, the parity bit is cleared. This parity technique helps detect one-bit errors. Since no error correction occurs, when a parity error occurs most systems do not report an error message. Normally, 1 parity bit is stored per data byte (8 bits). Attention: If external interference causes two bits in a data word to invert, the checksum remains the same and the error goes unnoticed! However, the probability that this occurs is very small.

ECC: Error **C**hecking and **C**orrecting
The ECC process is able to detect both one and two-bit errors. Moreover, one-bit errors can even be corrected. The underlying method is occasionally also referred to as "Error Detection and Correction" (EDC). Using particular algorithms, checksums are formed in blocks (similar to the CRC method) and are stored in separate memory sections. With the help of these checksums, individual bit errors can be detected and corrected. As for error correction of audio CDs, the software user will not notice this procedure take place.

Memory Bus
The memory bus is the system bus between I/O and CPU that addresses the main memory (memory) through the chip set. It consists of the address bus and the data bus.

Small **O**utline **J**-Lead (Chip)

Thin **S**mall **O**utline **P**ackage (Chips).

What memory for what computer?

In principle, the type of memory which can be used is specified by the hardware manufacturer. The manufacturer defines the memory to be used according to the specifications of their particular system. Wherever possible, servers should be fitted with ECC.

The modules must be identical within a given memory bank. However, depending on system specifications, different banks can be fitted with memory modules of different capacities and from different manufacturers. If in doubt, the system manual should be consulted.

10. Communication

10.1 Definition

Communication takes place between senders and receivers. Communication can be one-way or two-way and serves for the exchange of information. When such information takes the form of data, text, images or voice, one talks of telecommunication.

10.2 Prerequisites for Successful Communication

Three essential prerequisites must be met for successful communication:
- There is at least one sender and at least one receiver
- Sender and receiver are linked to a joint medium
- Sender and receiver speak the same language.

10.3 Communication Types

10.3.1 One-way Communication

In the case of one-way communication, information is exchanged in one direction only. The sender transfers its information to the receiver through the joint medium. If, as a consequence of having received information, the receiver answers the sender, this in turn is regarded as separate, one-way communication. The receiver becomes the sender and vice versa.

10.3.2 Two-way Communication

As opposed to one-way information, in two-way communication information is exchanged concurrently. Those taking part in the communication are at one and the same time: Sender and receiver.

10.3.3 Interactive Communication

Interactive communication is a developed form of two-way communication. The concurrent exchange of information interacts with information relating to the last exchange. A dialogue develops between the parties to the communication.

10.3.4 Open Communication

One talks of open communication (open system interconnection) when parties to the communication follow standardised communication rules. By standardising the communication rules, all potential parties to the communication can ensure that they can communicate with as many other potential parties as possible. Hence the possibility of communicating is open to all potential parties.

11. Standards

11.1 Definition of Standard

A standard is understood to be a universally valid definition in answer to a set of facts. Before having been adopted, a standard has always passed through an approved standardisation procedure recognised by law, which is made up of several instances and process steps. The prime criteria for a standard are technical maturity and user benefit.

11.2 Types of Standards

It is important to differentiate between **standards, industrial standards, manufacturer-specific standards** and **recommendations.**

Standards have passed through the defined process of standardisation (see definition of standard).
An industrial standard is a non-standardised procedure or operation in answer to a set of facts, which is manufacturer-independent and has as a rule proved to be correct by having been successfully implemented in practice for many years.
The manufacturer-specific standard differs from the industrial standard in that here manufacturer-specific procedures or operations in answer to a set of facts have been proven and tested in practice.
The recommendation is the most non-committal type of standardisation. Neither manufacturer nor user is obliged to conform with it.

11.3 Important Organisations for Standardisation in IT

11.3.1 International Standardisation Organisation (ISO)

The national standardisation institutes from 118 countries work together in the **I**nternational **S**tandardisation **O**rganisation (**ISO**) based in Geneva. The objective of ISO is to facilitate the international exchange of goods and services by standards of worldwide uniformity and to promote scientific, technical and economic co-operation beyond national frontiers. The ISO is responsible for the best-known standard from the field of data communication and is still developing it: The **OSI reference model**.

11.3.2 International Electronical Commission (IEC)

The **I**nternational **E**lectronical **C**ommission (**IEC**) was founded in 1906 and later affiliated with the ISO. It has more than 600 working groups dealing with standardisation. The standards

drawn up by the IEC enjoy worldwide validity. The IEC is based in Geneva, the national standardisation institutes operate in over a hundred countries. The IEC is a recognised standardisation organisation in the field of electrical engineering and electronics. Its field of reference includes electrical safety, electric strength and insulation, electrical components, installations for generating, transporting and distributing electrical power, domestic electrical installations, telecommunications and radio engineering and information technology.

11.3.3 International Telecommunication Union (ITU)

The **I**nternational **T**elecommunication **U**nion (**ITU**) was founded in Paris by 20 states on 17.5.1865 and since 15.10.1947 has been a suborganisation of the United Nations headquartered in Geneva. The ITU is a globally active organisation in which governments and the private sector of telecommunication coordinate the operation of telecom networks and services.

11.3.4 European Telecommunications Standards Institute (ETSI)

The telecom standards valid for Europe are issued by the **E**uropean **T**elecommunications **S**tandards **I**nstitute (**ETSI**) founded in 1988. The institute has its headquarters in Sophia Antipolis, a research centre in the south of France. The ETSI has approx. 800 members from over 50 countries. The ETSI European standards of telecommunication are mostly based on the corresponding international recommendations (e.g. of the ITU). The ETSI modifies these standards in line with the European requirements.

11.3.5 Institute of Electrical and Electronical Engineers (IEEE)

The **I**nstitute of **E**lectrical and **E**lectronical **E**ngineers (**IEEE**) in New Jersey, USA, addresses itself to standardisation tasks, among others, and is an association of chiefly American engineers. The IEEE is divided into various working groups, which are formed temporarily until a recommendation for a standard has been drawn up. The working group 802, for example, attends to the standardisation of local networks and has made major contributions including such for defining access procedures and security protocols.

12. The OSI Reference Model

12.1 Introduction

The OSI model serves as a reference model for all open and standardised systems. **OSI** stands for **O**pen **S**ystem **I**nterconnection and provides an abstract model as a basis for interpreting existing systems and as a reference for the definition of new protocols. To be able to understand the development of the OSI model, you have to bear in mind that the structure of the IT world before 1980 was non-uniform, governed by manufacturer-specific standards that were incompatible with one another. Any modifications in the interest of achieving compatibility were costly and involved a lot of effort, a situation corrected by the OSI reference model, which is based on an ISO standard and was adopted by the ITU in 1984. The objective of the OSI reference models is the unification of standards to enable communication between open and standardised systems. The OSI reference model is subject to ongoing, flexible updating, i.e. old elements are extracted and new ones added to prevent the reference model from becoming obsolete.

12.2 Structure

The OSI reference model is a layer model, that is to say the overall system is divided into subsystems: Layers. The individual layers conform with a specific structure. Subsystems of the same hierarchy form a layer, the layers are superposed, whereby a lower-level layer serves the communication function of the higher-level layer. Further, each layer provides certain services both within that layer and for lower-level layers.
In the OSI reference model there are seven layers in all, which can be divided into two superordinate categories: The transport-oriented layers 1 - 4, which are responsible for transmitting bits, and the application-oriented layers 5 - 7, which are responsible for the control of the operation and the nature of information presentation.

The layers communicate with one another according to specific rules, the protocols, the functions of elementary communication being called service primitives. The service primitives are classified into four families per layer: Request, Indication, Response and Confirm. In a communication between A (service user) and B (service provider) interconnected via layer N for service provision, A issues a Request for a service from B and the Indication informs B of this. Then comes the Response from B, which represents the acknowledgement of the preceding Indication, and finally the Confirm as the provider's acknowledgement of the Request.

12.3 Layers

Layer 1, Physical Layer
The Physical Layer defines the mechanical, electrical, functional and procedural characteristics of physical links between data terminal equipment (DTE) and data communication equipment (DCE). The bit stream is conveyed via a communication channel. Examples of characteristics will be set forth below. Mechanical: These may be the dimensions of the connectors or the arrangement of the pin structure. Electrical: The voltage level of the circuits, e.g. according to the X and V recommendations of the ITU. Functional: The assignment of the functions of individual pins. Procedural: the general rules for the use of interface circuits.

Layer 2, Data Link Layer
The Data Link Layer has to ensure transparent data exchange. The bits streams are combined to form so-called frames and forwarded to the Network Layer, passing through the processes of error detection and error correction.

Layer 3, Network Layer
The Network Layer realises the data transport between end systems, via any intermediate nodes, by means of routing. The data units are packets in connection-oriented transmision and datagrams in connectionless transmission. Therefore entire networks can be interconnected, the basic requirement being a unique address. X.25, Frame Relay and IP are typical protocols of the Network Layer.

Layer 4, Transport Layer
The Transport Layer realises the data transfer irrespective of the underlying network and provides reliable data traffic. Address-mapping, flow control between hosts with different speeds and multiplexing of transport connections in networks are typical functions of this layer. The best known protocols of this layer are **UDP** (**U**ser **D**atagramm **P**rotocol) and **TCP** (**T**ransmission **C**ontrol **P**rotocol). Layer 4 simultaneously forms the transition to the application-oriented layers.

Layer 5, Session Layer
The Session Layer realises the session, i.e. the communication of processes. Further, it organises the data exchange with added processes such as dialogue and communication control as well as synchronisation in case of systems crashing.

Layer 6, Presentation Layer
The Presentation Layer establishes the communication to application processes of heterogeneous computer systems. It deals with syntax only, the uniform processing and presentation

of the data. This entails, for example, the transformation of different data formats into a standard form, the so-called code conversion.

Layer 7, Application Layer

The Application Layer provides services for the external use of computer networks and is the most comprehensive layer. Possible functions are identifying communication partners, granting access rights and specific service qualities

13. Transmission Methods and Techniques

13.1 Introduction

Data transmission in networks can be categorised according to various characteristics. The character of the **signal**, i.e. analogue or digital, the **synchronism**, i.e. asynchronous/asymmetrical or synchronous/symmetrical, the **character transmission**, i.e. bit-serial or bit-parallel and the **mode** simplex, half-duplex, duplex and full duplex. These various categories of data transmission will be dealt with below. An important point is that these categories cannot be regarded individually in a network, rather more they complement one another and thereby describe the data communication in its entirety.

13.2 Signal

In principle, time-dependent signals can be subdivided into various classes. If a signal value is present at all times, one speaks of **time-continuous**, if a signal value is not present at all times one speaks of **time-discrete**. If all the signal values in the range of values are permitted, one speaks of **value-continuous**, if only certain signal values in the range of values are permitted, one speaks of **value-discrete**.
In the case of an **analogue signal**, the signal can assume any value between the negative and positive maximum value, i.e. it is a time-continuous and value-continuous signal. In analogue data transmission, the information is then represented by constant changes in the voltages. Here the modulation processes of frequency modulation, amplitude modulation and phase modulation come into play.
In the case of a **digital signal**, the signal can assume only certain values, i.e. it is a time-discrete and value-discrete signal. In the depiction of digital information only two characters are used. This procedure is also called dual-system. In this connection, binary signifies that only one of two states is assumed.

13.3 Synchronism

Synchronism applies to the data link layer, is used in synchronous and asynchronous data transmission and is a method for the time control of sending and receiving devices.
In **synchronous data transmission** distinction is made between bit-oriented and character-oriented transmission, whereby exact time intervals between sender and receiver are kept; the synchronisation of the clock generator is carried out via the data edges. A free number of data bits is combined into a frame. The beginning of the frame is marked by a special frame header.

In **asynchronous data transmission** the start-stop principle is employed. That means the frame is defined by a start-bit and a stop-bit with a fixed time slot in between. After the stop-bit an idle state is assumed to prepare for the next transmission.

13.4 Character Transmission

The character transmission differentiates between **bit-serial** and **bit-parallel**. In the case of bit-serial transmission, the bits of the characters are transmitted in time sequence. On the other hand, in bit-parallel transmission the bits of the characters are transmitted at the same time in parallel.

13.5 Operation

In sender - receiver communication in networked IT systems, distinction is drawn between three different modes of operation, these being **simplex**, half-duplex and duplex/full duplex. In the simplex mode two parties are connected through a common communication channel. However, each party can only be either sender or receiver, so that the communication can take place only in one direction. In the **half duplex** mode, two parties are likewise connected through a common communication channel, however here each party can be sender or receiver but can only perform one of these roles at a time. This signifies that prior to communication the parties have to agree who is sender and who is receiver, so that the communication can take place in both directions, but only one direction at a time is possible. In the **duplex** or full duplex mode, both parties are connected through two common communication channels. Each party can be sender and/or receiver at the same time, so that the communication can take place in two directions simultaneously.

14. Personal Area Networks - PANs

14.1 Definition

The term **P**ersonal **A**rea **N**etwork (**PAN**) covers the communication of devices of a single or small number of users within a range of about 10 metres. As a rule three areas can be defined. The linking of peripherals, the linking of external operator devices and the linking of multiple computer systems for data transmission. The latter can be disregarded as it borders on the classical LAN. Since most PANs function wirelessly, one also speaks of **W**ireless **P**ersonal **A**rea **N**etwork (**WPAN**), which is virtually a synonym. The two technologies most often used in this connection are IrDA and Bluetooth, which will be discussed in further detail below.
A further use to which the term **P**ersonal **A**rea **N**etwork (**PAN**) is put is in connection with research by the M.I.T. (Massachusetts Institute of Technology) concerning the transmission of digital information using the electrical conductivity of the human body.

14.2 Transmission Methods

14.2.1 IrDA (Infrared Data Association)

The term IrDA (Infrared Data Association) covers a network concept for short range, line of sight, point-to-point cordless data transfer, also referred to as IrDA Data. IrDA Control, on the other hand, provides the possibility of linking multiple peripherals for point-to-point or point-to-multipoint cordless data transfer without direct line of sight. It must be noted that IrDA Control provides lower speeds than IrDA Data.

14.2.1.1 IrDA Data

Since IrDA Data was passed as a standard in 1994, well in excess of 300 million electronic devices with these interfaces have been in use and the trend is on the up. This is due simply to the fact that there is an increased demand for communication between electronic devices, ranging from established devices of daily use to newer devices emerging from the development of the IT sector.

Protocols
The mandatory protocols of an IrDA Data implementation include: PHY (Physical Signalling Layer), IrLAP (Link Access Protocol) and IrLMP (Link Management Protocol).

Physical Signalling Layer (PHY)
The range for continuous operation is 1 metre to 2 metres. In a low power version, the range

is 20 cm to max. 30 cm for communication between low power and standard power versions. Bi-directional communication is the standard of all specifications, it being possible to attain data rates of 9600 B/s Infrared (IR), 115 KB/s Serial Infrared (SIR), 4 MB/s Fast Infrared (FIR) and 16 MB/s Very Fast Infrared (VFIR). The data packets are checked with a CRC (Cyclic Redundancy Check) procedure (CRC-16 and CRC-32).

Link Access Protocol (IrLAP)
The Link Access Protocol provides a device-to-device connection for reliable, ordered data exchange. It provides functions for establishing and clearing connections and for fault and flow control.

Link Management Protocol (IrLMP)
The Link Management Protocol provides multiple channels via an IrLAP connection. It also provides functions for multiplexing for the IrLAP connection. The Information Access Service (IAS), provides protocol and service discovery, forming the basis for the communication of IrDA components.

Optional Protocols
Further optional protocols are on layers 4 and 5 of the IrDA Data protocol stack, these being, for example, **Tiny TP** which provides flow control on IrLMP connections with an optional segmentation service. **IrCOMM** provides COM (serial and parallel) port emulation for legacy COM applications. Further, **OBEX**, which provides object exchange services similar to WWW Hyper Text Transfer Protocols and IrDA Lite, which provides methods of reducing the size of IrDA code. There is also the **IrTran-P**, which provides an image exchange service used in digital image capture devices/cameras. The **IrMC** protocol plays an important part in mobile communication, providing the specifications for data exchange between such devices. Finally the **IrLAN** protocol describes the interface between IR networks and an LAN (Local Area Network).

14.2.1.2 IrDA Control

IrDA Control is a communication standard enabling wireless peripherals to communicate with many different intelligent host devices.

Protocols
The necessary protocols of an IrDA Control implementation include: PHY (Physical Signalling Layer), MAC (Media Access Control) and LLC (Logical Link Control).

Physical Signalling Layer (PHY)
The range of operation of IrDA Control is comparable with that of present-day uni-directional infrared systems and is at least 5 m. In the same way as with IrDA Data, bi-directional communication is the basis for all standards, a data transfer rate of max. 75 KB/s being reached. The data packets are checked using a CRC (Cyclic Redundancy Check) procedure (CRC-8 and CRC-16).

Media Access Control (MAC)
The Media Access Control enables a host to communicate with multiple peripheral devices, but with a maximum number of eight peripherals simultaneously. Furthermore, a fast response time and low latency are ensured.

Logical Link Control (LLC)
The Logical Link Control ensures reliability features providing data sequencing and re-transmission when errors are detected.

14.2.2 Bluetooth

Bluetooth is a solution for the simple connection of peripherals via a wireless interface. The standard was introduced by the Ericsson company, which gave Bluetooth its name. The Danish King Harald II, who reigned around 1000 years ago and is still held in high regard in Denmark, was called Bluetooth. The standard is promoted by the BSIG (Bluetooth Special Interest Group) and was introduced for the first time in 1998. Along with Ericsson, nine companies form the so-called Promoter Group, while in the meantime a significantly larger number of companies belong to the lobby.

14.2.2.1 Bluetooth - Basics

Bluetooth operates in the 2.4 GHz ISM unlicensed band, the same frequency band as used by the wireless LAN Standards IEEE802.11b and IEEE802.11g. There are three different power classes: Class 3 operates with a sensitivity level of -70 dBm, an ouput power of 1 mW, a theoretical range of 10 metres and a network size in the piconet range. This is the class in which most Bluetooth devices today operate. Even if according to definition this is the range of PAN, Bluetooth is however not limited to this level. The output power can be significantly increased by added power amplifiers. Class 2 operates with a sensitivity level of 4 dBm, an output power of 2.5 mW, a theoretical range of 30 metres and a network size in the nanonet range. However, maximum power is achieved in class 1 with a sensitivity of 20 dBm, an output power of 100 mW, a theoretical range of 100 metres and a network size in the micronet range. Networks of classes 3 and 2 are to be associated with the PAN range, while class 1

networks form a borderline case to the LAN. It must be noted, however, that in the present-day Bluetooth solutions, data rates of 1 MB/s max. can be realised. Future expansions are already being worked on with the aim of doubling the data rate to 2 MB/s.
As already mentioned, Bluetooth operates in the 2.4 GHz ISM spectrum, the exact operating frequency range being from 2.402 GHz to 2.480 GHz. The fast method of **F**requency **H**opping **S**pread **S**pectrum (**FHSS**) is used for access. Frequency is switched 1600 times a second in 79 steps of 1 MHz each. Bluetooth provides a wide-band channel for voice and data transmission, it being possible to operate up to three voice channels with 64 KB/s each. In principle synchronous and asynchronous transmission can be selected. Synchronous transmission provides a data transmission rate of 433.9 KB/s. In the case of asynchronous transmission, a data transmission rate of 723.2 KB/s in the download direction and 57.6 KB/s in the upload direction can be realised. Bluetooth has three security modes, only mode 3 employing encryption methods. When the connection is established, a 128-bit authentication and an 8-bit to 128-bit encryption during data transmission is carried out.

14.2.2.2 Bluetooth - Network

In class 3 the network size is said to be piconet. In a piconet no more than eight Bluetooth devices can communicate with one another. The Bluetooth devices identify themselves using a 48-bit long identification number, the first active device in the network acting as master and being responsible for the control of the frequency hopping. That means that a master and seven slaves at most can operate in a piconet. However, up to 255 Bluetooth devices can also be parked as passive slaves and can then be activated through the master if necessary. Furthermore, it is possible to combine multiple piconets to form a so-called scatternet. In so doing, a common slave acts as intermediary between two piconets.

14.2.2.3 Bluetooth - Advantages/Drawbacks

The Bluetooth technology is an asset in the PAN segment and permits simple connection of peripherals. The advantage of such an open standard is the rapid acceptance on the market, however, this requires qualification as it varies greatly in sub-areas. If the development targets of a very economical, single chip solution for Bluetooth components are met, there will be no obstacle to widespread use. The brand name Bluetooth ensures that an interoperability with Bluetooth devices from different manufacturers can be fulfilled. However, the use of the 2.4 GHz frequency band means that there may be interference with WLAN components operating according to the IEEE802.11b or IEEE802.11g standards likewise in the ISM band. The use of Bluetooth for networked data transmission also has to be critically regarded, as this represents the borderline case from WPAN TO WLAN. In addition the industrial standards according to IEEE802.11 have significantly wider market penetration and a more sophisticated technology.

15. Local Area Networks - LANs

15.1 Definition

In the age of information technology, networks exist that span the globe. The best-known network is the Internet, which is becoming more and more popular. It is important that all of the advantages of a network apply to both small and large networks. Networks that have a scope that is limited to a company's premises or a building are called a **L**ocal **A**rea **N**etwork (**LAN**). A LAN contains at least two networked PCs and can stretch to cover a multitude of PCs and servers. The goals of this network can be divided into four important sections. These are "data interlocking", i.e. access to remote data, and "function interlocking", i.e. access to computers that provide special services, such as servers. The last two are "load interlocking", i.e. the load is spread out over several computers, and "availability interlocking", which guarantees failsafe operation with redundant computer systems. When talking about a LAN, you must differentiate between wired and wireless structures. If it is a wireless structure, then it is also called a WLAN (Wireless Local Area Network). The following will first explain the classic LAN.

15.2 Accessing

Since several computers access the same transmission medium, concurrent and reciprocal access must be regulated. Networked IT systems can be easily classified according to such access modes. Essentially one can distinguish between two access methods: The non-deterministic, random access mode and the deterministic, sequential access mode. CSMA/CD (Carrier Sense Multiple Access / Collision Detection) is the best-known method of a non-deterministic access mode, which is used in Ethernet and 802.3 LAN. Deterministic access modes generally work according to the token principle, e.g. token passing (token ring I), early token (token ring II), time token (FDDI) and polling mode (VG Any LAN). CSMA/CD and token passing will be explained in more detail below.

15.2.1 CSMA/CD

CSMA/CD stands for Carrier Sense Multiple Access / Collision Detection. Carrier Sense (CS) stands for the monitoring of the transmission medium and listening for communication. If a station would like to send something, it waits until the medium is free and then for a further time span, before it starts transmission. Multiple Access (MA) stands for equal, competitive access to the shared transmission medium. Collision Detection (CD) detects collisions by monitoring and listening to the transmission medium. If a station realises that a collision has occurred while transmitting its packages, it sends off a so-called jamming signal and all of the

sending processes are aborted. The station will only try to transmit again once a specific time span and a randomly set time span have both passed. Due to this, Ethernet is called a collision-oriented network, since collisions occur depending on the design of the system and the load on it. However, it is important to realise that CSMA/CD can only function if the doubled maximum round trip delay (RTD) between the stations that are furthest apart is shorter than the transmission time for the smallest permissable package in a 512 bit Ethernet. The 5-4-3 repeater rule has arisen in this context and is still important today. It states that a maximum of five segments with four repeaters and three inter-repeater links (IRL) can be coupled in an Ethernet.

15.2.2 Token Passing

In contrast to CSMA/CD, token passing is a set mode where collisions do not occur at all. Every station can only send if it has a specific bit pattern, the token. This token gives it authorisation to send and is passed from station to station to a physical or logical successor. The maximum sending time is limited, which means that response time is calculable in such networks Here a station checks after the token has been passed if its successor has received the token correctly. If it receives the token without any identification of its successor, then the station sends the token back on its way. If a second attempt also fails, then the system assumes that this station is defective or not active. After the token has checked the next station following this one, the defective or inactive station is skipped and no longer included. Networks based on the token principle are very useful for applications that require guaranteed response times.

15.3 Ethernet and 802.3 LAN

It is important to differentiate between the synonyms Ethernet and 802.3 LAN. Ethernet was originally developed by DEC-Intel-Xerox in 1980 with the CSMA/CD access mode. This standard was taken over by the IEEE (Institute of Electrical and Electronics Engineers) and then released as the ISO standard 802.3 LAN in 1990. The only essential difference between these standards are their frame formats, but they can co-exist in one network. The present-day network components are, in principle, manufactured only according the 802.3 standard, still referred to as Ethernet. The logical topology of an Ethernet is a bus system, whereby the physical topology depends on the various network standards and can also take on other forms, such as a star structure.

Layered models

In the case of an 802.3 LAN layer 2, data link layer, is divided once more into a MAC sublayer and a LLC sublayer. The MAC sublayer (Medium Access Control) controls access to the transmission medium and the allocation of the transmission capacity, whereas the LLC

sublayer (Logical Link Control) sets up the connection to protocols in higher OSI layers. There are three types of LLC sublayers. Type 1 provides a connectionless transmission, type 2 a connection-oriented transmission and type 3 a mixed form with connectionless and connection-oriented transmission.

Packet format

Ethernet is a network that conveys packets, i.e. the data that is to be transmitted is divided into small units, packets or even frames. Every packet searches for its own path through the network. They are only put back together at the end of transmission. The maximum frame length for an 802.3 LAN packet is 1518 bytes, the minimum frame length is 64 bytes. Larger packets are usually required in modern-day networks. These are called jumbo frames, which describe packets that are larger than 1500 bytes. Due to downward compatibility, an increase in packet size has been avoided in current Ethernet standards, which causes the performance of high-speed networks to suffer. Jumbo frames can be up to 9000 bytes, beyond that there are restrictions against larger packets, due to the 32-bit CRC process that is used in Ethernet. CRC is an abbreviation for **C**yclic **R**edundancy **C**heck, which performs a check using generator polynoms that is output in the frame check sequence (FCS) field. Another increase would be useful, but it not expected in the foreseeable future. Components that support jumbo frames are usually identified specifically, however, many components can support jumbo frames to a certain degree.

15.4 MAC Addresses

MAC or LAN addresses are the elementary addresses that are used to clearly identify components in a local network. There are generally two different kinds of addresses, a 16-bit and a 48-bit variant. In the 16-bit address, the first bit identifies whether it is an individual or group address, the other 15 bits are used for further identification. The group address can be a multicast group address or a broadcast address. However, 48-bit addresses are used almost exclusively today. The first bit in these addresses is also used for identification of individual or group addresses. In addition, there is a second bit, which identifies whether it is a global or local address and the other 46 bits are used for further identification. With global addresses, the manufacturers of controller chips have to buy the first 3 bytes of the MAC address from the IEEE. The manufacturers generally have a defined spectrum of addresses that they can use for their controller chips. The remaining 3 bytes of the MAC address are assigned by the manufacturer. This guarantees that the addresses are unique worldwide and that there are not any double addresses, which can lead to complications in a network. Since no one can be forced to buy the addresses from the IEEE, the adresses can also be managed locally.

15.5 Ethernet Standards

There is a variety of implementations from Ethernet, which mainly differ in terms of speed, transmission mode, maximum segment length, as well as connector types. Thus the nomenclature of the standard 10 BASE-5 according to IEEE802.3, refers to an Ethernet with 10 Mbit/s baseband transmission with a maximum segment length of 500 m. The most important standards that are already in practice and will possibly be put into practice are described in more detail below.

15.5.1 10 BASE-5

The standard 10 BASE-5 according to IEEE802.3 refers to an Ethernet with 10 Mbit/s baseband transmission with a maximum transmission section of 500 m. 10 BASE-5 is also known as Thick Net, since it uses a thick RG8 50-ohm coaxial cable as a transmission medium. Thick Net physically uses the bus topology and the 5-4-3 repeater rule has to be taken into consideration. This means that a Thick Net may have a maximum of 5 segments with a maximum segment length of 500 m over 4 repeaters that are connected together. This results in a maximum of 3 inter-repeater links (IRL). A maximum of 100 stations may be connected per segment, whereby the attachment unit interface (AUI) connector is used. The maximum length of the AUI cable is 50 m. The 10 BASE-5 standard has been used extensively for quite some time, but is not included in new installations.

15.5.2 10 BASE-2

The standard 10 BASE-2 according to IEEE802.3a refers to an Ethernet with 10 Mbit/s baseband transmission with a maximum segment length of 185 m. 10 BASE-2 is also known as Thin Net or Cheaper Net, since it uses a thin RG58 50-ohm coaxial cable as a transmission medium. Thin Net physically uses the bus topology, whereby the minimum segment length between two stations is 0.5 m and a maximum of 4 repeaters may be switched between two stations. A maxumim of 30 stations may be connected per segment and the BNC T adaptor is used as the connector. The 10 BASE-2 standard has been used extensively for just a short while, but it is not included in new installations.

15.5.3 10 BASE-T

The standard 10 BASE-T according to IEEE802.3i refers to an Ethernet with 10 Mbit/s baseband transmission with a maximum segment length of 100 m with copper-based cabling. Copper twisted pair cables of various standards are used as the transmission medium along with RJ-45 connectors. The various cable standards will be discussed in a later chapter. The

10 BASE-T physically uses the star topology, i.e. one active component is used to concentrate the stations into a star shape, the concentrator is also used as the amplifier. The 10 BASE-T standard has been used extensively for just a short while, but it is not included in new installations.

15.5.4 10 BASE-FL

The standard 10 BASE-FL according to IEEE802.23j refers to an Ethernet with 10 Mbit/s baseband transmission with a maximum segment length of 2,000 m with fibre-optic-based cabling. Fibre-optic duplex cables are used as the transmission medium, which often use ST connectors. The various cable standards will be discussed in a later chapter (15.9). The standard is an expansion to FOIRL (Fibre Optic Inter Repeater Link) and defines the connection between concentrators, as well as between stations and concentrators. The 10 BASE-FL standard has been used extensively for just a short while, but it is not included in new installations.

15.5.5 100 BASE-TX

The standard 100 BASE-TX according to IEEE802.3u refers to an Ethernet with 100 Mbit/s baseband transmission with a maximum segment length of 100 m with copper-based cabling. Copper twisted pair cables of various standards are used as the transmission medium along with RJ-45 connectors. The various cable standards will be discussed in a later chapter (15.9). The 100 BASE-TX physically uses the star topology, i.e. one active component is used to concentrate the stations into a star shape, the concentrator is also used as the amplifier. The 100 BASE-TX is widely used, very often in new installations.

15.5.6 100 BASE-FX

The standard 100 BASE-FX according to IEEE802.3u refers to a fast Ethernet with 100 Mbit/s baseband transmission with a maximum segment length of 400 m between stations and concentrators and 2,000 m between concentrators. Fibre-optic duplex cables are used as the transmission medium, which often use ST, SC, MT-RJ or VF-45 connectors. The various cable standards will be discussed in a later chapter. 100 BASE-FX is used with the fibre distributed data interface (FDDI), which works on the basis of the time token process. The 100 BASE-FX standard is widely used, also in new installations in LWL environments.

15.5.7 1000 BASE-T

The standard 1000 BASE-T according to IEEE802.3ab refers to a Gigabit Ethernet with 1000 Mbit/s baseband transmission with a maximum segment length of 100 m in the terminal section. Copper twisted pair cables of various standards are used as the transmission medium along with RJ-45 connectors. The various cable standards will be discussed in a later chapter (15.9). The 1000 BASE-TX physically uses the star topology, i.e. one active component is used to concentrate the stations into a star shape, the concentrator is also used as the amplifier. The 1000 BASE-T standard is an expansion to the 100 BASE-T2 standard and the 100 BASE-T4 standard, which specify characteristics for transmission over category 3 copper cables and in this case use more than two pairs of wires. In order to reach 1000 Mbit/s, 250 Mbit/s are transmitted over ever pair of wires. This standard is becoming more and more widely accepted and is used in new installations.

15.5.8 1000 BASE-SX

The 1000 BASE-SX standard according to IEEE802.z refers to a Gigabit Ethernet with 1000 Mbit/s baseband transmission via short wavelength. This means that a wavelength of 850 nm is used and, depending on the fibre-optic cable, can bridge a distance of 275 m maximum for 62.5/125 micron multimode fibres and 550 m maximum for 50/125 micron multimode fibres. This standard uses a point-to-point connection with CSMA/CD. SC connectors are commonly used, but MT-RJ and VF-45 may also be used. Besides the copper match, this standard is widely accepted and used in new LWL installations.

15.5.9 1000 BASE-LX

The 1000 BASE-LX standard according to IEEE802.z refers to a Gigabit Ethernet with 1000 Mbit/s baseband transmission via long wavelength. This means that a wavelength of 1300 nm is used and, depending on the fibre-optic cable, can bridge a distance of 550 m maximum for 62.5/125 or 50/125 micron multimode fibres and 5000 m maximum for 9/125 micron single mode fibres. Larger distances can be bridged, but these are usually manufacturer-specific solutions that are only compatible to each other. This standard uses a point-to-point connction and without the use of CSMA/CD. SC connectors are generally used. In addition to the 1000 BASE-SX, the 1000 BASE-LX is a solution for structured cabling in the LWL primary sector and has gained acceptance in this area.

15.5.10 Auto-Negotiation

The **A**uto-**N**egotiation-**P**rotocol (**ANP**) was created within the framework of the Fast Ethernet standard and is used when devices need to communicate with either 10 Mbit/s or 100 Mbit/s. The ANP is based on the Nway protocol from National Semiconductor, which is also often referred to in descriptions. The protocol automatically sets the highest possible speed that can be used for communication between the connected partners. The type of operation is also selected, either half-duplex or full-duplex, as described in the chapter on transmission methods. The ANP is optional for 10 BASE-T and 100 BASE-T. It cannot be used with 100 BASE-FX, since interoperability cannot be guaranteed during optical transmission, due to the different wave lengths. The ANP is required with the 1000 BASE-T. ANP problems can develop if the stations do not react to the sent control packets and the half-duplex operation is automatically set. If the communication partner should now be manually set to full-duplex, the connection cannot be opened.

15.6 Ring Topologies

In the ring topologies, the data is serially transmitted in one direction and usually sent from station to station sequentially. There is generally one station that takes over controlling in the ring. This structure will automatically cause the entire network to fail if one station fails, which can be avoided in practice by using concentrators or double ring structures. In ring topologies the time for transmission increases proportionally to the addition of more stations, which, on the one hand, results in a response time that can be calculated, but, on the other hand, often only allows a lower transmission speed.

15.6.1 Token Ring

A token ring LAN according to IEEE802.5 places a logical ring topology on a physical star topology. This means that the failure of one station does not lead to the failure of the entire network. The concentrator in such a structure is called multistation access unit (MAU). Transmission rates of 4 Mbit/s can be reached in the first token ring, but, in contrast, rates of 16 Mbit/s are possible in the second ring. Token ring I works according to token passing, token ring II according to the early token procedure. Token ring has this name due to a specific bit pattern, the so-called token, which goes around in a ring. Only the station that has the free token can transmit. The station that wants to transmit goes from the listening mode to send mode once it has received the free token. The data to be transmitted is attached to the free token, which is then occupied. This means that no other station can change its status while the token is occupied and they stay in listening mode. The station that is in send mode transmits until the transmission is complete or a specific time span has been reached; set

response times can be kept here. Once the transmission is finished, the station goes back to listening mode and releases the token into the ring, so that other stations can transmit. The maximum frame length in token ring I is 4500 bytes, the maximum number of stations per segment is 260 and the maximum expanse of cabling per segment is 150 m for UTP and 250 m for STP cabling. In comparison, the maximum frame length in token ring II is 16000 bytes, the maximum number of stations per segment is 188 and the maximum expanse of cabling per segment is 100 m for UTP and 150 m for STP cabling. There are, however, various proprietary manufacturer-specific standards that must be observed, also the types of cables are a decisive factor. A typical error often made with token ring LANs is including stations with different speeds. If this happens, the station does not receive a token that is valid for it. This causes a station to miss a beacon, referred to as beaconing, and disrupts the network. In addition to the token ring, there is also the token bus standard according to IEEE802.4, where a logical ring topology is placed on a physical bus topology. Token bus is not commonly used.

15.6.2 FDDI

FDDI stands for **F**ibre **D**istributed **D**ata **I**nterface and is an ISO Norm that has been generated from ANSI X3T9.5. FDDI is a standard for LANs and MANs and is based on a LWL double ring structure. It can reach data transmission rates of 100 Mbit/s. There are two types of stations that can be connected, class A stations and class B stations, also referred to as DAS (Dual Attached Stations) and SAS (Single Attached Station). Class A stations have two interfaces and can be redundantly connected to the LWL double ring, whereas class B stations have one interface. There is a primary and secondary ring in the standard configuration, which ensures that the data transmission continues if a cable or station fails. In this case, the secondary ring is activated and creates a new and complete ring with the primary ring. However, to do this, the affected stations that are directly connected to the ring have to be class A stations. In this manner, failsafe operation is provided by the double ring structure and reconfiguration during a disruption. FDDI is especially suited for use in MANs, since the maximum distance between two FDDI stations is 2 km with multimode fibres and 10 km with mono mode fibres, the maximum distance that can be bridged is 100 km. Although FDDI was originally created for fibre-optics, there is now a copper-based solution called CDDI (Copper Distributed Data Interface). With CDDI you can only reach a typical distance of 100 m between the individual stations.

15.7 Protocols/Standards

Every station that is connected to the network needs both the appropriate network driver and also a protocol driver, which creates the connections to the higher protocol layers. LAN protocols cover an entire family of protocols and not just one individual protocol. A protocol is necessary for communication between two instances and it presents, in principle, the mutual language from both instances. Several important LAN protocols will be presented below.

15.7.1 Protocols

15.7.1.1 NetBIOS

NetBIOS is an acronym for **Net**work **B**asic **I**nput **O**utput **S**ystem and was developed by IBM in 1984. This protocol is hardware-independent and can thus be used in both Ethernet networks and Token Ring networks. It implements parts of layer 2 and makes functions from levels 2

to 5 available. Layer 3 is only implemented as a zero layer and so this protocol is not routing-capable. Back then, routing was not so decisive as it is now. NetBIOS allows communication via symbolic names instead of networks, but the number is limited.

15.7.1.2 NetBEUI

NetBEUI is an acronym for **Net**work **B**ios **E**xtended **U**ser **I**nterface and was developed by IBM in 1985. It is an expanded user interface, which is based on NetBIOS. That is the reason why this protocol is also not routing-capable, since, as with its predecessor, layer 3 is only implemented as a zero layer. NetBEUI provides high perfomance in local networks with a maximum of 200 stations and is set up as the standard protocol in many operating systems. If communication beyond layer 2 is necessary, then other protocols come into play, such as TCP/IP, which is described in more detail in the WAN chapter.

15.7.1.3 NDIS

NDIS is an acronym for **N**etwork **D**river **I**nterface **S**pecification **S**upport and is a programming interface between layer 2 and the protocols based on it. Here NDIS works as a protocol-independent interface that functions as an agent betweent the network driver and the protocols from other levels.

15.7.1.4 ODI

ODI is an acronym for **O**pen **D**ata Link **I**nterface that was developed by Novell and ist suitable for NetWare protocol stacks. ODI works on level 2 and simulates a virtual network card, which allows various packet formats to be sent. ODI is hardware-independent and is the interface with protocols on higher levels.

15.7.2 Standards

A standard is understood to be a universally valid definition in answer to a set of facts. The standard has passed through an approved standardisation procedure and, in the end, forms a technically mature solution. The LAN standards from the **IEEE** (**I**nstitute of **E**lectrical and **E**lectronical **E**ngineers), an American committee for the creation of standards by engineers, which provides internationally valid standards, will be explained below. Work group 802 is responsible for standards for local networks. These standards make a very valuable contribution to the IT industry. The standards according to IEEE802.1, which cover extensive standards for LAN/MAN bridging and management are of special interest.

15.7.2.1 VLAN according to IEEE802.1q

VLAN stands for Virtual LAN and allows you to set up broadcast domains in an Ethernet that are separated from each other. Stations that are in a VLAN can only communicate with partners within the VLAN. If communication outside a VLAN is necessary, then active components have to be used that work in level 3. These components can be routers or multilayer switches. VLANs make it possible to limit traffic within an Ethernet, in order to increase performance throughout the entire network. VLANs are usually used to group together different units, such as departments within a company. VLANs according to IEEE802.1q and also VLAN tagging, work on level 2 and mark the Ethernet frames, that are sent by the individual stations with special tags, in order to assign them to specific VLANs. This means that they are logical and not physical VLANs, as is the case with a port-based VLAN. Thus, the administrator has considerably more flexibility, because as soon as components are being used that support VLAN tagging, every station on the device can be assigned to a specific VLAN. If, for example, the network structures change within the local network, this has no effect on the device.

15.7.2.2 CoS according to IEEE802.1p

CoS stands for **C**lass **o**f **S**ervice and is a way to divide the traffic within a LAN into various service classes, such as e-mail, file, transfer, voice, video, etc. Every class has its own priority,

which is then taken into consideration during queuing and processing. It is important, however, not to mix up CoS with QoS (Quality of Service). CoS cannot assign bandwidths, which is possible with QoS, but prioritises packets that are then sent according to the best effort principle. CoS works on level 2 and offers a simple way to use resources better. CoS marks the Ethernet frames by means of special tags, which have been assigned to the various classes. In addition to CoS, this technology is also used by ToS (Type of Service), which provides another class of service possibility besides DiffServ (Differentiated Services). DiffServ works in a different manner and analyses the PHB (Port Hop Behaviour) in order to be able to use the available bandwidth more efficiently.

15.7.2.3 Spanning Tree according to 802.1d

If more than two bridges are being used in a local network, the spanning tree protocol (STP) regulates the communication so that it only goes over one bridge and does not create so-called bridge loops. Bridges separate collision domains by analysing the traffic and deciding on which branch in the network the target address of the receiver is located. This way the traffic is not spread out over the entire network, but is limited to individual branches. The bridges are also called transparent bridges, since they analyse the network themselves. Individual hardware components in local networks are described in more detail in the hardware chapter. Since there are usually several bridges in a network and segments are connected in a variety of different ways, bridge loops may occur, which can have a lasting influence on traffic. The STP controls that the bridges communicate with each other using BPDUs (Bridge Protocol Data Units) and that only the most efficient section is used between two stations, and all other sections are not used. If there is a failure in a section, the network is reconfigured using an STP algorithm, in order to find the best section of the remaining sections. A disadvantage of the STP is the long time needed for reconfiguration, which can be from 30 to 60 seconds. The IEEE has solved this problem with RSTP (Rapid Spanning Tree Protocol) according to IEEE802.1w, which has substantially shorter reconfiguration times in comparison to STP, but which cannot be used in token ring networks.

15.7.2.4 Link Aggregation according to IEEE802.3ad

Link aggregation is a method to group together several physical links into just one logical link, in order to increase the bandwidth and also guarantee failsafe operation. This way the failure of a link between two communication partners does not lead to a complete failure of a connection. The connection can be maintained, whereby only the bandwith is reduced overall. The possibility to get higher bandwidths through aggregation is a simple way to use existing technology to attain higher data rates without having to modify the complete network technology. Link aggregation is used with copper- and fibre-optic-based 100 Mbit/s and

1000 Mbit/s Ethernet networks, whereby aggregation of Gigabit Ethernet links is especially important to be able to close the gap up to an economically feasible 10-Gigabit Ethernet network. The terms trunking and port trunking were often used in the past, but these terms are already used to describe regional trunked mobile networks in English. This is why the work group selected the term link aggregation.

15.8 Hardware - Active Components

15.8.1 Repeaters

A repeater is an active component that works on layer 1 in an Ethernet LAN and performs regeneration functions. A repeater is used if the maximum physical expanse of a segment has been reached and should be extended beyond that. The repeater performs, as already mentioned, regeneration functions, which means that it regenerates the level and cycle during signal progression. In addition, the repeater has a certain capacity for error detection, i.e. it can limit faulty signals or collisions to a specific segment, which leads to an overall increase in performance. Since a repeater works on layer 1, it is completely protocol transparent and is only used to extend the segment beyond the physical limits of the topology.

15.8.2 Hub

A hub is basically a multiport repeater and an active component that works on layer 1 in an Ethernet LAN. The hub is the concentrator in a star topology and can connect various stations and segments using different connectors and transmission media. Hubs are differentiated according to the location where they are used, e.g. workgroup, departmental, or enterprise hubs. The size and extendability of hubs also depend on where they are used. The performance of a hub is largely determined by its internal busses, which form the backplane.

15.8.3 Bridge

A bridge is an active component that works on layer 2 in an Ethernet LAN. Most bridges in Ethernet LANs usually work on the MAC sublayer and very seldom on the LLC sublayer. The devices often work, however, on the LLC sublayer when different topologies are used, e.g. token ring, since a bit more intellegence is required here. Bridges are used if the maximum physical expanse of a segment has been reached and should be extended beyond that. It takes over regeneration functions and also has error detection functions. Furthermore, with the help of a bridge collision domains are separated from each other and data packets are analysed on the basis of their MAC addresses. Over time, a bridge learns which stations are located in which segments and decides whether it is necessary to expand the traffic to the entire

network. If the sender and the receiver are located in the same segment, then communication is limited to just this segment. These are known as learning bridges, since they learn the addresses of the local network on their own. If a learning bridge uses the spanning tree protocol (STP), then they are known as transparent bridges. Transparent bridges use the STP to find a loop-free communication path in a network. Loops can arise if several bridges are being used and if there are various paths for data transmission. Transparent bridges are essential, since such loops can bring data communication to a standstill. This name also refers to the transparency found here, with regard to the connected stations in a network. This is based on the fact that the bridges do not change the packets and only work on layer 2 in a LAN. Besides these types, there are additional types of bridges, such as encapsulation bridges, translation bridges and source routing bridges. The encapsulation bridge encapsules the frame of a protocol in the data area of another protocol, in order to transmit data over sections where this type of protocol is not supported. In this manner, an Ethernet frame, for example, can be transmitted over an FDDI network. This process is often referred to as tunneling. A disadvantage of this process is that a protocol overhead develops due to this nesting, and bandwidth is then wasted on the transmission medium. Translation bridges translate the addresses from one type of frame standard into another frame standard. In order for communication to take place, the frame types must be adjusted in terms of minimum and maximum lengths, so that both standards can work with these frames. Source routing bridges are used to facilitate transmission between different token ring networks.

15.8.4 Switch

A switch is basically a multiport bridge and an active component that works on layer 2 in an Ethernet LAN. Switches perform, as the name implies, switching functions and are used as the concentrator in networks with star topologies. The core of a switch is the switching fabric, which determines the performance. The forwarding rate and filtering rate also contribute to the performance of a switch. The forwarding rate describes the maximum number of data packets that can be forwarded per second whereas the filtering rate indicates the maximum number of processed data packets and MAC addresses that can be administered. Switches are additionally divided according to the switching process into cut-through, fragement-free and store-and-forward switches. With cut-through switches, the data packets are simply sent through to keep the switch delay time as short as possible. After the target address has been determined, the cut-through switch does not wait for the transmission of the data to be completed, but immediately forwards the next data. Other than the target address, no other information in the packet is necessary to complete this process, which can dramatically reduce the delay time. In the fragement-free process, the switch waits for a specific time period in order to avoid collisions in the target segment. The most common process is the store-and-forward process where the data in the switch is buffered before it is forwarded. The ad-

vantage of this is that the data goes through an error detection and error correction process, whereas the disadvantage is that a longer delay time develops.

Layer 3 Switching

The term layer 3 switching and other synonyms for this technology are being used more and more. This paragraph should clarify this technology a bit more. There is an entire group of synonyms, other than layer 3 switching, which are used for this technology. This is why terms such as switching routers, layer 4 switching and multilayer switching are also used. However, all of these terms describe the same technology, but the term layer 3 switching will be used below. It is a mistake to believe that a network's performance can be increased purely by the bandwidth or the packet switching performance. An advantage of layer 3 switches is that the decisive components, besides packet switching, are the route processing and intelligent network services. These three components are the foundation on which a layer 3 switch is built. A layer 3 switch is a high-performance router that is especially used on local networks. In contrast to a classic layer 2 switch, it also works on the higher layers in the OSI reference model. The physical implementation also differs from that of classic routers. Thus with classic routers, packet switching is based on microprocessor-based engines, but layer 3 switches are based on ASIC hardware-based technologies (ASIC=application specific integrated circuits). For this reason, classic routers work in the Kpps range and layer 3 switches in the Mpps range. Besides these different implementations, classic routers and layer 3 switches have a lot in common in one respect. Using route processing, the hardware-based packet switching of a layer 3 switch is turned in the right direction. Route processing is a software-based implementation that should, however, be regarded as independent from packet switching. Protocols such as RIP, RIPv2, OSPF, OSPFv2 are used with route processing, which control the packet and route and then create routing tables that depict the constantly changing network structures. A layer 3 switch that is not equipped with the appropriate routing protocol is, if you prefer to see it this way, a device without a leader. In addition to packet switching and route processing, the intelligent network services make up a large part of the advantages in a layer 3 switch. They are so-called value added services, such as management, security, IP multicasting, and quality of service (QoS). The quality of service aspect is especially emphasised in this context. QoS covers all processes that have such an influence on data traffic, that the service arrives at the receiving end with a set quality level. The fact that QoS plays an ever growing role can be traced back to the type of data that has to be transmitted. In the past, file transfer and e-mail were the primary types of data, but, in the meantime, applications from the voice and video sectors have become more and more important. These applications require specific service qualities in terms of latency, rate of packet loss and bandwidth. There many different solutions from institutes for standardisation on how to implement QoS. At the moment, there are very few QoS models that are uniform across all manufacturers, and quite often a proprietary solution from a specific manufacturer

is preferred, which limits flexibility in terms of network equipment. Using intelligent network services, the overall performance of a network can be further increased, independent from additional increases in the bandwidths. Due to this and all the reasons described above, layer 3 switches will become more prevalent in the future.

Cascading/Stacking

The terms cascading and stacking are often used in conjunction with solutions in the switching sector. Both terms describe an option of connecting one or more switches to one unit, but in different ways. The basic difference between cascading and stacking is that in cascading the switches are connected by standard ports and by special stacking ports in stacking.
In cascading this can lead to the interface between both devices becoming a bottle neck in the configuration which has unfavourable effects on network performance. Stacking, in comparison, uses ports that are provided especially to connect several switches. Due to this, both devices can be connected to each other with the complete bandwidth and form a unit. The stacking solutions from the manufacturers make it possible to communicate with the physical switch stack as a logical switch via an IP address. This makes it much easier to do maintenance work on the device. You should, however, note that stacking solutions are manufacturer- and product specific. Other than classic stacking, many manufacturers provide solutions where you can administer many switches that are dispersed across the network in just one stack, even if they are not physically connected. But this is a borderline case to classic stacking.

Modular Switches

In terms of hardware, you can differentiate between two basic types of switches. One of them are fixed configuration switches, that are set by their port structure and are primarily used in the edge area. The other type are modular switches, which have a freely configurable port structre and are mostly used in the core area.
Many of the switches that are used in the edge areas of local networks can be expanded to modular switches. These switches have a fixed port configuration with additional extendable module slots, that can be equipped with various interface modules. Since a variety of modules are usually on offer, the installation can be carried out flexibly with Ethernet standards and media interfaces. There are two common types of modules. One is an uplink module that is usually only suitable for a specific device from a specific manufacturer. Additionally, network manufacturers are making every effort to develop a uniform module standard. The result of these efforts is the GBIC, which stands for **G**iga**b**it **I**nterface **C**onverter. GBIC modules are Gigabit Ethernet modules, that can basically be integrated into all GBIC module slots. However, different manufacturer-specific developments have led them to a point where general interoperability cannot be guaranteed, in spite of uniform module standards and form factors. Even edge switches can be flexibly connected with these modules.

Modular switches that can be found in the core area, can, in contrast, usually be modularly configured. Consisting of a chassis, switch controller slots and interface slots, a variety of port configurations can be implemented. Due to the central function of such switches, the existing components are normally redundant to prevent a total failure of the network.

15.9 Hardware - Passive Components

When passive components are mentioned in conjunction with network technology, they are talking about the network cabling. Since network technology is constantly changing, demands on the physical transmission medium are also changing. Selection criteria for cabling are immunity to interference, options for expansion and, above all, the distance that has to be bridged, the transmission speeds that have to be implemented, as well as the cost of the cabling. Cabling must provide investment security and also be suitable for use with future standards. The highest demands must be taken into account when selecting the passive components, since they are the foundation for a high-performance network.

15.9.1 Structured Cabling

A structure model for universal cabling methods has been developed in cabling technology. In this case, structured means that as few transmission media as possible have to be able to transmit as many different applications as possible. Implementation of this is carried out by creating various steps in a hierarchy. The **primary section** is cabling that is not limited to just one building, which is often provided by means of redundant cable lines that begin or end at the building distributors. The primary section generally uses fibre-optic-based cable technology, due to the distances that have to be bridged. In the **secondary section** the building distributors are connected to the floor distributors, which is why the term vertical section is also used when describing this. Last of all, in the **tertiary section** the floor distributors are connected to the terminals, which is why the term horizontal section is used here.

15.9.2 Backbone Concepts

Two important backbone concepts have arisen from the development of structured cabling. The backbone is literally the backbone of a network over which all of the data flows to the connected stations.
The concept of the **collapsed backbone** compresses all of the floor backbones on one powerful concentrator; all stations are directly connected with this concentrator. Advantages of this are that there are no more active components that could cause delay times and every connected station has, in theory, access to the complete bandwidth of the connected section.

With this centralising, costly building structures with distributor rooms are not as necessary and a redundant network can be created with less effort and installation costs. However, even this concept has some disadvantages. A lot of cabling is required, since every station has to be individually connected. Due to this, the central concentrator also has to have many ports, at least enough to correspond to the number of devices that have to be connected. Additionally, this structure creates a single point of failure.
The concept for a **distributed backbone** is worked out differently. Here many floor distributors bundle the horizontal section and run the floors to the central building distributor. An advantage of this is that the sections are more fail-safe due to the decentral structure. Plus much less effort is needed for cabling and it can also be used in the horizontal section with more inexpensive copper cabling. However, the number of active components, which bring about delay times is a disadvantage. The bandwidth that is available at the terminal also decreases according to the number of stations that are connected to the floor distributor. Another problem is that structural changes can only be accomplished with a lot of effort.

15.9.3 Copper-Based Cabling

Copper cables are widely used and many Ethernet standards require them. The benefits of copper-based cabling are fast, flexible, and inexpensive installation. However, the disadvantages are that they are sensitive to electromagnetic influences and shorter distances that can be bridged. This is why copper-based cabling is often used in the tertiary and secondary sections. In practice, two transmission media are widely used, the coaxial cable and the twisted pair cable. Coaxial cables are used for the 10 BASE-5 and 10 BASE-2 Ethernet standards. 10 BASE-5 uses thick RG8 cable with AUI (Attachment Unit Interface) connectors and 10 BASE-2 thin RG58 cable with BNC (Bayonet Neil Concelman) connectors. The cable is set up as follows.
There is a core line that is surrounded by PVC or Teflon insulation, as well as by a copper or aluminium line. All of this is then covered by an outer shield. Twisted pair cables for all newer copper-based Ethernet standards since the 10 BASE-T standard are used and have a RJ-45 connector. Twisted pair is the designation for a copper cable with one or more pairs of twisted lines. The copper-based Ethernet standards have four twisted pairs of wires or a total of eight wires. Almost all services require two pairs of wires for transmission. Twisted pair cables are differentiated according to the type of shielding and standardised categories. With regard to the shielding, the following types of cables are distinguised from each other. UTP stands for **U**nshielded **T**wisted **P**air where neither the individual wire pairs inside nor the entire cable is shielded. S/UTP stands for **S**creened / **U**nshielded **T**wisted **P**air, where the individual wire pairs inside are not shielded, but the entire cable outside is shielded. STP stands for **S**hielded **T**wisted **P**air where the individual wire pairs inside are shielded but the entire cable is not shielded on the outside. S/STP stands for **S**creened / **S**hielded **T**wisted **P**air where the individual

wire pairs inside are shielded as well as the entire cable outside. Furthermore, copper cables are divided into various performance classes. These performance classes define the minimum requirements of all electric characteristics in a transmission section. The common categories are category 3 with a bandwidth of 16 MHz according to EN 50173, category 5 with a bandwidth of 100 MHz according to EN 50288-2-1, category 6 with a bandwidth of 250 MHz according to the Draft International Standard July 2002, category 7 with a bandwidth of 600 MHz according to EN 50288-4-1. Category 3 copper cables are suitable for the Ethernet standard 10 BASE-T, category 5 copper cables are suitable for the Ethernet standard 100 BASE-T, category 5 enhanced copper cables, and especially cables from categories 6 and 7, are suitable for Ethernet standard 1000 BASE-T.

15.9.4 Fibre-Optic-Based Cabling

Fibre-optic cables are widely used and many Ethernet standards require them. Benefits of ibre-optic cables are insensitivity to electromagnetic influences, high transmission rates, as well as the ability to bridge large distances. Disadvantages are that the cables are not very flexible and difficult to install, as well as the expensive active and passive LWL components. This is why copper cabling is generally used in the secondary and primary sections. The fibre-optic cable is composed of a plastic jacket divided into secondary and primary buffers that protect the fibre-optics. The light signals are transmitted in the core, which is surrounded by a glass cladding. The difference between the fibre-optics in the core and the jacket is their refractive index. Two types of fibre-optics are differentiated. These are the multimode fibres and the single or monomode fibres. With multimode fibres, the light is transmitted in several modes. With the multimode step index fibre, the fibre-optic core has a higher refractive index than the fibre-optic jacket, so that a total reflection of the light occurs between the core and jacket. On the other hand, the multimode gradient strand has a refractive index that is shaped like a parable, which results in wave-shaped spreading of the light. Multimode gradient strands are used primarily today. With single mode fibres, the light is transmitted in one mode. Due to the very small core diameter of a single mode fibre, the light can spread in an almost straight line, which results in lower attenuation and thus makes it possible to bridge larger distances. A decisive criterion for differentiating between fibre-optics is the diameter of the fibre-optic core and jacket. There are two common diameters that are used with multimode fibres: 62.5/125 and 50/125 micron. In contrast, single mode fibres only have a core diameter of 9/125 micron. A variety of transmission sections can be bridged, depending on the Ethernet standard and the transmission medium that is used. Using 1000 BASE-SX, 220-275 m can be bridged by using 62.5/125 micron fibres or 500-550 m can be bridges using 50/125 micron fibres. With 1000 BASE-LX, 550 m can be bridged using 62.5/125 and 50/125 micron fibres, or 5000 m with 9/125 micron fibres. A variety of connectors are used, such as ST, SC, MT-RJ or VF-45 connectors, which are not dependent on the type of optic fibre used.

15.10 Management

Network management is an essential tool for network administrators to control complex local networks. The simple network management protocol (SNMP) is decisive to accomplish this. SNMP was developed in 1988 and is the foundation for administration in TCP/IP networks; it is also based on the application-oriented layers in the OSI reference model. The goals during the development of SNMP were software and hardware independence, as well the mimimum use of system resources. An SNMP environment is made up of a managment station, on which the corresponding SNMP management software is installed, and the network components that are equipped with the appropriate SNMP agents. The SNMP agents rely on the management information base (MIB), in which the characteristics of the network components are described in order to guarantee that the management station only requests data that the agent can deliver. The main tasks of the management are monitoring, controlling, diagnosis and problem analysis. The management station requests the data from the SNMP agents in order to convert them appropriately so that the administrator can evaluate them. There are now three different versions of SNMP, which are constantly readjusted to fit current demands. There is also a second version of MIB, MIB-II, which supports a variety of new protocols that deliver even more detailed information. Due to the object identifier, MIB-II is downward compatible to MIB-I. We also have to mention RMON, which stands for **R**emote **Mon**itoring. RMON is a MIB that was developed to enable proactive network management. Information has to be called up regularly in MIB, which could lead to an increase in the network load. With RMON, the information is collected and saved by the device itself. You can also set limits with RMON, so that a message is only sent if this limit is exceeded. This also reduces the network load considerably. Due to the fact that RMON has to analyse every frame, delay times, which have a negative influence on the network performance, may occur. RMON1 works on the levels 1 and 2 and RMON2 works on levels 3 to 7 in the OSI reference model. RMON2 does not supersede RMON1, but is an expansion to achieve transparency for higher protocol layers. All of these components make it possible for the administrator to manage local networks, which is essential in complex structures. However, you must note that the use of management components can also lead to delays in the network and pay attention to the physical principle that every reading changes the current status.

16. Metropolitan Area Networks - MANs

16.1 Definition

Networks which, in terms of geographic breadth, are between local-area networks (LANs) and wide-area networks (WANs) are called **M**etropolitan **A**rea **N**etworks (MANs).
From the conceptual viewpoint, a MAN is a network interconnecting individual networks, generally individual LANs. By connecting these LANs to form a MAN, all the advantages inherent in LANs can be exploited for a larger area.
Regarded from the specific standpoint, a MAN is in turn an individual network, delimited against other networks, which is set up by an operator and is used for providing services.
From a geographic viewpoint, a MAN extends as far as 150 kilometres, but at least over 2 kilometres. MANs are typically set up for conurbations, towns and their facilities, or as campus networks for universities.

In addition to classical services of data transmission, the task of a MAN consists in providing value-added services such as IP telephony, IP image and video transmission or economical routing for all connected users. In addition, virtual local networks can be set up in a MAN for cross-location and cross-facility purposes.

SDH today forms the basis of town networks, a technology for logical cabling and for establishing logical connections between the users. Since SDH is merely a technology for providing lines and logical connections, but is not a transport technology, a technology such as ATM has to be employed if a service is to be offered.
Thanks to the FDDI, DQDB or ATM high-speed technologies used, service provision in a MAN is realised with very high quality and, above all, shorter transmission time than in a WAN. A further technology is meanwhile venturing into the high-speed regions of the MANs: Ethernet, thanks to its latest stage of development: 10-Gigabit Ethernet.

16.2 Transmission Technologies for Setting Up a MAN

16.2.1 Distributed Queue Dual Bus (DQDB)

DQDB was defined by the working group IEEE 802.6 in the framework of MAN standardisation and can be used for distances of up to 100 kilometres and as the backbone of a MAN.

This method permits both asynchronous and isochronous transmission of coded language, for instance. The concept of DQDB is that two optical fibre unidirectional buses transmit in opposite directions. Both star and ring topology can be implemented. Each node in a DQDB

ring structure is connected to both buses and communicates with other nodes by sending information on one of the two buses. The selection of the bus is based on the direction of communication. At the lowest level the double bus supports full duplex connections, as one bus in each case is used as read or write connection.
G.703 with the speed of 34 or 139 Mbit/s or SDH with 155 Mbit/s according to the G.707, G.708 and G.709 standards are proposed by the IEEE as standardized transmission systems through DQDB.

The FDDI and ATM transmission technologies, as well as G.703 and SDH together with the corresponding standards are also used in LANs and WANs and are therefore discussed in the respective chapters.

17. Wide Area Networks-WANs

17.1 Definition

Geographically dispersed networks are called **W**ide **A**rea **N**etworks (WANs) and provide communication solutions for organisations or private parties. The geographic range can extend to 1,000 or even 10,000 kilometres.
The basic technology employed is either data packet switching or circuit switching. WANs often consist of multiple individual networks interconnected, for example, through ISDN, X.25 or ATM. Secure transmission in WANs is today realised through Virtual Private Networks (VPN). An example of a WAN without which life cannot be imagined is the analogue network in use for decades for telephony. In this network data is transmitted at a transmission rate of 56 Kbit/s. Due to increased demands on transmission rates in WANs due to new services such as video and voice via IP, transmission rates of up to 622 Mbit/s are meanwhile realised with the use of ATM. With the development of 10 Gigabit Ethernet, the migration into even faster transmission rates is just round the corner.
In Europe WANs are usually operated by telecommunication providers.

17.2 Addressing in a WAN

Addressing in WANs is based on the Internet Protocol (IP) addresses assigned to the active network components connected to the WAN via the **netid** and to the hosts via the **hostid**. IP addresses are unique and in IP version 4 consist of 32 bits subdivided into four octets for simpler handling.
Two IP addresses are however excluded from this addressing scheme:
The IP address in which all the bits of the host ID are set to 0 and the address in which all the bits are set to 1. The first one serves as the general description of a network address, the last one is a reserved broadcast address.
The **Network Information Centre (NIC)** manages and distributes IP addresses. The NIC allocates the netids, the hostids are allocated by the respective owner of a netid.

Distinction is made between three classes of addresses:
Class A addresses can address 128 networks, because the first octet forms the network address but the first bit in this octet is 0. The other three octets are for assignment to hosts, which corresponds to the theoretical possibility of assigning unique addresses to 16,777,214 hosts each.
Class B addresses can address 16,384 networks with 65,534 hosts each. The first two bits in the first octet of IP addresses in Class B networks are defined by 10 and thereby characterise Class B networks as such. Generally the first two octets form the netid, the last two octets

form the hostid.
Class C addresses can address 2,097,152 different networks with up to 254 hosts each. In Class C networks the first three bits of the first octet indicate 110 for unique identification. Only the fourth octet is used for assigning IP addresses to hosts, the other three octets serve for assigning netids.

Due to the explosive development of the Internet, freely available addresses according to IPv4 are meanwhile running low and therefore a unique IP address is almost always assigned only to the active network components directly connected to the WAN. All the other devices connected behind the respective WAN interface receive addresses from the Private Name Space which, however, are then no longer unique worldwide. To prevent any misrouting, the IP addresses are identified as belonging to the connected network at the interface to the WAN via Network Address Translation (NAT).

To counter the problem of insufficient addresses according to IPv4, a new address structure has been developed.
In the Internet Protocol version 6 (**IPv6**), compatible with IPv4, the addressing is based on the hexadecimal system. An address according to IPv6 is 128 bits long and consists of a series of eight 16-bit long parts, providing four times more bits than in IPv4.
Apart from the increased address space and the attendant possibility of assigning an address that is unique worldwide to every host, a decisive advantage of IPv6 addressing is that so-called site multihoming is possible. Thanks to site multihoming, multiple IPv6 addresses can be assigned to hosts, allowing a host to use the services of multiple providers. In addition, IPv6 addressing offers a defined header format, making the processing of the addressed packets far more efficient. What is more, IPv6 also offers the option of using security headers.

17.3 Protocols in a WAN

WAN protocols are protocol families at Layer 3 of the ISO/OSI reference model. The Layer 3 protocols most widespread today will be introduced below, including the transport and routing protocols.

17.3.1 The Internet Protocol (IP)

The task of the platform-independent **I**nternet **P**rotocol (**IP Protocol**) is the transfer of data across multiple networks from a sender to the receiver. IP is at Layer 3 and is responsible for regulating the transfer of data fragments forwarded from Layer 4 to Layer 3. An IP packet is defined as having a minimum size of 576 bytes and a maximum length of 65,535 bytes. Due to this variability and the variously defined transmission structure and packet length per

subnetwork to be traversed, an IP packet may be split in the course of the transmission path. Such splitting is called fragmentation. Reassembling refers to putting the fragments back to restore the original state.
The structure of an IP data packet (also referred to as datagram) always remains basically the same regardless of the fragmentation.
The main constitutents of an IP packet are in consecutive order: The header with the information on the packet proper, the sender and receiver address and all the information serving for transfer control and troubleshooting. The header is followed by optional information such as security rules for the data transfer. The data information to be transferred is associated to the last bits.
Transfer of IP packets is packet-oriented, connectionless and not guaranteed. In the event of a buffer overrun along the transmission path, an IP packet is rejected when arriving at the overloaded node. However, thanks to the incorporated frequency check sum and the directly built up communication between sender and receiver, retransmission takes place in the event of an error. Since IP packets are not regarded as an aggregate of information units, but independently of one another, their transfer is arbitrary, i.e. connectionless. Each data packet takes its optimal path from the sender to the receiver. The TCP protocol based on the IP protocol is responsible for sorting the IP packets into the correct order at the receiver.

17.3.2 The Transmission Control Protocol (TCP)

The **T**ransmission **C**ontrol **P**rotocol (**TCP**) is at Layer 4 of the OSI reference model and is responsible for processing the data stream from Layers 5-7 in fragments capable of being processed for the IP protocol at Layer 3. TCP divides the flood of data, for instance of a file to be transferred, into packets by virtue of the prescribed rules and transfers them to the transmission protocol.
The **TCP/IP** (**T**ransmission **C**ontrol **P**rotocol/**I**nternet **P**rotocol) protocol stack defined by the U.S. Departament of Defense is available on every major computer platform. It was not designed for a special information transport system, such as a LAN protocol, but rather for use across different media and computers. Hence, TCP/IP is the most suitable protocol for networking heterogeneous systems. It is capable of connecting computers running under Unix (or any of its variants such as SunOS, Digital Unix, HP-UX, AIX), OpenVMS, DOS, or Windows. A range of applications have been based upon TCP/IP, which are usually included under the title of the protocol: E.g. **ftp** for file transfer, **telnet** and **rlogin** for remote control or remote login, electronic mail, Web browsers, etc.

17.3.3 User Datagram Protocol (UDP)

The **U**ser **D**atagram **P**rotocol (**UDP**) is, like TCP, at Layer 4 of the OSI reference model and, like TCP, is a transport protocol based directly on the IP Protocol. UDP is responsible purely for the transport and was developed to give applications the possibility of sending datagrams direct. UDP operation is connectionless. UDP sends data to the receiver, but without receiving acknowledgement that the transfer has been successful. UDP hence does not guarantee reliability of the transfer, lost packets are not re-sent.

17.3.4 Routing Protocols

The routing protocols are responsible for the sequence and routing of packets between a sender and a destination network. The forwarding decision of a router in a WAN is based on the routing table set up through the routing protocol and the metric derived from it.
The metric of a routing protocol is defined as the value based upon which a router takes its forwarding decision. Metric calculation varies from protocol to protocol.
Distinction is made between static and dynamic routing protocols.
Static routing protocols are generally used only in relatively small networks or for dial-up connections, because the burden on resources for manually updating routing tables is very high. Every new route and every change in a route has to be entered manually in the routing table.
Dynamic routing protocols on the other hand are often to be found in larger, growing networks. They are predestined for this use by virtue of involving less input on updating routing tables for the administrator. Dynamic routing protocols recognise changes in the network and update their routing tables in reaction to such changes. These protocols hear of changes from their neighbours in the network. In a WAN set up on the basis of a dynamic routing protocol, each router advertises its route information, i.e. the IP addresses it knows, to its neighbours. The best route through the network is selected on the basis of the information advertised specific to protocol.
A second general distinguishing feature between routing protocols is whether it is a distance vector or link state protocol.
Distance vector procotols are characterised by their comparatively slow convergence and the limited scalability, accompanied however by simple installation and convenient operation in service. They operate according to the "routing by rumour" principle and send their full routing tables to their direct neighbours at periodic intervals, who process this information by replacing their own routing tables in full with the new information. In large networks having highly complex routing tables this procedure leads to excessive traffic and excessive load on the routers. RIP, RIPv2 and IGRP are examples of distance vector protocols.
Link state protocols on the other hand support triggered updates, i.e. new routing information

is advertised as a single change, not as an entire new routing table, signifying very little comparison of routing information for links and routers. By advertising a topology change in the network directly after its discovery, networks based on link state protocols are significantly more convergent than those based on distance vector protocols. OSPF and IS-IS are examples of link state protocols.
EIGRP combines the features of both protocol types.

17.3.4.1 RIP

The **R**outing **I**nformation **P**rotocol (**RIP**) is a distance vector routing protocol, the metric of which is based purely on so-called **hop counts**. A hop is the transfer from one routing component to the next. RIP does not heed further metric criteria, such as bandwidth or delay in traversing a route. RIP is the oldest routing protocol and its modus operandi is very simple in structure. It is a flat, classful routine protocol, which advertises all the routing information throughout the network. RIP does not support VLSM (variable-length subnet masking) or manual route summarisation. The maximum distance for a network operating with RIP is 15 hops.

17.3.4.2 RIPv2

RIPv2 is an extended version of RIP and supports VLSM. In addition to the periodical updates typical of distance vector protocols, it can also advertise triggered updates, i.e. unscheduled changes in the topology of a network. RIPv2 is mainly implemented in small networks operating with point-to-point connections. It is also used for dial-in connections, because in RIPv2 the routing table may be frozen until a connection is requested and only then is the routing information exchanged.
Two weaknesses of RIP still persist in RIPv2 though: The hop count is limited to max. 15 hops and the metric criterion is still based only on the hop count.

17.3.4.3 IGRP/EIGRP

IGRP, the **I**nterior **G**ateway **R**outing **P**rotocol, was developed by Cisco and hence is a proprietary distance vector routing protocol. Compared to RIP, its metric calculation is more complex in that minimum bandwidth and the delay time to a link are added. Due to the pure distance vector functionality, it has pretty slow convergence time and is used only in small networks.
The weaknesses of IGRP were eliminated in the extended version, **EIGRP**, **E**nhanced **I**nterior **G**ateway **R**outing **P**rotocol. EIGRP combines the characteristics of distance vector and link state protocols. Thus EIGRP supports VLSM and triggered updates, i.e. the immediate advertising of

unscheduled changes in the topology of a network. What is more, routers with EIGRP support have routing tables not only identifying the best route, but recording the overall topology of the network. Manual IP address summarisation is also supported. EIGRP provides the following parameters for metric calculation: bandwidth, delay, reliability, throughput and MTU. However, only bandwidth and delay are used as standard for determining the best route.
Due to the proprietary, both IGRP and EIGRP can be employed only in networks based on Cisco hardware. Only Cisco hardware supports these protocols.

17.3.4.4 OSPF

Open **S**hortest **P**ath **F**irst (**OSPF**) is a link state protocol and was developed for routing IPv4. OSPF is optimised for use in large, scalable networks in which the limitations to which RIP is subject are not acceptable. OSPF offers enhanced convergence functions, supports VLSM and manual IP address summarisation. In OSPF the metric calculation for determining the best route is based on the path costs and the sum of bandwidths the connections provide between the first and last node in the transfer.
In-depth knowledge of the network is a basic requirement when using OSPF. In OSPF environments, a backbone has to be defined as area 0, to which all the other areas have to be directly connected as non-backbone. All routing traffic between non-backbones has to be conducted via the backbone. The routers between the various areas are called area border routers, that permit IP summaries allowing the routing tables to be kept relatively small and thus minimising the load on the routers.

17.4 Data Switching in a WAN

WANs transfer data either on the basis of circuit switching or data packet switching.

17.4.1 Line Switching

In circuit switching, the switching points are arranged according to a hierarchical approach, those lower in the hierarchy are linearly associated to those higher in the hierarchy. In such networks data can be transferred with the entire bandwidth of a circuit, since a separate circuit is set up across the hierarchies for every communication between senders and receivers. If, however, the bandwidth is not fully utilized by the transfer, the bandwidth is lost for the duration of the communication since this one set up circuit can be used only for this single communication. The connection is reserved only for the parties to the communication (pre-allocation).
Advantages of circuit switching are that the transfer is almost entirely delay-free, there is full protocol transparency and uncomplicated routing with little burden on computing resources.

The very long time required for making and dropping the connection has negative repercussions. Further, no redundant paths are provided and the poor utilisation of the bandwidth is to be regarded as a drawback of circuit-switching.

17.4.2 Data Packet Switching

Data packet switching splits the information (data) to be transferred into chinks, so-called packets. In **single packet switching** the sender allocates clear addressing, send sequence, flow control and error correction characteristics to the packets to enable the receiver to evaluate the entire information transfer and to establish whether and when transfer was fully concluded and how the overall information is to be composed from the individual packets. The individual packets are forwarded from the sender via various nodes to the receiver in a technology based on buffers. Each node stores the packets, checks them and then forwards the checked packets to the next node. This method of receiving, buffering, checking and forwarding is also called **Store and Forward**.

The transfer time of data-switching technologies is significantly shorter than that of circuit-switching technologies, particularly due to the shorter time for making and dropping connections. What is more, the bandwidth utilisation is optimised and redundant paths can be used. A drawback of the technology is that it is more error-prone.

17.5 Transmission Methods in a WAN

17.5.1 Synchronous Transmission Methods

Synchronous transmission methods transfer data within prescribed time intervals synchronised using a clock pulse. Each connected end station reads out the clock pulse in the header of the data packet and subordinates itself to it. Synchronous transmission methods are faster and more efficient than the less ordered, asynchronous transmission methods.

17.5.2 Asynchronous Transmission Methods

Compared with synchronous transmission methods, asynchronous transmission methods do not have a prescribed clock pulse. The data packets of asynchronous transmission methods have start and stop bits for error recognition. A transmission error occurs if the difference between a start and a stop bit upon arrival at the receiver is longer than it was at the sender.

17.6 WAN Technologies

In general WAN technologies are based on the lower three layers of the OSI reference model: the Physical Layer, the Data Link Layer and the Network Layer. As a rule the services are provided via one of the following technologies:
- By leased lines (dedicated lines)
- By circuit switching
- By packet switching
- By cell switching

All the switching techniques mentioned are based on virtual channels. The channel to which a data packet belongs has to be capable of being read out by the service provider on the basis of the information in the packet.

17.6.1 Leased Lines

Dedicated lines are point-to-point connections permanently reserved for data transmissions. They are not just set up when a transmission has to be made, but are established permanently regardless of whether or not data is flowing. For leased lines it is therefore advisable to optimise the bandwidth utilisation as far as possible to guarantee the best return on investment. The carrier of the leased line establishes the connection either via a physical connection or via an assigned channel with the use of either frequency modulation or time multiplexing. Leased lines generally enable synchronous data transmission.

17.6.2 Circuit-Switched Networks

Circuit-switched networks provide a dedicated connection between sender and receiver for the duration of the transmission. Thus, for example, analog connections across the telephone network (PSTN) run on circuit-switched networks. For the duration of the connection the line is not available for any other connection, even if the available bandwidth is not fully utilised. Further examples of circuit-switched connections are ISDN and asynchronous, serial connections.

17.6.3 Packet and Cell-Switched Networks

In packet- and cell-switched networks, the carrier creates permanent virtual circuits (PVCs) or switched virtual circuits (SVCs). In these networks the users share the resources available and use various paths through the network for their data transmission. This enables the carrier to improve the rate of capacity utilisation of its infrastructure compared to point-to-point connections and thereby to optimise the return on investment for all involved. It is employed

in networks that are intended to use X.25, Frame Relay or SMDS (Switched Multimegabit Data Service).
ATM is a technology used only in pure cell-switched networks. ATM divides data into cells of defined size and transfers them via a physical medium with the use of digital signal technology. ATM cell transfer can be either asynchronous, queued or with the use of a multiplexing technique.

17.6.4 Digital Subscribe Line (DSL)

DSL is a transmission technology enabling high transmission rates through the traditional telephone cabling. The four versions of DSL are Asymmetric Digital Subscriber Line (ADSL), High Data Rate Digital Subscriber Line (HDSL), Single-Line Digital Subscriber Line (SDSL) and Very high Data Rate Digital Subscriber Line (VDSL). Due to the fact that DSL and telephone transmission can operate on the basis of different frequencies, both data streams can be transferred in parallel. The most common DSL version is ADSL.
ATM and DSL will be described in further detail in a section below.

17.7 Data Transmission in WANs

17.7.1 Analogue Data Transmission in WANs

Probably, the most extensive and familiar WAN is the public telephone network, in which dial-up connections are established to provide mainly voice communication links. The use of this network as a platform for data transmissions via modem or facsimile will continue to be significant.

17.7.2 Digital Data Transmission in WANs

17.7.2.1 Integrated Services Digital Network - ISDN

ISDN combines the use of voice, text, images and data into one single network. This means that only a single ISDN connection is needed for access to a variety of telecommunications services such as telephone, fax, file transfer utilities, etc. The high-speed, error-free data transmission, quick connection and growing popularity make ISDN not only attractive for newcomers to data communications, but also for users who up to now have been dependent on analogue modems. ISDN operates through the already installed telephone lines for the end device. The main difference with respect to the current telephone network is that signal transmission is entirely in digital form, right up to the end device. Data communications between central relay stations and local relay stations is already completely in digital form.

17.7.2.1.1 Compatibility

If a connection between two ISDN devices is made, first of all control information is transmitted over a separate channel (D channel). This information includes the caller's number, which then appears on the display of an ISDN telephone, an indication of the type of the service desired (e.g. voice or fax) as well as other connection-related technical data. Following the introduction of ISDN, several standards were developed for the protocols used for sending control information over the D channel. This is the reason why, in Germany for example, two D channel protocols for ISDN exist: One is the national standard 1TR6 and the other is the European standardised protocol DSS1 or E-DSS1. The E-DSSI standard is known as the Euro-ISDN protocol, even though nowadays it is also used in many non-European countries. The number of network providers offering Euro-ISDN has grown to 30 worldwide, 20 of which being in Europe. When using only the telephone, no problems should be encountered if there is a difference in D channel protocols used at the respective locations. However, problems may arise when data is to be transmitted or if particular service features or line types are needed in ISDN, which are only available on a national basis. Meanwhile, Euro-ISDN has become the standard in many countries, and consequently, wherever possible, this type of connection should be used for new installations.

17.7.2.1.2 Connection Types

The ISDN system includes the following distinct connection types: For the **ISDN dial-up connection**, a B channel is provided for an ISDN service by the dial-up procedure to any ISDN subscriber. Dial-up connections to other nets are also possible (analogue telephone system, mobile phones, X.25). Digital **fixed lines** provide permanently switched digital 64-kbit/s channels. A maximum of 30 B channels (PRI) are available per connection. These dedicated permanent connections can be used both with and without a control channel.

17.7.2.1.3 CAPI

CAPI stands for Common ISDN Application Interface and was developed with the aim of creating a uniform interface between application software and the ISDN end devices (e.g. ISDN card). In principal, CAPI enables any ISDN card supporting CAPI to be used together with any ISDN application software which also supports CAPI. German companies from the ISDN sector and the Deutsche Telekom were at the forefront of the development of CAPI. There are at present two versions of CAPI in use. CAPI versions 1.0 and 1.1 have been designed for PC-based ISDN solutions. Practically all ISDN cards for PC installation support this standard. The more recent CAPI version 2.0 provides cross-platform support, which is a considerable improvement. Cross-platform support is also the reason why Windows NT, for instance, also

uses CAPI 2.0 in its Remote Access Services (RAS).

17.7.2.1.4 Communication via ISDN with analogue distant ends

If the distant end is being accessed via the analogue telephone network, for example via an analogue modem or analogue access to any other network, then data communication is considerably more complicated. Connections with analogue distant ends only work if the ISDN end device supports them. This is why several ISDN cards and ISDN terminal adapters feature an integrated analogue modem chip or a digital signal processor, allowing the device to be used like an analogue modem. It is also possible to communicate with analogue fax devices using ISDN cards, which do not have a modem chip, if the appropriate fax software is available for this card. This software converts digital data for the analogue fax device. However this method is complicated and under certain circumstances can result in errors in the fax transmission.

17.7.2.1.5 Communication via ISDN with digital distant ends

If a digital end device, such as a PC with an ISDN card or a digital fax (fax group 4), is located at the distant end, or you are dealing with ISDN access to a network or service provider (Datex-J-Service, CompuServe, Internet Provider), then the data exchange in the ISDN B channel usually takes place according to either the X.75 or V.110 standard. The X.75 protocol completely uses the ISDN bandwidth, and the data transmission rate amounts to 64,000 bit/s per ISDN B channel. If the distant end is not capable of supporting communications at 64,000 bit/s, then the data rate has to be adapted. This is called bit rate adaptation. For that purpose, fill bits are inserted to make up the difference between 64,000 bit/s and the required bit rate. This procedure is defined in the V.110 norm for 2400 to 38400 bit/s. Absolute flexibility in ISDN communication to the digital distant ends can only be attained by using hardware that supports both standards.

17.7.2.2 Digital Multiplex Hierarchies

17.7.2.2.1 Plesiochronous Digital Hierarchy (PDH)

The plesiochronous digital hierarchy (PDH) defines a digital technology for transmission between nodes that do not have identical clock speeds, and are hence asynchronous. PDH was standardised in 1972 by the CCITT at that time and has various bit rates all based on the bit rate of 64 kbit/s. This basic bit rate is defined in G.702 of ITU-T and is also called the E0 interface. PDH has meanwhile been superseded in new installations by the synchronous digital hierarchy (SDH). The European digital multiplex hierarchy PDH works on the basis of a

bit rate of 64 kbit/s per interface. This basic bit rate is also called an E0 interface. Additional E-X interfaces are defined thanks to a multiplex hierarchy, where the higher-level interface forms the quadruple multiplex of the lower-level interface.

17.7.2.2.2 PDH Standards

G.703

The G.703 ITU recommendation describes an interface used from transmission rates of 64 kbit/s upwards. It is a bit-oriented digital interface for transmitting unstructured data streams. The unstructured G.703 is as a rule used for transmissions with the maximum transmission rate of 2 Mbit/s possible with G.703, but this transmission rate can also be utilized only in part. Unstructured means in this context that all 32 available 64 kbit/s data channels are combined to form a data channel with 2 Mbit/s. G.703 is mostly used for connecting communication equipment such as routers or multiplexers. Transmission is performed either through balanced (120 ohm twisted pair) or unbalanced (dual 75 ohm coax) cable. Balanced service is the most common, with the exception of the Netherlands and the U.K.

G.704

G.704 is the structured version of G.703. G.704 operates exactly like G.703, only here all 32 available 64 kbit/s data channels (also time slots) are individually used for transmissions. In G.704, however, time slot 0 is reserved for the synchronisation.

17.7.2.2.3 PDH Multiplexing

To increase the bandwidth 32 E0 lines can be combined to form an E1 interface. One then speaks of multiplexing. As a result of multiplexing, the data of multiple lines can be sent on via only one line without any reduction in throughput. The E-1 transmission interface (also CEPT1) is the Primary Rate Interface (PRI) of the PDH multiplex hierarchy in Europe and thereby the first stage of multiplexing. It transmits up to 2,048 Kbit/s or 2 Mbit/s. These 2,048 kbit/s are made up of 32 E0 service channels for voice and data transmission each with 64 kbit/s. Thanks to multiplexing, the following increases are possible as variations of E interfaces:
E-2: A route over which four E1 lines are combined to form a line with a data throughput of 8,448 MbpS.
E-3: A route over which four E2 lines are combined to form a line with a data throughput of 34,368 MbpS.
E-4: A route over which four E3 lines are combined to form a line with a data throughput of 139,264 MbpS.
E-5: A route over which four E4 lines are combined to form a line with a data throughput of 565,148 MbpS.

"Plesiochronous" is equivalent to "almost synchronised". This is indicative of the fact that in PDH certain unavoidable differences in clock speed can occur in the various multiplex stages, which are made up for by filler bits for compensation in the upper hierarchy stages.

17.7.2.2.4 Multiplex Techniques

There are five different types of multiplexing: Space division, frequency division, time division, code division, and wavelength division.

17.7.2.2.4.1 Space Division

With space division multiplexing (SDM) the physical transmission media are combined into one frequency band or cable. In relation to digital transmission, space division multiplexing is used to combine several ISDN connections at one local exchange in a primary rate interface.

17.7.2.2.4.2 Frequency Division

With frequency division multiplexing (FDM), a broad frequency band is divided into several smaller frequency bands that can all transfer data parallel, but independent from each other. FDM is used, for example, for telegraphic connections. FDM is not commonly used today, since it has been superseded by time division multiplexing

17.7.2.2.4.3 Time Division

With time division multiplexing (TDM), the individual transmission requirements are processed one after the other. Every channel is assigned a time slot to transfer its data. If the time slot is exceeded, transmission is disrupted and the next channel is given the right to transfer for the length of its time slot. However, in TDM you must differentiate between synchronous time division multiplexing (STD) and asynchronous time division multiplexing (ATD).
The synchronous time division multiplexing (STD) technique defines transmission frames, of which each are made up of a specific number of time slots with fixed sizes. Every user is assigned a specific time slot within the transmission frame, and during this period it can send or receive data. The asynchronous time division multiplexing (ATD) technique transmits the data asynchronously in information units that have a fixed or variable length. Information unit assignment takes place with channel identification numbers, so-called channel identifiers, that are included with every packet. ATD is thus also called address multiplexing. This procedure is also called cell switching if data packets with fixed lengths are used during transmission, and data packet switching if the packet lengths are variable.

17.7.2.2.4.4 Code Division

Code division multiplexing (CDM) allows you to simultaneously transmit data from several senders over one medium (frequency or channel). This is made possible by individual coding of the data using a bit pattern per sender, which enables the receiver to reproduce the information using the code. CDM is used in UMTS.

17.7.2.2.4.5 Wavelength Division

Wavelength division multiplexing (WDM) is used for transmission over fibre-optic cables. In WDM, various fibre-optic wave lengths are transmitted paralled as a light current. Simultaneous transmission can even be done in full-duplex mode.

17.7.2.2.5 Synchronous Digital Hierarchy (SDH)

SDH, or the American version SONET, is a flexible transparent multiplex structure and is characterised by its high availability and reliability. Transmission rates between 155 Mbit/s up to 40 Gbit/s can be realised with SDH. In comparison with PDH, SDH provides the advantages of bundling lower transmission rates (e.g. 2 Mbit/s) into a higher transmission frame. For the information transmission SDH uses STM signals (Synchronous Transport Module) transmitted in the STM basic data frame STM-1 at 155 Mbit/s. The higher transmission rates can be realised by virtue of multiplexing. The clock pulse in SDH, compared to the clock pulse in PDH, is controlled by a uniformly, centrally generated network clock pulse. Accordingly all the hierarchy stages have the same clock rate and lower hierarchy stages can be directly accessed without having to pass through all the multiplex stages. This direct access is realised through STM transport units.

17.7.2.2.5.1 SDH-Multiplexing

The STM basic data frame STM-1 allows for a data transmission rate of 155 Mbit/s. An increase of the data rate is always possible by quadrupling the lower multiplex hierarchy. This results in the following data transmission rates:
STM-4 with 622 Mbit/s transmission rate as the second hierarchy step
STM-16 with 2.5 Gbit/s transmission rate as the third hierarchy step
STM-64 with 10 Gbit/s transmission rate as the fourth hierarchy step
STM-256 with 40 Gbit/s transmission rate as the fifth hierarchy step

17.7.3 Serial Interfaces in WANs

Serial interfaces are often used for dedicated line connections in WANs. The most common ones are described below.

17.7.3.1 V.24

V.24, developed by ITU, is an interface transmitting 20 kbit/s over 15 metres. It is probably the most widespread interface technology for Data Terminal Equipment (DTE), generally a telecommunications connection (e.g. a modem). V.24 is independent from the type of transmission and supports both synchronous and asynchronous transmissions. The physical interface is the 25-pin miniature D-sub connector.

17.7.3.2 RS-232

RS-232 is the serial, bi-directional interface, which is based on the functional characteristics of V.24 and transmits data asynchronously. RS-232 defines serial transmission in both the half-duplex and in the full-duplex mode. RS-232 is specified for two serial channels which can both be active independently of each other and can send and receive at one and the same time (full duplex). In practice RS-232, with the full name RS-232-C for the current version, is used for the communication of a computer with another computer or another serial device. Since the computer internally uses a parallel data stream, RS-232 specifies how the Universal Asynchronous Receiver/Transmitter (UART) chip at the Data Terminal Equipment (DTE) interface of the computer converts these parallel data streams into a serial data stream, which is then sent serially, one bit after the other, to another computer. As the DTE of the computer, the UART also communicates with other serial devices that then have the counterpart to the DTE as the communication interface: The Data Communication Equipment (DCE) which, for example, modems and routers use.

17.7.3.3 X.21

X.21 is the ITU-T standard for access to digital networks. It applies to all packet-switching and circuit-switching networks, but also for dedicated lines, with digital access. X.21 applies to both symmetric and asymmetric signal lines. Throughput rates of up to 10 Mbit/s can be achieved with X.21, as a rule however 2 Mbit/s dedicated lines are set up with X.21. A 15-pin connector is the standard interface for X.21.

17.7.3.4 V.35

V.35 operates with bandwidths greater than 19.2 kbit/s and is likewise used as an interface between packet-oriented data networks and access equipment to these networks. Several telephone lines can be combined with V.35.

17.7.4 X.25 and Frame Relay

17.7.4.1 X.25

The probably best-known standard for packet-switching transmission methods is **X.25**. X.25 was developed by ITU-T. In its recommendation for X.25, the ITU-T describes the synchronous operation of an end device operating in the packet-switching mode and connected to a public data network. It also takes the timing and transmission format for X.25 transmissions into consideration. The ITU-T recommendation is oriented to the three lowest layers of the OSI reference model. In X.25 layer 1 the physical connection and transmission proper take place, in layer 2 the security and control schemes for checking and error detection of the transmission. In layer 3, establishing and closing the connection and the terminal system connections for the higher application layers are defined for X.25 transmissions. This interface between the third layer and the transport layer above it in the OSI model is called Network Service Access Point (NSAP).In X.25 up to 4,096 virtual connections are set up between the linked end stations and are divided between 16 logical channels max. These connections are either permanent virtual circuits or switched virtual circuits. X.25 packet services are particularly suitable for linking networks with X.25-compatible remote routers. The main advantages are the short call set-up time of one second and the good transmission quality. Transmission speeds of between 300 to 64,000 bit/s are possible. There are separate data networks for data transmission with X.25, e.g. Datex-P in Germany, Telepac in Switzerland, Datanet 1 in the Netherlands or ISTEL, British Telecom.

17.7.4.2 Frame Relay

Frame Relay is a transmission system similar to X.25, which is characterised by its comparatively low protocol overhead. In contrast to X.25, Frame Relay only operates on levels 1 and 2 of the ISO/OSI model. Thus, the protocol does not guarantee the correct delivery of data packets, but leaves this task to the terminal devices.

In this way, the efficiency problems associated with X.25 are bypassed by moving the error correction function to the upper protocol layers. Considerably higher data rates are achieved compared with X.25. Therefore, Frame Relay is a very good choice for improving the data

throughput in a WAN. In comparison with fixed lines, Frame Relay offers economic advantages wherever the capacity of a point-to-point connection is not being used to its full potential. Depending on the country, Frame Relay technology can be employed in various public or private networks, e.g., in the Datex-M Service of Deutsche Telekom. The **s**ubscriber **n**etwork **i**nterface (**SNI**) is used for network acess. The small entry interface operates at a minimum transmission speed of 64 kbit/s with the next category at 2 Mbit/s (E1). In terms of the standard interface to the router or bridge is X.21. The highest speed class at present achieves 34 Mbit/s, which conforms to the E3 standard.

Although Frame Relay offers clear advantages over X.25, this technology is less suitable for time-critical applications such as voice and video transmission. A few years ago, this problem would not have arisen since voice, video and data networks were set up separately. Since multimedia applications such as video conferencing and video distribution have been increasingly incorporated into LAN and WAN networks, the integration of audio and video information is now essential in networks which have until now only been used for data transmission. The packet-switching transmission method - the Asynchronous Transfer Mode, ATM, plays a role here. Various providers are offering Frame Relay connections.

The prices quoted by these providers differ significantly, but the performances vary even more so. It is not usual in this market for CIRs (Committed Information Rates) to be actually provided or for Round Trip Delays to be within a reasonable range. Before deciding upon a Frame Relay provider, the quality of the offered network should be carefully checked and contractually binding penalty clauses should also be agreed on in case guaranteed characteristics are not provided.

17.7.5 Asynchronous Transfer Mode (ATM)

Asynchronous **T**ransfer **M**ode (**ATM**) is a fast, service-oriented, asynchronous, packet-switching transmission method developed mainly for use in wide area networks. ATM has however also found its way into some local networks.

17.7.5.1 Data Transmission in ATM Networks

As a rule, it does not make much difference whether a file transfer takes a second more or less in the case of time-critical data, such as audio and moving pictures. Even a slight delay in the audio data stream will cause an audible discontinuity or click, and in video information, jerkiness of the image sequence is the unavoidable result. Therefore, if multimedia data are to be integrated into the network, all the hardware and software components involved must possess a predictable real-time behaviour. ATM offers this feature, and is therefore a

technology intended for multimedia applications. The key to ATM is the transmission rate, however, which reaches the gigabit range. Network managers are pressing their suppliers and system integrators to provide solutions to their bandwidth problems in the backbone. This is where ATM represents a promising investment, and will shift from being a high-end technology to an everyday solution. Companies should study the market and look for migration paths that lead the user to ATM. In addition to its speed, there is the scalability of ATM, meaning the flexible provision of the bandwidth that happens to be needed. The present-day generation of ATM switches typically supports 155 Mbit/s or 622 Mbit/s per port. But this is nowhere near the upper limit: ATM permits speeds in the multi-gigabit range. ATM is the first network technology that can integrate LANs and WANs seamlessly. This makes additional gateways transposing from LAN to WAN protocols unnecessary.

In ATM, all types of information (audio, video and data) are conveyed in packets of fixed length (53 bytes), which are called cells (cell relay). 48 bytes are available for user data, 5 bytes are reserved for check information. Some essential characteristics of ATM result from this data structure. Thanks to the uniform length of all cells, there is a calculable delay in the transmission of any information, which means that guaranteed bandwidths can be assigned if there are several competing data flows for individual applications. With other topologies, long packets can cause problems, e.g. blocking other applications during file transfers. Due to the cell format used by ATM, such problems are avoided. Very short blocks are more suitable for voice transmissions, long blocks for data and video transmission. The ATM forum has found a salomonic compromise with the cell size for ATM, which in theory is suitable for all services. As most information cannot be accommodated in a single ATM cell, the different length packets of higher network layers are partitioned into several ATM cells and reconstituted at their destination using the adaptation mechanism **S**egmentation **a**nd **R**eassembly (**SAR**) of the sending ATM adapter.

The transmission itself is based on the principle of virtual links, the same principle on which packet switching networks such as X.25 are based. Using this principle, no permanent channels are established between the end devices involved. Instead, the cells of a particular connection are transported along a previously determined path through the network when they are generated. In this way, the resources of the network elements can be used efficiently for several connections. ATM recognises permanent connections (**PVC**s, **P**ermanent **V**irtual **C**ircuits) as well as switched connections (**SVC**s, **S**witched **V**irtual **C**ircuits). These in turn can consist either of virtual point-to-point or point-to-multipoint connections.
Connections can be established with a specific grade of service, **QoS** (**Q**uality **o**f **S**ervice). Each device connected to the ATM can be allocated the required bandwidth statically or dynamically, by means of the PVCs or SVCs mentioned, this being limited only by the capacity of the ATM hardware. After the link has been established, one can be sure of being

able to use the bandwidth of the transmission channel requested alone, without being disturbed or even interrupted by other stations waiting to transmit.

The system attempts to meet the respective demands in the best way possible. In principle, ATM offers three different categories of service: Guaranteed services, predictable services and services that utilise an available bit rate to the best possible effect. A guaranteed connection is needed for supporting constant bit rates (**CBR**, **C**onstant **B**it **R**ate), for which all cells must reliably transfer data to their destination with only a tiny time deviation, as is the case for speech information. But there are also applications with variable bit rates, for example file transfers and e-mail, for which a service that utilises the available bit rates as best possible, suffices. Here, there are no hard-and-fast guarantees. Rather, the cells are simply transmitted as best as possible on the bandwidth that is left available by connections with a higher priority. In the ATM Forum's specification, these services are listed under the term Traffic Management.

17.7.5.2 Interfaces

A large number of ATM transmission interfaces are defined both for wide area networks (WAN) and local area networks (LAN). The transmission speeds for the physical media range from 1.5 Mbit/s to 622 Mbit/s. The differences between LAN and WAN specifications are primarily in the transmission medium. While single mode fibres and coaxial cables are primarily provided for the WAN sector, multi-mode fibres and Twisted-Pair cables are in more prevalent use in the LAN sector.

17.7.5.3 Availibility and Standardisation

The ATM Forum, an interest group of ATM manufacturers, is the driving force behind the standardisation of ATM. It provides the rather hesitant ITU (successor to CCITT) with proposals. The definition of cells was completed some time ago as well as the definition of the functionality of the hardware components of ATM systems. Also defined was the way in which the hardware of ATM systems has to function, how connections are to be set up, how they should be monitored and which logical interfaces are needed. However, ATM covers other specifications. According to the manufacturers, it presents no problem to adapt the products to different standards by means of firmware upgrade. This ensures compatibility of products from the same manufacturer with specifications from the ATM Forum. If ATM products from more than one manufacturer are to be used together, the products used must be compatible.

17.7.5.4 Migration to Existing Networks

Due to the substantial differences between ATM and previous network technologies, flexible integration of ATM products in existing networks is necessary to protect hardware investments. This long-term process of adaptation is called migration. The migration of ATM will take place in several steps. ATM switches can be used in large networks as a collapsed backbone. Previously installed ring or bus-type backbone cables are brought together at the ATM LAN Access switch, which connects all segments and devices of the network in a star-type topology. Additional ATM switches or ATM routers can be connected to such an ATM switch, as well as end devices using ATM adapters.

Existing network topologies are connected to central ATM switches by means of routers with ATM functionality. Most of the manufacturers of large network concentrators offer modules for ATM connectivity. Setting up a continuous connection to terminal devices is the last step in this process. For this to happen, the added costs per port of ATM networks will have to drop dramatically. However, for applications such as high performance workstations, servers, image processing or virtual reality, ATM is already an interesting alternative to other topologies. But how can existing applications be integrated into an ATM network? The easiest way would be to exchange the existing computer's LAN interface for an ATM board and install the ATM driver. Unfortunately, things are not that easy. ATM lacks one property that conventional LANs offer: Support for broadcasts, multicasts, and the use of MAC addresses. These are basic prerequisites for the use of most network protocols, such as TCP/IP or IPX. The solution agreed on by the ATM Forum to this dilemma is called **LAN E**mulation **LANE 2.0**.
In a network combining ATM with standard LAN technologies, LANE consists of four software components. The **LAN E**mulation **S**ervices (**LES**s), consisting of three components, are made available to the computers connected to the ATM network as ATM applications or as an integrated part of an ATM switch. The **L**AN **E**mulation **C**lient (**LEC**) is installed on the end device as the fourth component. LAN Emulation Services:
The **LAN E**mulation **C**onfirmation **S**erver **LECS** coordinates all LESs and LECs.
The **L**AN **E**mulation **S**erver **LES** translates MAC addresses into ATM addresses and vice-versa.

The **B**roadcast and **U**nknown **S**erver **BUS** continues to forward broadcast and multicast pakkets using ATM point-to-multipoint connections and forwards packets to unknown computers, until the destination address is found. At the beginning of a transmission, the LEC contacts the LECS using a uniform network address to exchange general LANE information. The LECS provides the client with the address of the corresponding LES. With this, the LEC establishes a connection with the LES. Then, ATM address, MAC address, frame size and LAN type are agreed on. The LEC passes broadcasts and packets with unknown destination addresses to the BUS which forwards them. The missing MAC level features are then emulated for clients, for

whom ATM then appears either as an 802.3 (Ethernet) or 802.5 (Token Ring) network. This enables communication between conventional LANs and ATM.

17.7.6 Digital Subscribe Line (DSL)

DSL is a transmission technology enabling high transmission rates via traditional telephone cabling. The four standardised versions of DSL are Asymmetric Digital Subscriber Line (ADSL), High Data Rate Digital Subscriber Line (HDSL), Single-Line Digital Subscriber Line (SDSL) and Very High Data Rate Digital Subscriber Line (VDSL). The term **xDSL** is often used when referring generally to DSL in all its various forms. Due to the fact that DSL and telephone transmission can operate on the basis of different frequencies, both data streams can be transferred in parallel.

17.7.6.1 Asymetric Digital Subscriber Line (ADSL)

ADSL is the most widespread and most frequently used DSL standard. ASDL is an asymmetric wide-band data transmission technique using conventional twisted-pair copper lines. For ADSL communication an ADSL modem has to be installed at either end of the line, both at the local exchange and at the subscriber's end. In the ADSL method the copper line is divided into three channels: The downstream channel from the service provider to the end user, the upstream channel in the opposite direction and a channel through which PSTN and ISDN communication takes place simultaneously by adding a splitter. The signal to be transmitted is divided into diverse parts transmitted using different carrier frequencies. In addition to the discrete multitone (DMT) method specified in the ADSL standard, the CAP/QAM method is also used. With ADSL transmissions up to 6 Mbit/s can be realised using normal telephone cable, for bi-directional transmission a second frequency band with transmission rates up to 640 kbit/s is available. ADSL can be used for distribution services and interactive video services alike. Examples of this are pay-per-channel, pay-per-view, video-on-demand and information-on-demand applications. The distance to the next node can be 4 kilometres for transmission speeds above 1.5 Mbit/s and below that up to 6 kilometres.

17.7.6.2 High Data Rate Digital Subscriber Line (HDSL)

HDSL is ideal for private branch exchanges and is an alternative to local Frame Relay. In addition, it can replace amplifier-supported or E-1 lines. With HDSL, transmission rates of up to 2 Mbit/s can achieved over distances of up to 5 kilometres.

17.7.6.3 Symmetric Digital Subscriber Line (SDSL)

Like HSDL, SDSL is a method for transmissions in the full-duplex mode with symmetrical transmission speeds of 64 kbit/s to 2 Mbit/s (E-1 line). Unlike HDSL, only one pair is needed for transmission. The maximum distance for SDSL is 3.5 km.

17.7.6.4 Very High Data Rate Digital Subscriber Line (VDSL)

VDSL is used wherever symmetrical or asymmetrical data streams have to be transmitted at high speed over relatively short distances. This variant of DSL can be transmitted over both copper and fibre optic lines.
With VDSL, depending on class, symmetrical transmission bit rates of up to 72 Mbit/s are reached. In terms of classes, distinction is made between asymmetrical and symmetrical transmission, the class for the asymmetrical operation modes having bit rates of 8 Mbit/s, 16 Mbit/s and 28 Mbit/s. The symmetrical class has aggregate bit rates of 12, 24, 48 und 72 Mbit/s,can hence also be used for high-resolution television transmission.
With VDSL, downstream transmission speeds, hence from node to user, of 13 Mbit/s to 52 Mbit/s are reached with twisted-pair copper lines. The upstream transmission speeds are between 1.5 Mbit/s and 2.3 Mbit/s. VDSL is divided into three ranges: Long-range between 1 and 1.5 kilometres, mid-range between 300 and 500 metres and short-range less than 300 metres.

17.8 Security in WANs

A basic prerequisite for the success of a company is that network security be implemented. The unique solution underpinning a company's success must remain in sole possession of the company. Should the respective information and data come to the knowledge of third parties, the company runs the risk of having to share its turnover and profit with third parties or to lose them altogether. Willfull attacks levelled at disrupting smooth operations in a company are also to be regarded as posing threats to security. Therefore care has to be taken that WAN-based communication cannot be tapped, that no infiltrators from the WAN can penetrate the internal company network, thus gain access to the network resources and possibly alter them, and that public servers like Web servers or mail servers cannot be attacked.

17.8.1 Attacks and Targets of Attack

So-called integrity or confidentiality threats constitute the greatest danger to a company. Integrity threats occur when an attacker tries to alter corporate data without the respective authority for access or editing, e.g. when an attacker tries to gain access to development plans, to alter or delete them. The owner may not even notice this attack before it is too late and all the data are lost. Another example is attack on a Web server, altering the attacked party's presentation on the Internet. It is impossible to put a figure to the damage this may to do to the image of the victim of such an attack. In many companies integrity threats are regarded as those to be taken most seriously, because it is very difficult to detect and repair any changes. Deletion of data may even destroy the basis for the company's entire success. Confidentiality threats are levelled at gaining access to secret or highly confidential data in order to learn their contents. These attacks are even more difficult to identify because, once the attacker has left the network again, no noticeable changes to the data can be found. If the attacker has copied the secret data, this can be ascertained only with great difficulty, or not at all. A further attack of this kind is intercepting data transferred through the WAN, e.g. from a branch to the head office, altering and forwarding such.
Service availability or denial of service (DoS) attacks, targeted at making services unusable, constitutes a further threat. In a DoS attack, a vast amount of data or requests is sent to the destination to occupy the recipient to such an extent that the service can no longer be rendered or the service is even discontinued. Servers with public access, such as Web servers or WWW servers, are for example flooded with e-mails or inquiries and thus forced to give up. Internet connections can also be targets for DoS attacks. For example, the attacker sends so many packets to a router that inquiries from the actual users can no longer be accepted and processed. Active network components can also be the target of attack. If an attacker gains access to a switch or a router, for instance, there are a multitude of ways in which this illegal access can be used for operations inflicting damage:
- Intercepting data
- Attacking security services set up on the basis of confidence between network components
- Altering configurations for example.

The attacker could copy data flowing through a backbone switch and misappropriate it. If a remote access server becomes the victim of an attack, the attacker could, for example, read out passwords and use them for further attacks, gaining access to the network via the password as a regular user. In case of attack on a firewall, the attacker could alter the configuration in such a way that the firewall identifies him in future as having access authority. An access router could be reconfigured so that access to the Internet is no longer available. Entire networks become targets for attackers when the attacker thereby hopes to gain information for specific subsequent attacks. In this way a network might be combed

for potential, unprotected targets of attacks. The attacker tries to get a picture of the entire network, including the topology, the services used and applications employed. From this analysis he derives the weaknesses and points of attack, which are then exploited for further, more specific attacks. The latter may be attacks on hosts, applications or network availability. These attacks follow the same targets as those already described: Gaining unauthorised read or write access to secret data or rendering services unavailable.

17.8.2 Protective Measures and Strategies

Unless any appropriate security measures are taken, the potential dangers described represent an almost incalculable risk for every network provider. The basic requirement for effective protection against outside attack is to know the potential points of attack in one's own network, to classify them and take appropriate measures for their protection. An internal security policy defines the potential danger and how it is handled in the company. A security policy documents the risks to which a network is exposed and how these risks are dealt with. The development and implementation of such a policy is an ongoing process. Policies are as a rule broken down to individual areas with the objective of minimising the identified risks. Area-specific security measures are taken on the basis of the respective policy. The subject matter of such a policy might be, for example:

- Definition of network access control. What data is accessed when, by whom and via what route?
- Definition of a data hierarchy. What data is of what level of importance to the company and who gains what access to it?
- Definition of a user policy. What conduct is expected of a user who has access to the network?

A further example:
A company wants to develop rules for Internet access as part of a fully comprehensive security policy. The security policy says that Internet access is available to users only for certain, necessary services, that no user is allowed to install applications himself and then use them, and that all transactions contravening the defined security measures are to be fully documented via an audit. As a result of this definition of the rules for Internet access, the company implements the following precautionary measures:

- The firewalls allow the passage only of HTTP traffic and e-mails
- Internet access is allowed only for business purposes
- The protected network does not permit any download of executable programs from external sources
- The network components keep log files on all Internet traffic.

17.8.3 Access Control in Networks

Data integrity and data security are ensured by a well thought-out access control based on a guideline defined in a security policy, for example, specifying what user is granted access to what resources. Network access control is based on two measures: Authentication and authorisation.

17.8.3.1 Authentication

Authentication checks and determines the identity of a user. After successful authentication, administrators gain access to network components, hosts and network management facilities. Users gain access to the network and to the resources for which they were authorised. Authentication is especially important if public access facilities, such as access points, are implemented in the network. It is vital at such access points to distinguish users having access authority from those who do not. The identification of a user is based on a unique code:
- Which the user knows, for example a password or a PIN
- Which the user owns, for example a token or a smart card
- Which forms part of the user, for example a finger print.

An efficient access control combines at least two of the above-identified criteria: Two Factor Authentication. These methods which are commonly in use today in the EDP field, form the basis for a whole series of user identification and authorisation possibilities. The various procedures, together with their corresponding advantages and disadvantages, are described below.

17.8.3.1.1 Authentication Methods

The most common procedures for user identification today are:
- Stand-alone passwords
- Dial-back systems
- Authorisation by query/reply systems
- Software tokens
- Time-synchronised tokens.

17.8.3.1.1.1 Re-useable Passwords

Use of the protection method most widespread today, i.e., passwords, has the following disadvantages:
- Passwords can be manipulated
- In practise, they are not changed at sufficiently regular intervals

- They can be known to more than one user
- They depend on different social components
- They can be decoded by special programs.

Besides, their circulation and use are difficult to control and to track.

17.8.3.1.1.2 Dial-Back Systems

Dial-back systems used with passwords offer greater protection. They have the advantage that access can be traced via the listing of telephone calls. An additional advantage is that the caller (normally the company switchboard) pays the cost. This is useful but irrelevant from a security viewpoint. Nevertheless, dial-back systems still have disadvantages:

- The device, not the user is authorised
- Dial-back seldom works for mobile users
- Authorised users not using a pre-defined device can be locked out
- The system can be interfered with by call forwarding
- They are troublesome and time consuming for the user
- They are useful for switched line access only (e.g. they offer no protection for access via X.25 or the Internet)
- They offer no protection on the application or protocol level.

Dial-back systems in general are difficult to control and do not yet offer extensive security.

17.8.3.1.1.3 Query/Reply Systems

Query/reply systems consist of two partial systems, one installed on the user's machine the other installed on the host. When an access check is performed, the host system sends information to that of the user. This information is altered at the user side via an integrated algorithm and possibly completed with data. This message is returned to the host system as the result of the operation. If the result is satisfactory, access to the host is granted. This process can be performed several times in order to achieve even greater security. The hardware dongle systems integrated into some software products are also types of query/reply systems. These systems offer quite a high degree of access security. However, depending on the system type, there are certain disadvantages:

- The user may have to perform several operations
- Depending on the type of system, they can be problematical to use
- Apart from the PIN stored as an exchangeable token, a second PIN is needed for the host
- Authentication systems can be loaned or stolen - Export is problematical

17.8.3.1.1.4 Software Token

If access protection is integrated into the operating system or applications software, no external device is needed for authentication. This is the main advantage of software tokens. The disadvantages of this solution are:
- They are dependent on the data end device and/or operation system
- The software token can be copied or altered
- PIN, key and algorithm are all stored in the program
- Management is problematical.

Audit trails and controls are arduous to perform and can point to the wrong person. For this reason, there is no reliable way of controlling their circulation and use.

17.8.3.1.1.5 Time Synchronised Tokens

Time-synchronised characters or number strings which are constantly recalculated using sophisticated algorithms, offer a considerably higher degree of security. The hardware required can be accommodated on an area of the surface the size of a credit card.Advantages of this system:
- Independent of the data end device
- A card reader is not required
- Login can be carried out in one step
- No query/reply process
- It is relatively secure - based on two independent factors (user's own PIN and token)- At any given time, only one person can have the token card; in case of theft, the card can be deactivated immediately. Thus, access can still be controlled even when PIN and card fall into the wrong hands.

17.8.3.1.2 Authentication Protocols

17.8.3.1.2.1 Password Authentication Protocol (PAP)

The PAP is a simple authentication protocol, that is used to authenticate a user by means of an ID and password. It is part of the PPP tunnelling protocol. PAP is a two-step authentication process, in which the client sends an uncoded ID and password to the server after the connection has been made, which verifies it. If the ID and password are correct, the server confirms this and allows access to the network. The connection is closed if the entries are incorrect.

17.8.3.1.2.2 Challenge Handshake Protocol (CHAP)

The challenge handshake protocol (CHAP) is used to authenticate systems and is implemented by links that use the PPP tunnelling protocol. This deals with a three-step authentication process, where the participant logs on in the first step, the distant end, which also provides the key for the code, is asked for the password in the second, and the external participant enters the encrypted password and finally receives access authorisation in the third step.

17.8.3.1.2.3 Terminal Access Controller Access Control System (TACACS)

TACACS is an authentication protocol developed by Cisco, which allows a remote access server to forward authentication data to a special authentication server. The passwords for user authentication are then administered in central database.

17.8.3.1.2.4 TACACS+

Although the name might cause one to think so, TACACS+ is not comparable to TACACS. TACACS+ is a totally new authentication protocol that works with a message digest function for passwords. TACACS+ also supports PAP and CHAP authentication. TACACS+ uses the transmission control protocol (TCP).

17.8.3.1.2.5 Remote Authentication Dial-In User Service (RADIUS)

RADIUS is a client/server protocol which allows a remote access server to communicate with a central server to authenticate dial-in users. A database that contains all authorised user profiles is stored on the central server. Since the user database can be found centrally using RADIUS, this can also be administered centrally, which is much easier for the administrator. Additionally, a company-specific security policy can be easily implemented. Besides all that, centralisation also allows for simplified monitoring, statistical evaluations and, if necessary, calculation.

17.8.3.1.2.6 Secure Shell (SSH)

SSH is a security protocol that ensures the authentication and communication through encoding, if, for example, a user logs on to a UNIX computer. Here the authentication of the user can be done with encoding, in addition to verifying the password. The public key used for this is stored on the target system, and the private key is used by the user for authentication.

17.8.3.2 Authorisation

Authorisation defines who is allowed to do what in a network. Based on the rights assigned to every user, authorisation mechanisms limit the access to resources and define how every user is allowed to work with the resources. To achieve effective protection, only those resources absolutely necessary for a user to perform a task should be accessible to him. Redundancy is also to be considered in authorisation questions. One security mechanism should be confirmed by another one before final access to a resource is freed. Central control on authorisation is vital for it to be truly effective. Authorisation can be realised by implementing, for instance, the following security measures:
- IP addresses are assigned to clients requesting access from a creditable centre
- Address filters limit the access to resources
- Application filters allow only permitted applications.

17.8.3.3 Intrusion Detection

Intrusion detection systems operate with the use of so-called attack signatures, i.e. typical patterns of an attack or type of attack. These signatures define certain conditions that have to be met before data traffic is identified as an attack.

There are two complementary intrusion detection technologies. The network-based intrusion detection system (NIDS) monitors all the packets passing through a collision domain. As soon as the NDIS recognises a packet or series of packets that correspond to a predefined pattern or suggest an attack, the NDIS raises an alert and/or the session is ended.
In a host-based intrusion detection system (HIDS), agents are posted in the host to be protected. This technology comes into play when an attack is made on that specific host, an alert is raised and access to the host is ended or refused.

17.8.4 Protection of Data Transmission

Protection of data transmission is necessary when data is transferred over unsecured networks. The danger in such transmission is that an attacker can listen in, read out and misuse data. Protection can be realised only by data encryption. Encrypted data can be read out by third parties, but the contents cannot be reconstructed if the encryption has been done effectively. The data can be decrypted only by authorised parties, i.e. the users having the decryption code. An example of a widely used encryption algorithm is IPSec, which effectively encrypts the payload of IP packets and is sent through a VPN tunnel from the sending device, which encrypts the data, over an unsecured network, for example the Internet, to the receiving device, which decrypts the data.

17.8.4.1 Virtual Private Networks (VPN)

The introduction of **V**irtual **P**rivate **N**etworks (**VPNs**) enables mobile users and branch offices using unsecured public networks to be securely linked to the head office. From the viewpoint of security, a VPN can be compared with leased lines or lines owned by the user, but it provides more flexibility and - when properly used - also yields cost advantages, because VPNs are set up only when data is to be transferred. VPNs can be used either through dial-up access in the analogue telephone network or ISDN network, or over the GSM network and, for large data volumes, also through dedicated lines. Generally speaking, VPN is a closed, logical network, which is usually based on layers 2 or 3 of the OSI reference model and is established for a specific group of users. VPNs use tunnelling mechanisms for IP traffic.

17.8.4.1.1 Tunnelling

Tunnelling is the process in which two different protocols are encapsulated on the same layer. The data of the one protocol are packed in data packets of the second protocol. This method is employed at the transition from one protocol to another. It serves for transporting data over an unsecured, public network, such as the Internet, between a central VPN gateway and a remote VPN client. A virtual link is set up between the endpoints. A tunnel is set up by each data packet receiving an additional IP header, and also one or more special header fields. The start point of the tunnel is where the IP header is added, the endpoint is where it is removed again. Authentication and encryption take place within the tunnel. Tunnelling can take place on layer 2 and layer 3.

17.8.4.1.1.1 Tunnelling-Standards and Protocols

17.8.4.1.1.1.1 Generic Routing Encapsulation

Protocol-independent tunnels can be set up with GRE, a standard in which a tunnelling process is described in more detail. The GRE tunnel packet is comprised of the tunnel and GRE header and the payload. The tunnel header contains the address information; the information on the encapsulated tunnel protocol is in the GRE header and the encoding algorithms and payload cover the protocol header of the encapsulated tunnel protocol and the user data.

17.8.4.1.1.1.2 Point to Point Protocol (PPP)

The point to point protocol (PPP) was created to encapsulated datagrams over serial connections and supports transmission of all common LAN protocols. The PPP protocol overrides the limits of interoperability, which is caused by the encapsulation technology used by bridges

and routers during transmission over WANs. The PPP protocol enables data transmission over synchronous and asynchronous dial-up and dedicated lines. It is thus able to work independently from the respective physical interface. The only prerequisite that is necessary for using the PPP protocol is a completely transparent, full-duplex-capable data circuit.

17.8.4.1.1.1.3 Point to Point Tunnelling Protocol (PPTP)

The leading point to point tunnelling protocol, developed by Microsoft, was recommended in 1996 by the IETF as the standard protocol for Internet tunnelling. PPTP is an expansion of PPP. PPTP encapsulates PPP packets in IP packets, so that protocols such as IP, IPX and NetBEUI can be transmitted over the Internet. PAP and CHAP are used to control access. The data encryption standard (DES), with keys that have a depth between 56 (DES) and 168 Bit (3DES), is used as the encoding algorithm. With PPTP, the user is responsible for controlling the end of the tunnel.

17.8.4.1.1.1.4 Layer 2 Forwarding (L2F)

Layer 2 Forwarding (L2F), developed by Cisco Systems, is specially used to connect individual computers. In conjunction with PPTP, it forms the foundation for the layer 2 transport protocol (L2TP), which is a further development of both systems. L2F supports various protocols and several independent, parallel tunnels. User identification is somewhat weaker than that for PPTP and extra data encoding is not provided

17.8.4.1.1.1.5 Layer 2 Tunnelling Protocol (L2TP)

The layer 2 tunnelling protocol (L2TP) is only different in a few respects from PPTP. On the one hand, L2TP supports several tunnels, just like L2F. On the other hand, the user does not control the end of the tunnel as in PPTP, but is done by ISP.

17.8.4.1.1.1.6 IP Security Protocol (IPSec)

The IP security protocol (IPSec) was specially developed to connect two local networks. In doing so, the IPSec protects the data packets of the IP protocol from possible modification or copying. IPSec influences neither the communication protocol, nor the application program, so that routing is not interfered with. Authentication processes, that have been created with IPSec, can differentiate between data from authorised and unauthorised communication partners. The authorisation processes are based on MD5 hash algorithms with 128 bits and the encoding on the DES algorithm with 56 or 168 bits. All data traffic can be protected with IPSec and even Ipv6 can be transmitted via IPSec.

17.8.4.1.1.1.7 Layer 2 Security

The layer 2 security protocol (L2Sec) is supposed to eliminate specific weak areas that are exhibited by IPSec in remote access solutions. With L2Sec, all data packets are packed into a tunnel and then saved as a whole. 17.8.4.1.2 Encryption Standards

17.8.4.1.2.1 Data Encryption Standard (DES)

Data Encryption Standard (DES) refers to a block encryption, which encrypts the data in 64-bit blocks. The algorithm receives a 64-bit block that is not encrypted and then outputs a 64-bit block that is encrypted. DES is a symmetric algorithm. Encryption and decryption uses the same algorithm and key. The key length is 56 bits. The key is usually referred to as having 64 bits, but every eighth bit is used for a parity check and is ignored. These eight parity bits are the lowest-level bits in the individual bytes for the key. Every number that is 56 bits in length can be used as a key, and the key can be changed at any time. The DES standard was superseded in 2001 by the AES standard, since the so-called DES crackers compromised the integrity of the key.

17.8.4.1.2.2 Triple Data Encryption Standard (3DES)

The triple DES key refers to multiple encoding based on the DES. The symmetric encryption algorithm 3DES uses two keys and three DES cycles. 3DES thus works with a 168-bit key length, or 56 bits for a DES cycle.

17.8.4.1.2.3 Secure Socket Layer (SSL)

SSL was specially developed to encrypt information in the Internet and is based on TCP/IP. SSL encodes with public keys which are usually acknowledged by a third party. High security is guaranteed, since the key for decoding must be individually set and is only stored by the user and not transmitted over the Internet. The developers of SSL have designed the protocol to have two levels. One level is for encrypting the data. It permits various algorithms, including DES and 3DES, and requires both communication partners to have the same secret key that is generated for every connection. Data authenticity is also verified by a checksum test. The private keys are exchanged on the second level. Participants for the communication are authenticated, arrange an encryption algorithm and then send each other the encoded session key.

17.8.4.1.2.4 Advanced Encryption Standard (AES)

The advanced encryption standard was developed in 2001 to deal with security problems that arose from the DES standard ("DES crackers"). The AES is based on the Rijndael key and thus recognises three key sizes with 128, 192, or 256 bits. The AES is so secure, that it can create 10 to the 21st power more 128-bit keys than the DES 56-bit key.

17.8.4.1.3 VPN-Configuration

The following section deals with the configuration and set up of VPNs for various applications.

17.8.4.1.3.1 End-to-End-VPNs

End-to-end VPNs provide a direct connection between several workstations. This type of VPN can be used, for example, to safely connect customers with an online shop or to help employees at different locations to work on the same project. When using end-to-end VPNs, you must remember to install the corresponding VPN protocol on the connected computers, since the workstations are directly connected to each other and not via a VPN server. L2F, L2TP and IPSec are particularly suitable protocols for setting up end-to-end VPNs. IPSec is, however, best suited for applications that require the highest level of security.

17.8.4.1.3.2 End-to-Site-VPNs

End-to-site VPNs, or Remote-Access VPNs, are mainly used to connect external employees to a central company network. The greatest advantage of such a network is that the employees can dial into the network using any POP from the company's service provider. End-to-site VPNs can help to reduce the costs of expensive long-distance connections, since you do not need to provide a large modem pool for employees to dial in.

17.8.4.1.3.3 Site-to-Site-VPNs

Site-to-site VPNs are the classic VPN variant. Several LANs from various locations are connected with this method. This configuration is suitable to connect company networks or public utilities. All the protocols referred to here can also be used for secure connection to a company network (remote access).

17.8.4.2 Remote Access

Remote access means access to applications or data on computers from a system which is not directly connected with these computers via a LAN. Successful access requires a range of factors which are described in detail below. The first prerequisite is a **network connection** between the computer seeking access and the system on which the data or applications are located. The easiest way of establishing a connection is via modems connected to a serial interface on both systems. However, communication servers can also be used with modems connected to them. If ISDN is selected as a transmission service, internal ISDN controllers or serial ISDN terminal adapters are used. In theory, all other types of connections already described in the sections above, such as GSM, X.25 or Frame Relay, can be used. The Internet became ever more important as an economical transport medium for connections using the PPTP. PPTP allows setting up secure tunnels for so-called virtual private dial-up networks (VPDN). The second prerequisite is a **communications protocol** which is used to transport data between the systems or via the network. PPP is most important here. PPP is a special line protocol for WAN connections over serial lines. It enables LAN protocols such as TCP/IP, IPX or DECnet to be transmitted asynchronously or synchronously. The modus operandi and set up of PPP is described in further detail elsewhere. The older protocols SLIP and CSLIP (Compressed Serial Line Internet Protocol) are of little importance nowadays. The so-called **Low Bandwidth X Protocol (LBX)** has been developed for X Windows-based systems. For cases where only a small bandwidth is available, e.g., serial connections, this protocol accelerates the data transmission by means of data compression.

Typically, remote access applications are classified into one of the following three main categories: **Remote Control, Remote Node** and **Data Transfer**. All of these applications are feasible with different systems at each end. However, since the great majority of accessing systems run under a Windows operating system, only these cases will be considered below. The statements made for this operating system generally also apply to any other system.

17.8.4.2.1 Remote Control

Remote Control is used to describe operations in which a computer is externally controlled by a second computer via a WAN connection. The remotely controlled computer can be a stand-alone device or a workstation within a local network. Depending on whether the remote computer is a single-user or a multi-user system, remote operation is exclusive or can be parallel. Because data transmissions are limited to keyboard and screen display information, remote control technology permits relatively high processing speeds as long as the applications used do not involve frequent complete screen refreshes or a large number of graphics. One major application of Remote Control is the maintenance of a single computer or a LAN via a

WAN link. This provides a very simple and inexpensive means of solving software problems, and frequently dispenses with the high travel costs incurred for visits by an expert.

17.8.4.2.2 Remote Node

For a remote node solution, the remote PC becomes a node in the network into which it is dialling. Communication takes place, as in a LAN, via the network protocol operated by means of a carrier protocol via the dial-up line. The remote PC therefore has its own network address like a direct node in the network. This is either assigned when dialling-in or fixed in advance. With remote node, all applications run on the accessing computer. To ensure that the application programmes do not always have to be loaded over the comparatively slow line, it is sensible if they are also available locally.
As LAN protocols such as TCP/IP, IPX or DECnet cannot be transferred directly over a serial line, a special communication or carrier protocol is required. The protocol, which is most commonly used, is the **Point-to-Point Protocol (PPP)**. The task of PPP is therefore to transport data together with the LAN protocol. Consequently, PPP is not another LAN protocol, but is essentially a carrier protocol for the available LAN protocols. PPP performs various tasks during data transfer, such as negotiating the connection parameters, compression, determining the block size, extracting control characters, monitoring the line quality or password validation using PAP (Password Authentication Protocol) and CHAP (Challenge Handshake Authentication Protocol). With PPP, different network protocols such as TCP/IP, IPX or AppleTalk can also be used simultaneously. It is also possible to have several tasks running in parallel via a PPP connection, e.g. a telnet window for a Unix system plus a Web browser via TCP/IP.

17.9 Hardware in a WAN

17.9.1 Analogue Modems

Modems are used for data transmission in the analogue public telephone network. The modem will continue to be a significant data transmission medium for as long as a large number of the installed connections worldwide are sent to the user terminal via the analogue telephone network rather than having direct access to the ISDN network.
Modem is a word coined from the original functions of **mo**dulation and **dem**odulation. Today's modems have a number of additional functions such as error correction, data compression and - together with appropriate communications software on the connected PC - a fax function and an automatic answering machine function.
The modem receives computer data from the computer in digital format and converts them into analogue signals suitable for transmission through the telephone lines. The receiving modem demodulates these analogue signals, reconverting them back into digital format. The

installation arrangements vary according to the type of modem; modems can be installed as slot cards designed to be inserted in the PC, or they can be external devices, which are connected via the RS-232 serial interface. Yet another, more recent alternative is a modem with a PC-card interface (PCMCIA). These are very popular, as virtually all modern portable personal computers are equipped with the corresponding connection.
Because modems can also transmit and receive facsimile (fax) transmissions, they have effectively driven separate, individual fax cards from the market.
The Hayes-compatible command set has become the standard tool for configuring and controlling modems. It is also referred to as the AT command set, as the commands are prefixed with the letters AT.

17.9.1.1 Transmission Modes

The two major factors that distinguish the wide variety of modems currently on the market are their transmission mode (asynchronous or synchronous) and transmission rate.
Asynchronous transmission uses characters as the basic data unit. Each transmitted character is assigned a supplementary start bit and one or two stop bits. In synchronous transmissions all of the data bits are sent in a continuous sequential stream. Synchronisation is achieved via separate clocking lines. The corresponding synchronous connections with the appropriate protocols are nevertheless required for synchronous modems and consequently asynchronous modems are normally used.
The unit used to measure a device's transmission rate is the number of bits per second (bit/s). The baud rate, as a unit, indicates the signalling rate over the transmission route. However, as more recent transmission methods no longer merely use frequency but also amplitude and phase to convey information, several bits can be transferred per baud.
Modems can never be fast enough for transmitting large data volumes. The high procurement investments for faster modems will soon pay for themselves due to the low operating costs. Even the V.34+ transmission standard, which specifies a gross transmission rate of 33,600 bit/s, has not yet reached the maximum capacities of traditional telephone lines although experts predicted that a transfer rate of 28,800 bit/s would be the maximum. The speed category 56,000 bit/s is standardised as V.90. Using V.90 allows a download transmission rate of up to 56 kbit/s. A prerequisite for this is a so-called 56k host. The higher transfer rates are achieved by dispensing with the analogue conversion of the signal between the 56k host and the digital switching centre. Only the user's connection continues to be analogue.
Older, manufacturer-specific protocols such as PEP, HST, V.32terbo and V.Fast are no longer of any relevance, since the V.90 standard surpasses all of them with its transfer rate.

17.9.2 ISDN Adapters

Communication between existing applications is generally adapted to serial analogue modems which recognise the Hayes AT command set. To use ISDN, in such cases, the application of an **ISDN terminal adapter** with a serial interface is recommended. As the ISDN terminal adapter is an external device with a serial port in accordance with RS232, it can provide a direct replacement for an analogue modem. Similarly to the analogue modems, control is performed using the extended AT command set. For this reason, ISDN terminal adapters are suitable, in principle, wherever data is to be transferred from the serial interface via ISDN. Examples of usage include workstations without an internal ISDN controller, as well as remote bridges and routers with a serial interface. ISDN terminal adapters are also suitable for use with serial communication servers to implement high-performance remote access applications via ISDN. Note that not every serial interface is appropriate for the high transmission rate of 64,000 bit/s (up to 256,000 bit/s or more with channel bundling/data compression) used in ISDN. Refer to the technical specifications for the particular interface for more detailed information. In any event a serial interface with incorporated data buffer (UART 16550-compatible interfaces) is of advantage.
It should be added that optionally most ISDN terminal adapters can also be operated through a CAPI interface in addition to the AT command set. More information on CAPI can be found in the CAPI section.

Common telephones, answering machines, group 3 facsimile devices and modems are analogue end devices and cannot be operated directly on the ISDN connection. In these cases, an **ISDN terminal adapter with an a/b interface** is the solution. It allows the use of available analogue end devices even with ISDN. Small ISDN PBX systems offering up to 16 analogue PBX lines are, in principle, equivalent to a 16-port terminal adapter. However, the analogue lines limit the quality and transmission speed of this type of device.

ISDN controllers represent the most inexpensive and simplest option for entering the world of ISDN data communication. They can be obtained for most bus systems and hardware platforms. Transmissions through BRI as well as PRI are possible, depending on the card. However, the most common form is PCI insert cards for PCs.

17.9.3 Routers

Routers, in contrast to bridges, also link networks of different topologies in OSI layer 3. They are the cornerstones of well structured LANs and WANs. Their ability to route different protocols and network types provides optimal traffic control and network load distribution. Routing becomes necessary when communication between stations in different subnetworks

needs to be established.

Routers are not protocol-transparent, but must be able to recognise all the protocols used, since they translate information blocks according to protocol specifications. Thus, the protocols used must be either routable or be correspondingly translatable, i.e., re-packageable into other protocols. Since not all protocols can be routed, most routers are also able to bridge packets.
In comparison with switches, routers ensure improved isolation of the data traffic, since for example they do not forward broadcasts as standard. However, routers generally slow down the data transfer. Still, in branched network connections, especially in WANs, routers actually transfer data more efficiently. On the other hand, routers are usually more expensive than switches. Therefore their implementation should be carefully analysed for each specific case.

The logical addresses in a network can be evaluated by routers and the optimal path (route) from sender to receiver can be determined with the aid of internal routing tables. Routers adapt the packet length of the data to the maximum possible within a network segment, i.e., they change the packet length for example at the transition from Ethernet to X.25. Routers not only modify the packet length, they also adjust transmission speed at the transition point between LAN and WAN. To this end they need a suitably large buffer, which in most models can be flexibly configured.

Many different types of routers are available. **Local routers** are primarily used for security considerations, e.g., if communication in a LAN is to be restricted to specific nodes. Alternatively, they are used if the data load on a network is to be reduced, that is to say, when collision domains and thereby the number and distribution area of broadcasts are to be minimized. Local routers are available for all Ethernet speeds.

Due to the large number of ISDN lines available and ISDN's attractive tariffs, ISDN routers are often used as **remote access routers**. However DSL and, given commensurate data volumes, transmission technologies for up to 2 Mbit/s are also used for remote access. To ensure redundancy and thereby high availability, ISDN connections are provided as backup for remote access routers of high transmission rates.
The router types described so far are all based on a combination of hardware and software, but there are also pure software solutions running on servers, workstations or PCs. What solution is most suitable for a specific application must be evaluated for each individual case. The basic components and the technical characteristics of routers are discussed below.

17.9.3.1 LAN Interfaces

Most routers are provided with one or more LAN interfaces, according to topology for Token-Ring, Ethernet, 100 BASE-T Fast Ethernet, FDDI or ATM. For connecting corresponding media either multipurpose ports (e.g. Ethernet AUI, BNC, RJ-45) are available or the connection is implemented as a slide-in unit which can therefore be customised for the requirements.

17.9.3.2 WAN Interfaces

Today, a number of companies offer WAN lines supporting different transmission speeds. The tariffs as well as the interfaces offered vary correspondingly.
For smaller networks (e.g. workgroups), connections with transmission speeds of 64 Kbit/s are advisable. There are, of course, applications that can manage with lower transmission rates. If higher transmission rates are required, the best solution is the use of E-1 connections (in principle ISDN primary multiplex connections), which are common in Europe and offer a transmission speed of 2,048 Kbit/s, or connection via DSL.
Routers are equipped with one or more WAN ports, which are either built-in, or in the case of modular versions can be upgraded with corresponding interface modules. Standard physical interfaces for synchronous mode are RS449, V.35 and X.21; for the asynchronous mode, the RS232 interface. The S0 interface is used for ISDN connection.

17.9.3.3 Protocol Support

Routers route one or more network protocols. The network protocol most often supported by routers is IP. Due to the widespread use of heterogeneous networks, most routers also offer multiple protocol capabilities. Depending on the router and the equipment, IPX, DECnet, AppleTalk, OSI, XNS, VINES and Apollo Domain may be supported in addition to IP. Depending on the router and the equipment, DECnet, AppleTalk, OSI, XNS, VINES and Apollo Domain may be supported in addition to IP and IPX. All non-routable protocols such as LAT and NetBIOS will be bridged, provided the router has this function.

17.9.3.4 Software

Router software is generally stored in Flash PROMs, which permit easy upgrading. When booting the router, the software is then loaded into the RAM and executed. The range of software supplied differs according to the strategies of the manufacturers. Routers are either supplied with a basic version that can be expanded if needed, or come with the complete software package.

17.9.4 Firewalls

Firewalls are computers which operate as dedicated security gateways between networks and are intended to increase security in a corporate network by means of various mechanisms. They assume the central control of company access to the Internet by authenticating users on the basis of their identity and allowing passage only to approved services. The firewall rules define the criteria by which firewalls allow passage or block data. If such a rule is contravened, the data flow is interrupted and, depending on configuration, an alert is raised in case of serious contravention.
A firewall generally consists of multiple hardware and software components individually configured depending on user-requirements for services and security. Security management, as well as monitoring and control functions, are significantly simplified by concentration of access on a single component.

17.9.4 1 Firewall Technologies

17.9.4.1.1 Packet Filter

Packet filters monitor at IP level whether or not a packet has access authorisation, checking sender and receiver address and the IP service. The hardware demands made on packet filters are relatively small and therefore packet filters are often implemented on routers.

17.9.4.1.2 Stateful Inspection

Stateful inspection is a packet filtering technology in which the data packets are analysed and the connection status is included in the decision. In this technology the data packets are analysed, during transmission, on the communication level and are stored in dynamic state tables. The decisions on forwarding the data packets are made on the basis of comparing multiple data packets and by determining the correlation between data packets belonging together. Therefore in security-relevant applications, firewalls incorporating stateful inspection technology are superior to pure packet filter firewalls.

17.9.4.1.3 Circuit Level Gateways

Circuit level gateways associate packets to existing TCP connections. They operate with a subnet and an external and an internal router with a host through which all communication takes place. The circuit relay concept is comparable with a packet filter, but operates at a higher level of the protocol stack. A computer wanting to set up a connection has to register with the host and prove it has access authorisation. By virtue of the separation of internal

and external network, no internal IP addresses can be read out from outside.

17.9.1.4 Application Gateways

Application gateways represent the most secure, but also the most complex alternative to a firewall. In application gateways, adequate security mechanisms are implemented across multiple levels. They can decouple networks logically and physically and expect prior identification and authentication of every user.
The application gateway receives the data packets at the respective ports. If only a service is to be possible via the port, software (a proxy server) is activated on the application server and transfers the packet from one side of the network to the other. From the viewpoint of the accessing user, it appears as if he were communicating with the server process of the service on a destination computer. In point of fact, however, the user communicates with the proxy server acting as mediator on either side so that a connection between destination computer and user is never established.

18. LAN Core Solutions

18.1 Introduction

Nowadays fast communication and common access to information, ideas and resources is becoming increasingly important. For this reason data networks are vital for business success.

This economic scenario requires modern companies to be more competitive and to increase their productivity. Optimisation of the data flow within the company as a whole has become an important constituent of this process. The provision of bandwidth-intensive content and sophisticated multimedia applications calls for efficient use, signifying, in turn, the need for a stable network infrastructure.
These new requirements mean that networks are designed in such a way that they support the current business operations, but are also geared to future growth. The increase in switched networks, substantially reducing network complexity, is an important development for companies of every size. Routing technologies formerly playing a predominant role are taking a back seat. With high-performance LAN traffic, switched networks provide a cost-effective basis for providing applications of the following generation, including video streaming and sophisticated Web services.
The currently available Layer 3 switches perform all the operations of a conventional router. They can be used at any point within a network core or backbone and operate in the network layer of the OSI (Open Systems Interconnection) reference model, ensuring significantly higher data transfer. The new switch generation has been developed specifically for operation in higher OSI levels. It boasts even more flexible functions and hence copes with a wide spectrum of data traffic modes. In addition, it enables enhanced network management. The latest switches are even more intelligent and achieve a higher throughput. They offer substantial advantages for all network types, for example user authentication and Quality of Service (QoS). Layer 4 switches, which identify and forward data traffic on the basis of multilayer packet information, permit data traffic administration at network level and not just in the central core.
The advantage: Faster replies as well as maximization of server use and availability. These enable you to identify and establish priorities for critical data traffic based on application.
In the last ten years companies have been constantly harnessing new potential for building and for the administration of corporate networks. This development is forging ahead to support critical business requirements. Switched networks are important to ensure optimum data access and services. At the same time they are guarantors for high-grade security, mobility, provision of information and, what is so important in the present-day business world: Teamwork. This chapter is intended to give you an understanding of the technologies currently available.

18.2 New Strategies for Networks

Nowadays a company relies on the network living up to the demands posed by daily business life: Canvassing customers, transacting business and gaining a competitive edge. How can a network nevertheless keep pace with the demands of the daily growth in business - especially without exceeding the financial limits?

These are the challenges confronting many network managers today. At the same time there is a call for the business requirements to be met within the framework of the budget available. This can only be done if the planned network is expandable without going to great expense. Future network expansions should be possible without any major restructuring, because such are sure to be considerably more costly. To achieve this, the selection of the network core is decisive.

18.2.1 Structure of a Network Core

Changes in corporate structure, whether they take the form of a merger, takeover or simply growth, put considerable demands on the network backbone. In an ideal case, the potential afforded by the budget should meet the increased data traffic demand due to new locations and new users, yet without imposing any constraints on future growth. Such demand often comes entirely unexpectedly. When, for example, a large number of contractors work together on a large project, this involves an increase in data traffic. These short-term peak loads should not impact the long-term network strategy. Furthermore, the operator has to be in a position to set up new technologies cost-effectively. And, at the same time, the focus should be on secure connections and more cost control, despite constant expansion or redesign of virtual LANs.
What are the crucial economic factors in developing or selecting core technologies? Here consideration has to be paid to the individual requirements of each company.

18.2.2 Cost Control

The requirements are: Achieving more and more and securing growth with tight resources. That is why low purchasing costs are a vital factor. In the present-day business environment it is equally important for the network to be geared to the future for the long term. The overall operating expenses should form a significant part of the equation. If it is possible to achieve savings here and at the same time ensure fail-safety, this has a significant impact on the return on investment. Companies then secure better control over their expenditure.

18.2.3 Scalability and Flexibility

The scalability of networks is an important factor in the overall resources. Where extra functionality is needed, special requirements and funds should be carefully balanced, thus optimizing future-proofing of a network. In many companies the demands on capacity are dynamic and unpredictable. That is why flexibility figures large in the configuration of the network. The set-up of a scalable network gives companies maximum security for existing investments.

18.2.4 Security

Switched networks indeed have their advantages, but they have involved a heightened security risk at the very point where the network is not protected. End-to-end security has today become a critical task in network administration, particularly when the phenomena accompanying e-business, the use of extranets and the increase of data traffic in private networks are included. Today networks have to protect users, systems and intellectual capital more than even before. This calls, above all, for extended authentication and control.

18.2.5 Availability

The workstation today looks entirely different from a few years ago. It is no longer simply a nuisance if the network becomes slower or even fails, it can seriously jeopardise critical business operations. Most companies nowadays employ hardly any IT staff trained in system administration. Network failure therefore has catastrophic consequences. Resilient network infastructure is indispensable for problem-free communication and uninterrupted operation. The network also has to be capable of self-regenerative operation, it has to display high-performance and support a multitude of applications and high data traffic. High-grade availability and fault tolerance depend on the configuration and functions of the devices distributed over the entire network. To ensure access to critical network applications, chassis-based switches and stackable configurations are of advantage.

18.3 Current Solutions

The network capacity in very large companies is as a rule predictable or static. The network core is affected only by small or scheduled modifications. These generally require the scalability, availability and performance of a centralised core backbone. Large networks demand extended multilayer switching functions in order to support sophisticated applications and to prioritize real-time data traffic such as voice communication and video. Central high-end core switches offer companies the necessary processing functionality, port density and storage

capacity. When networks demand support for routing protocols (e.g. IPX and Appletalk), this is certainly the most effective solution.

In a rapidly expanding company of medium size with unpredictable capacity requirements, a conventional network backbone is today put to insufficient use and is tomorrow already overloaded.

Let us look at the following example. The chassis-based implementation comprises a routing fabric with several slots for interface modules. When users are added, you just add extra modules to the chassis. The increase in network capacity should be thoroughly planned in advance so that the most appropriate chassis is selected. When the capacity is not used to full, the upfront costs are higher. If, however, the network outstrips the capabilities of the switch when it is expanded further, you will need a second, similarly configured switch.
If the anticipated increase in capacity can be predicted with accuracy and a network device having the number of ports matching the requirements is available, a chassis is the right solution. It is important to select the network type that meets the requirements best. For most companies a strategy with end-to-end availability, covering the entire network, should be the real target. For growing companies a dynamic network architecture is important so as to attain a realistic equilibrium between support of the present business operations and future growth. How this can be done is explained below.

18.4 Networks with Distributed Core

New generations of devices are hitting the market with the need for scalability and output bandwidth. Important functions formerly associated only with the core switch are now moving in the direction of network edges. Scalable switches, wireless access points and further intelligent devices in the edges sector of the network now ensure service prioritization for voice or video and end-to-end performance. Hence the information spreads to an increasing degree over the entire network and offers every user immediate access to the data wanted. Companies thereby secure easier management of branches. At the same time they support programmes for teleworkers and provide secure, cost-effective remote access.
This approach relates to distributed core networks. A distributed network is no longer supported on a central control point. It is based on the principle of aggregation instead of hierarchy. In this way network providers can add devices whenever increases in performance or further ports are needed, meaning they expand in steps! In a network with a high degree of distribution, all switching and routing is managed by way of the various devices.
If one withdraws the intelligence, functionality and performance of the core from a single, central, high-end core platform and distributes them between several switches, one can ensure that the core does not have a single point of failure. A flexible core configuration dispenses

with expensive and complex redundancy functions. Thus the danger of complete network failure is substantially reduced.
XRN technology (eXpandable Resilient Networking) from 3Com is a new solution based on this concept of a distributed fabric. The performance features and functionality of XRN-enabled switches deliver practical scalability to network managers for the design of high-performance core backbones of high availability, offering redundant connections, protocols and routers. The scalability of XRN technology signifies:
Expansion on demand only! This reduces the up-front costs and you safeguard network investments.
When XRN technology is employed, multiple interconnected Gigabit switches behave as a centrally managed, distributed switching fabric that grows with the network, without the physical limitations of a centralized core device. Core functions such as QoS and authentication are within the distributed fabric and hence provide for top performance. Every switch in the fabric constitutes a fast Layer 3 and Layer 4 switching entity. The performance is increased in line with the expansion of the network, as every switch in the fabric physically forwards the data traffic. The individual processors do not have to bear the entire routing load. The consequences are: Unprecedented performance and resilience.
As XRN technology supports link aggregation in the entire distributed fabric (distributed link aggregation), this provides superior performance and availability. The distributed resilient routing technology also permits high-performance forwarding between the interconnected switches. This ensures router redundancy for the core backbone. Since the interconnected switches compose a single entity, administration across the network is much easier, freeing IT resources and reducing administration overheads.

18.4.1 XRN Components

18.4.1.1 Distributed Link Aggregation (DLA)

DLA offers the possibility of implementing link aggregation via ports extending in an XRN distributed fabric to two units at present. DLA permits layer switches to be configured with paths to both core switches. Optimum utilization of available capacity is ensured.
DLA supports standard-based IEEE 802.3ad link aggregation. Devices supporting this standard are directly connected to the XRN distributed fabric and immediately benefit from the enhanced performance and resilience.

18.4.1.2 Distributed Resilient Routing (DRR)

DRR is a sophisticated routing implementation. Two interconnected switches in an XRN distributed fabric behave as a single active routing entity. DRR provides for intelligent distribution of the routing capability across both switches in the distributed fabric. Maximum utilisation of routine performance and optimisation of bandwidth capacity. Entirely new potential for network design opens up for network managers faced with dynamic requirements. In the event of switch failure, the other switch automatically takes over, averting network outage and eliminating any requirement for end-station reconfiguration. Since every switch in the fabric provides Layer 3 local forwarding for connected devices, this also results in significant performance benefits.

18.4.1.3 Distributed Device Management (DDM)

DDM is the control system of XRN technology. It is responsible for distributing management and control information across the XRN distributed fabric. Regardless of which edge devices are connected, DDM allows the entire XRN distributed fabric to be managed as a single logical entity via a single IP address using SNMP, TELNET or Web management.
Management tasks, such as software upgrades, VLAN configuration, spanning tree parameter modification, multicast filtering and QoS configuration are all performed across the distributed fabric, minimizing complexity and administration overheads. In addition, the management IP address is shared across both units, ensuring continuous device management and monitoring in the event of an outage in one of the interconnected switches.

18.5 Setting up a Distributed Fabric in a Network

The implementation of a network with distributed fabric opens up entirely new potential for network design to network managers confronted with dynamic requirements.

18.5.1 Basic Network

Basic networks usually demand minimum outlay in network administration and management resources. They simultaneously offer a high-performance, highly scalable and redundant network solution, enabling business expansion. Special users and key applications require redundant connections to the core switch and to business-critical servers. Low costs are a further decisive factor.
A typical example of appropriate architecture would be a Layer 2 infrastructure with a network architecture on two levels - Layer 2 - Layer 2 (i.e. workgroup switch to core switch), scalable and capable of expansion to form a full IP routing network without the otherwise usual

complexity.
For basic networks the implementation of a distributed fabric offers higher security for network availability. In case of a further business expansion and increasing user numbers or subdivision of users into different departments, with core switches Layer 3 switching software is easily activated and thus ensures minimal interruption of operations. Additional network components due to increasing office area - whether permanent or temporary - can be administered more cost effectively. Furthermore, authentication of the users before they access network resources is possible.
Advantages:
- Delivers the right performance level for the given applications at low cost
- Supports network telephony where needed
- Low demands on resources for network administration and maintenance
- High network availability from the edge all the way to the core
- High resilience through the deployment of appropriate workgroups and core switches
- Dual home trunking function provides for comprehensive networking.

18.5.2 Standard Network

Standard networks are usually based on the IP protocol. They often have to be converted to Gigabit Ethernet in the core or at the aggregation points while safeguarding existing investments. The more intensively Internet is used, the more important are redundancy/resilience as standard and an intuitive network administration. Standard networks often extend to several buildings in proximity as well as to remote sites. A common network topology is a network architecture on three levels - Layer 2 - Layer 3 - Layer 2, comprising workgroup aggregation, core connectivity and server aggregation.
A Layer 3 Gigabit Ethernet solution on three levels consists of workgroup aggregation, core connectivity and server aggregation.

Advantages:
- Economical performance in Wirespeed from the edge all the way to the core with 2:1 blocking
- Supports network telephony where needed
- Low demands on resources for network administration and maintenance
- High network availability from the edge all the way to the core
- High resilience through the deployment of appropriate workgroups and core switches
- It is easy to assign priorities to applications and to block them in the network
- Dual home trunking function provides for comprehensive networking.

18.5.3 Network Expansion

Extended networks demand high-grade network functionality relating to Layer 3 switching. These corporate networks can be used through a common infrastructure for performing all business applications (voice, data and video). The support of multiple protocols (as well as Multicast support) and a larger number of VLANs than usual are often necessary.
Availability, resilience and redundancy are extremely important for extended networks, because outage even of short duration can entail substantial loss of business. High performance is an important point, encompassing the capability of supporting complex applications, e.g. video streaming, CRM, ERP and Unified Communication.

The possibilities arising from the use of extended networks at multiple sites and from the utilization by many remote users reinforce the demand for a secure and resilient VPN service. Support of multiprotocol routing and the possibility of controlling the flow of data traffic by prioritisation, classification and access lists are important factors.
For extended networks, a Layer 3 Gigabit Ethernet network on four levels, comprising workgroup aggregation, core connectivity and server aggregation, is the right solution. This ensures the acceptance of legacy protocols and the transition to a seamless IP infrastructure.

Advantages:
- Complete redundancy of edge, workgroup, core and at the server avoids a single point of failure
- Stacking resilience
- Trunking at stack level (dual connection to the workgroups)
- DLA (distributed link aggregation) supplies XRN resilience
- Redundant switching fabrics offer complete redundancy and enable rapid switching
- Client availability through VRRP support in the core
- Dual connection of the servers to the server aggregation switches ensures 100% availability of the server farm
- Prioritisation of the desktop switch applications makes for end-to-end perfomance
- Network telephony is realisable immediately.

18.6 A New Direction for Corporate Networks

Efficient utilisation of WAN bandwidth, the desktop and workgroup switches automatically rerouting the web data traffic to a web cache and thus cutting overall WAN costs.
The further development of switched networks is running at full speed since there is a rapid increase in bandwidth demands. Therefore the need for expanded network functions must be met. Network connections continue to be wired or wireless. Flexibility and mobility must be

ensured without detracting from security or user-friendliness, enabling the user to establish connection to a network from remote sites.
The wireless "roaming function" can also be implemented at the company premises. Fast and secure access to business applications and services can be gained from any location.
To cut overall operating expenses, 3Com's solutions support a new direction in corporate networks. Earlier, network intelligence was available only to core devices, today any switch, wireless access point and further endpoints of the next generation can tie in, bringing functions like traffic prioritisation, authentication and encryption closer to users.
"A pay as you grow" cost structure is the result of an intelligence distributed across the network components. It represents improved cost control and administration as well as targeted provision of new services and functions.
The enhanced interoperability between the various network devices permits more effective use of existing network investments and the addition of voice, data or wireless elements where needed. Diverse applications, including Voice over IP (VoIP), Unified Messaging and a series of business-specific tools can benefit from the strength of a distributed network design without high cost.

18.7 Résumé

A network has to create a stable basis for the wide spectrum of applications and services providing strategic benefits and promoting business success. The new switched networks are meeting challenges in their development in respect of scalability, extended QoS and security, simple administration and minimized operating costs.
By virtue of the distribution of network intelligence to switches of the next generation, wireless access points and further intelligent devices in the edge sector of the network, you can build scalable, administrable networks permitting cost-effective growth, without compromising performance or availability.
3Com Corporation has been developing innovative networks for over two decades and always comes up with entirely novel solutions, allowing you to meet the demand for new services and functions driving productivity, to achieve better cost control and reduce administration complexity.

19. Input Devices

19.1 Keyboards

For normal use in an office, a keyboard is expected to provide long service life, high functional reliability as well as ergonomic and secure handling.
The mechanical stability of the keys is decisive for the long-life functionality of a keyboard. As a rule, keyboards have gold metal-reed contacts, which ensure a long service life even with continual usage. In contrast, low-cost keyboards use only a thin contact foil, which is delicate and has only a short lifetime.

Ergonomic keyboards - e.g. the Microsoft Natural Keyboard - ensure a healthy position for hands and wrists, thus reducing the risk of injury (carpal tunnel syndrome CTS or tendovaginitis). The result is complaint-free effective work. In general, it is important to guarantee that the seat, work desk and keyboard are correctly positioned in relation to each other.
Keyboards with other integrated additional functions all offer the advantage that they provide several functions while occupying only one interface on the computer. There are a variety of keyboards which combine several functions in one unit. Amongst these are keyboards with chip card readers, bar-code readers, magnetic card readers and multi-functional card readers. Another alternative is freely programmable keyboards for reading credit cards, EC cards and bar codes. For access control, chip card reader keyboards with integrated fingerprint sensors can be used.

19.2 Mice and Trackballs

A **mouse** essentially consists of a rolling ball and sensor mechanics/electronics, which detects the rotary motion of the ball, converts it into a stream of data which is transmitted to the computer. The latest models work without a ball. Instead of having the traditional moveable parts, an optical sensor detects the movement of the mouse. At the core of this technology is a small chip, which contains an optical sensor and a digital signal processor (DSP). The sensor produces several hundred images per second from the desktop surface, which the DSP uses to calculate the movements. Since the problem of dust, dirt and grease, which the moving parts pick up, has now been overcome, optical mice no longer need to be cleaned. Additionally, the mice can work on almost every desktop surface, and a mouse pad is no longer necessary. Additionally, the mouse has a number of buttons on top, which are either used as function keys by the application used, or can be assigned with functions by the users themselves.
The current mouse generations, e.g. from Genius, Logitech or Microsoft are equipped with a

small wheel or key rocker in addition to the function keys. With this, the user can zoom and scroll within various applications without moving the mouse. Scrolling through the information windows becomes even more comfortable and Internet navigation, e.g. with Microsoft's Internet Explorer or Netscape Navigator, is faster than ever.

Mice without restrictive cable connections are free to move on the desktop surface. Today, wireless technology for cableless mice has replaced infrared technology. The advantage of the wireless solution is that no direct visual contact with the receiver is necessary, the wireless mouse works within a radius of about 2 m.

Should desk space be a problem, a **trackball** is a good solution, as the input is performed by rotating the ball located on top, and pressing the buttons. In essence, a trackball is an upside-down mouse, which performs the same functions as a mouse, but needs less desk space.

Standard input devices, such as a mouse or trackball, are often unsuitable for 3-D applications, as they only allow the simultaneous control of two degrees of freedom. Many 3-D applications, such as virtual reality and 3-D modelling, require the control of six degrees of freedom. To meet the demands of such applications, special 3-D mice have been designed. They combine the functions of an ordinary mouse with those of a device for motion control of 3-D graphical objects.

19.3 Scanners

Scanners have become an important aid for importing data in the form of printing texts, handwritten documents, photographs or drawings. Scanners operate according to the following principle: A light-sensitive CCD element (charge-coupled device), driven by a stepping motor, moves over the original document while it is being illuminated underneath by a lamp. The CCD element detects the brightness differences for each colour and transforms these into voltage values. The analogue voltages are then translated into digital information by an analogue/digital converter and transmitted to the computer.

Depending on the scanning method used, the CCD read head either passes over the original document once (one pass), or once for each of the primary colours red, green and blue (RGB) (three-pass). One-pass scanning can be carried out in two ways. Firstly, by directing a white beam of light onto a CCD which performs a direct RGB colour differentiation by means of filters. Secondly, by passing the beam of light through a prism which splits it into the three colours (RGB), and then each beam is directed on to one of three CCDs. The first method mentioned has become the standard for scanning.

The CCD technology exclusively developed further by HP enables high resolution and high speed. It consists of two sensors (CCD) in one: A linear 600 dpi CCD sensor for high quality scans at top speeds and a 2400 dpi CCD sensor for high resolution scans. In most cases no more than 600 dpi resolution is needed, where the 600 dpi CCD is then used. For films, negatives and enlargements, however, a resolution of 2400 dpi is required. In such cases the 2400 dpi CCD sensor is then used to secure performance including big-scale enlargements without any loss in quality of the scan.

There are a number of criteria to be considered in selecting a scanner. The **resolution** of a scanner is measured in dpi (dots per inch). As a general rule, the higher the resolution, the better the quality of printed image. The dpi figure refers to the number of pixels per inch recorded by the sensors (1" = 2.54 cm). If an image is recorded with 100 dpi for instance, this means that every inch is divided into 100 pixels. In terms of surface area, the scanner provides 100 x 100 = 10,000 pixels per square inch, at 200 dpi, this figure reaches 40,000 pixels. This illustrates that a higher resolution permits a printed image of higher detail.

Note, however, that doubling the resolution results in quadrupling of the data volume. Resolution as described above, refers to the physical resolution of a scanner. It determines the actual scanning of the pixels. In addition, a mathematically calculated so-called interpolated resolution of a scanner is always specified. Software allows further intermediate values between two recognised dots to be calculated by means of **interpolation**. This provides the computer with additional pixel information which improves the quality of the printout.
Another criterion for rating a scanner is the **colour or bitdepth**. It determines the number of shades of grey or colours recorded during the scanning process. A 1-bit scanner for instance, can only distinguish between black and white. On the other hand, a scanner with 8-bit colour depth is capable of differentiating 256 shades of grey or colour (2 to the power of 8). With a 24-bit scanner, this figure reaches 16.7 million colour possibilities. 48-bit colour depths are common at present.
Twain has become the most accepted software interface standard for scanners. Twain (**T**oolkit **w**ithout **a**n **i**mportant **n**ame) is a software interface standard created by leading peripheral manufacturers so that scanners from different manufacturers can be used from within a single application. This means that any scanner with a Twain driver can be accessed by any programme with Twain support. Twain includes an **API** interface (**A**pplication **P**rogramming **I**nterface) and a special protocol that sets up the connection to the scanner's Twain driver and controls the scanner process.

Twain emphasises the need for powerful scanning software. The scanner provides digital data to the computer, representing only what it sees. What the computer does with this and how the data is processed further, depends on the software. A particular problem is character

or **text recognition**. Scanned-in text must first be processed with a special text recognition (OCR) programme, so that it can be further processed as a text file.

Usually, a standard software package is supplied together with scanners. Note, however, that these scan software packages can differ considerably in the performance they offer. Often only a Limited Edition of the software is supplied.
It is evident that the software is a decisive factor in user-friendliness. With the HP scan software and the "copy, e-mail, fax or archive" operating buttons, HP offers a simple solution for reading extensive texts into the PC and for processing them further. Images scanned into presentations and letters also make them more attractive - without involving effort or complicated handling of the software and the device.
A further important point for scanners is whether there is a **transmitted light option** for scanning in transparent media (e.g. slides, films). As a rule, a transmitted light accessory with fluorescent lamps is mounted on the scanner.

19.4 Bar-Code Readers

Nowadays, bar codes can be found in every section of the merchandise economy. They permit small quantities of data to be imported quickly, easily, and above all reliably, into computers for further processing.
There is a large variety of bar codes in the various application fields. The codes appear as a sequence of narrow and wide bars or gaps. By optical sensing, the different intensities reflected from the dark bars and light gaps are converted in the receiver into corresponding electronic pulses. These are then translated into computer readable characters by a micro processor, and then transmitted.

Bar-code readers connected between the keyboard and the computer or display terminal via the keyboard lead are in widespread use. Such bar-code readers require neither an additional interface in the computer nor a separate power supply. Specific programming or changes in the hardware or software are not necessary, because the bar-code reader data is treated just like keyboard input. The computer does not distinguish between data from the bar code reader and data from the keyboard. Switching between the two input modes is performed automatically.

Wireless solutions have become the modern alternative, since their use can considerably reduce the time and effort involved, particularly for handling awkward items (e.g. pallets of monitors) in the incoming goods department. The wireless base station is installed between the keyboard and the computer and bar codes are recorded in a way analogous to that of the conventional reader units. Major benefit is the mobility, since wireless scanners have a

maximum range of operation of about 30 m.
The various bar-code readers differ in terms of the number of readable codes, in terms of their code-recognition function (automatic or adjustable), reading distance and reading mode. Available bar-code readers include pen, CCD and laser gun designs. Reading pens are ideally suitable for small data volumes. A red LED serves as a light source. In order to recognise 4 codes correctly, the pen must be brought into direct contact with the code that needs to be read.

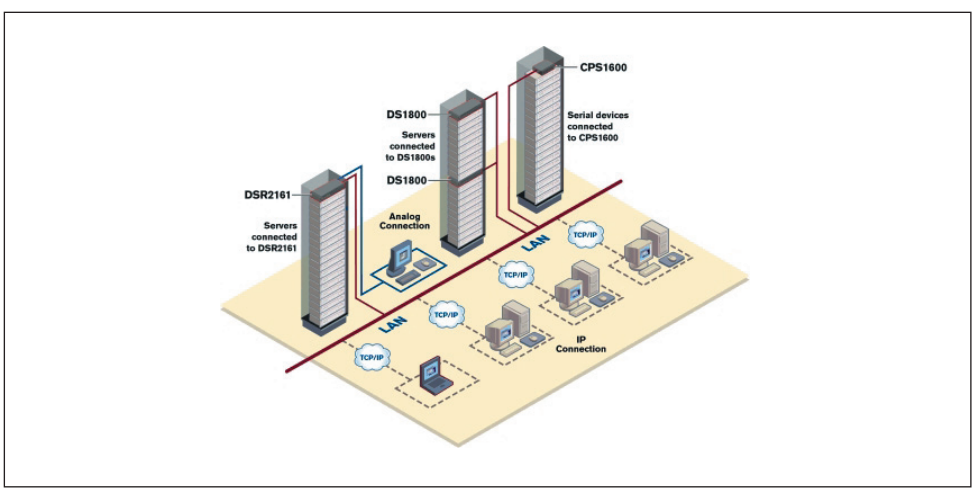
A CCD bar-code reader is a near-contact device. The reading distance is between direct contact and a few centimetres. Laser bar-code readers are capable of reproducing codes from large distances (70 cm) and from uneven or curved surfaces. These readers employ a laser diode as a light source.

20. Data Communication

20.1 KVM (Keyboard, Video, Mouse) Switching

The function of every **KVM (keyboard, video, mouse) switch** is based on a simple principle: Several servers can be controlled and monitored via only one or multiple workstations, as if one were sitting directly in front of each individual unit. For this purpose, the keyboard, video, and mouse signals are just tapped directly at the computer and transmitted to a central workstation, from where a system administrator can define BIOS settings, carry out system diagnostics or restart his computers. Even if network protocols or operating systems no longer function.

This technology, originally analogue, has already helped many computer centres to save superfluous display workstations, to gain more space in the computer centres, to cut operating costs, and, above all, to reduce the economic loss for server downtime.
These systems were limited, however, with a view to the number of servers to be managed and by the restricted potential for adding further operator desks. Here the so-called **matrix switches** provide the solution: Their modular design allows the number of servers and users to be successively increased. However, access is possible only through analogue cabling within a radius of a few hundred metres from the computer centre. The **Digital Data Centre Management over IP** forms the cutting edge technology today: No limitations on distance are posed for remote access via IP connections. Administration of any number of servers and serial devices is possible. And there is no limit to the number of users either.



21. Terminals

PCs or workstations are permanent fixtures in today's workplaces. The use of conventional display terminals, which are connected either directly or through a terminal server to the computer via the serial interface, is in decline. Their use is limited to applications which do not require any graphics, which do not place high demands on the computer and where the cost per workplace is the decisive factor. The main competitor of these terminals, particularly in the Unix domain, are powerful workstations or PCs which can process data locally. However, these are significantly more expensive and require more intense maintenance efforts. In terms of their functions and connectability, X Window terminals rank between display terminals and workstations.
Network Computers (NC) have, in the meantime, become available on the market. However, they have not yet been properly accepted due to the fact that an alternative is available, the so-called Windows-Based Terminals (WBTs).
At first glance, both NC and WBT appear to be a step backwards, since, as with X terminals, data is stored centrally and applications are presented via a uniform user interface. On closer inspection, the advantages with respect to decentralised programme processing become obvious.
The selection of which type of unit and system to be used is especially important because of the high initial purchase price and the subsequent costs for technical support, maintenance, energy consumption, etc.

21.1 Alpha Terminals

Alpha terminals are best used in situations, where graphics capability is not necessary. This means, for example, when the terminal is employed only for database registry. To communicate with the connected computer, the respective terminals only need to support the required protocols. For historical reasons, different manufacturers have developed specific protocols and terminals for communicating with the computer in their system environments. Alpha terminals are available in single and multi-session versions.

21.2 X Window Terminals

X Window terminals were developed in order to provide a uniform user interface for different computer platforms. The need for such a window system, which can distribute information over the network independent of hardware or manufacturer, arose as early as the 1980s at **MIT** (**M**assachusetts **I**nstitute of **T**echnology) in Boston. Based on these requirements, the X Window system was developed, and was then distributed free of charge to third parties with version 10 (X10) in 1986. Many manufacturers seized the opportunity and supported the

current version **X11R6** as the graphics and window protocol for their products. However, the classical X Window terminals are being gradually replaced by Thin Client solutions. This solution expands PCs or NetPCs with X Window terminal functions by using corresponding software (Exceed from Hummingbird). Xware applications can be installed under Windows 3.1/95/98/ME and Windows NT/2000/XP.

The X Window system is used to enable easy communication between systems and applications running under different architectures and operating systems. Through the graphical user interface, it is possible to call up several applications on any X Window terminal or Thin Client with Xware software without having to exit running applications. In this way, information can be imported from one application to another (copy and paste function), which is particularly useful for office workplaces with changing applications. Through the X Window system, the network appears to the user as a single, large computer system.

The **X Window terminal concept** is based on the client-server model. In this case, the X Window terminal is referred to as a **server**, whereas the **client** is an X application, that e.g. runs on the processor of a UNIX host computer. The server, as the displaying part, is responsible for displaying and processing inputs and outputs of the screen, keyboard, and mouse. The actual application runs on the functional part, that is, the client. The client and server can run on separate systems, as long as these are interconnected via a network. It is important for server and client to use the same network protocol for communication; nowadays, this is normally TCP/IP.
Compared with workstations, X Window terminals represent a cost-effective alternative since system administration is centralised, which dramatically reduces the effort required for system management. However, potential network bottlenecks should be avoided when using X Window terminals, e.g. by implementing fast backbones. The performance of X Window terminals is calibrated through a standardised benchmark and specified in so-called XStones. The storage volume needed differs according to the number of open windows and graphics applications, therefore no general specifications can be made in this respect.
Some developments attempt to distinctly increase the performance and area of usage of X Window terminals. Given the enormous amount of business and technical software for Microsoft operating systems, efforts are underway to integrate these applications on X terminals. The X terminal is then able to simultaneously process and display **Windows applications** from a Windows server at the same time as the X applications from a Unix server. It is even possible to exchange information between X and Windows applications with the copy & paste function. For this purpose, there are a variety of software solutions on the market with very different functions and system requirements. The Xware software, mentioned above, represents a different approach by adding the X Window functions to PCs.

21.3 PCs and Workstations

As an alternative to display terminals, **PCs** with terminal emulation are often used today when the user wants to have access to applications on a host computer. The advantage is that smaller tasks, such as word processing, can be performed without overloading the central computer, and the PCs can be used for other tasks as well. The disadvantage of such a solution is the significant increase of the expense of system administration since part of the data is only stored locally, so updates must be carried out separately for each computer. The effort required for system maintenance increases substantially in relation to the number of PCs that can be connected to a network.
Another disadvantage of PCs is the greater user effort involved in learning to handle the central computer in addition to the PC. This disadvantage can be compensated by using **workstations** with the same CPU architecture or at least the same operating system, as the host computer. Examples of these are the SPARCserver and SPARCstations using Solaris. These workstations can ease the load on the host computer by providing the possibility of local processing. In contrast to the X Window terminals, workstations can provide local disk space. The advantage with respect to PCs arises from their uniform use and simplified system administration. On the other hand, the disadvantage still remains that the common personal applications, such as Office packets, are difficult to use.

21.4 Windows-Based Terminals (WBTs)

A Windows-Based Terminal (WBT) is a Thin Client hardware with Windows CE.NET as the operating system, which is linked to a Windows Terminal Server (for Windows 2000 or Windows NT 4.0) or a MetaFrame Server. A WBT uses either the **R**emote **D**esktop **P**rotocol (RDP, previously referred to as T.SHARE) developed by Microsoft, or the **I**ndependent **C**omputing **A**rchitecture protokoll (ICA) from Citrix for communicating with the server. Both protocols are generally supported.

An interesting characteristic of Windows-Based Terminals is that they are, as the name indicates, completely adjusted to the features in the Windows platform. Thus, a multi-user version of Windows 2000 or Windows NT is essential for operation, and they cannot co-operate with other systems. This interaction, e.g. with Java applications or special company software, has to take place with the help of a browser.

When using WBT, all applications are run on the server; even data backup is done centrally on the server. Data is never available locally. Therefore, WBTs provide you with security against data loss, as well as data theft.

Licenses are a very important part of this. **Every Client** first needs a Server Client Access License (CAL). In addition to this, terminal services CAL is needed on the client. Systems with Windows 2000 Professional and Windows XP Professional automatically have such a license and do not have to be additionally licensed.
To also work with non-Windows-based hardware, Citrix offers the **MetaFrame** product. The successor product to WinFrame extends the Windows Terminal Server both on the server and client sides. The ICA protocol also supports Java-based computers, MACs, UNIX computers and ICA WBTs, just to name a few.

ICA technology contains three important types of components:
The **server component**: This software component that is installed on the server works closely with the respective multi-user operating system. The following operating systems are supported: Windows 2000 server family, Windows NT 4.0 Terminal Server Edition, HP-UX, IBM AIX, Sun Solaris. The operating system on the server ensures that the application is run and the ICA server component ensures that the application logic is decoupled from the application user interface. The **protokoll component:** This component ensures that the decoupled application user interface is transported to the end device, transmits keyboard entries and mouse clicks from the user to the server and sends the answer from the server back to the end device. With the Citrix ICA technology, the applications only take up approximately one tenth of the network bandwidth that would normally be required in a client/server architecture. The **client component:** The ICA component (the ICA client) that is installed on the end device is used to display the application user interface and allows it to interact with the application that is run on the server. ICA clients are available for almost any type of end device. They allow you to work on all sorts of platforms that would otherwise have incompatible technologies.
ICA clients can be downloaded free-of-charge: http://www.citrix.com/download

21.5 Network Computers (NCs)

A **network computer** refers to a Thin Client system, that uses Java-based technology. Applications can be run on the server and also on the client as a Java applet.

They are designed for use in intranets and the Internet. They do not use the ICA and RDP protocols that are Windows platform-specific (see Windows Based Terminals), but the protocols that are common in the Internet. However, this does not mean that NCs are not suitable for ICA and RDP environments.

A client in this architecture no longer requires an upgraded PC, but simply the NC , which is particularly characterised by the following features:
- NCs need no hard disks
- NCs are ready to be used in the intranet
- High-performance CISC or RISC processors
- Extensive graphics and audio capability.

When compared to a PC, the NC offers significant advantages, which the user should be aware of before making a purchasing decision. This means among other things:
- Far lower administration costs
- Security and integrity of the network-oriented system is distinctly higher (e.g., in terms of central data backup)
- Maintenance of clients or client software is no longer necessary; far less PC help desk functions are required
- Central management of applications and thus, always uniform versions across the entire company
- The latest software versions can be up and running without time-consuming installation and maintenance since the relevant modules must only be loaded onto the server
- Web browsers as client software are available at economical prices
- Low training costs.

To summarize, these items can be reduced to one common denominator:
The general advantage of Thin Clients is the significant reduction of the Total Cost of Ownership (TCO).

A "Network Computer Reference Profile" developed by Apple, IBM, Netscape, Oracle and Sun defines exactly the minimum requirements an NC must meet. In the meantime, the profile has been adopted by **The Open Group (TOG).** This international organisation is also responsible for regularly updating the profile. Further information on TOG and its members is available at www.opengroup.org

NCs from manufacturers like NCD (NCD NC), Tektronix (Netstation), IBM (Network Station), Neoware (Workstation) or Sun (JavaStation) fully meet these requirements, however, they may offer different equipment and additional features.

21.6 Network PCs

In contrast to NCs, which are equipped primarily with Intel i960 and PowerPC processors (NCD, IBM, Neoware), or MicroSparc-II processors (Sun), the NetPCs (e.g. from Microsoft) will

be equipped with Intel Pentium CPUs. The main differences with respect to NCs or WBTs are that the operating system of the NetPCs and the typical hardware equipment (e.g., local drives and hard disks) remain very extensive. This results in a high initial purchase price for the client and higher cost of ownership. So the concept of central data management and administration has not yet been fully realised for this product area. Network PCs, rather, represent an additional workplace computer category positioned between PCs and Thin Clients.
As of yet, it is not possible to predict which manufacturers and which products will eventually prevail on the market. The only thing that is certain is that PCs as well as NCs and WBTs will have a niche to fill and more and more applications will be influenced by the Internet.
The aim when developing Internet appliances is to create independent and communicating units from ordinary units. The Internet appliances are networked together and data can be exchanged between the appliances. There have already been certain interesting partial results of this development. For example, household units have been networked together so that each unit can be controlled independently from the site of the user. This means that the entire home electronics can be controlled from nearly every Internet terminal. Therefore, the emphasis is no longer placed on the underlying technology, but instead the aim is for the computer to meet the needs of the user with respect to ease of application and operation.

22. Output Devices

22.1 Monitors

22.1.1 Units of Measurement and Terms

The **resolution** of a monitor refers to the number of pixels that are displayed on the monitor. One picture element, or pixel, represents the smallest unit that can be addressed.

The **refresh rate** specifies how often the entire picture is redrawn per second. The higher the rate, the more stable the picture appears. Optimum rates start at about 73 Hz beyond which the eye cannot detect a flicker: Thus limiting the eye strain - especially during relatively long sessions in front of the monitor. The recommended guideline value nowadays is 85 Hz.

The **line frequency** refers to the time required to draw a monitor line. It is calculated from the refresh rate and the number of lines for the particular video mode and is measured in kHz. It therefore depends on the resolution and the refresh rate of the monitor.

The **video bandwidth** reflects the quality of the video amplifier. If the video bandwidth is insufficient, shadowing may occur in the case of light vertical lines against a black background.

The **contrast:** The contrast control is used to adjust the brightness wanted. The contrast should be adjusted so that the picture is sufficiently bright, but is not of uncomfortable intensity, as is evidenced by the righthand image edges being distorted in colour.

The **focus** refers to the shape of the electron beam which, when it strikes the phosphor, should be as close as possible to perfectly round over the entire monitor area for all resolutions and frequencies. If the focus is poor, the displayed characters and lines become blurred, particularly in the screen corners. It should be borne in mind that the picture definition is also influenced by the video bandwidth, the convergence and the selected screen resolution.

The **picture stability:** Keeping the high voltage of a monitor stable under the wide variety of operating conditions demands high technical input. Poor stabilisation manifests itself particularly in bright/dark changes. A weak high voltage stabilisation then leads to constant changes in picture width, so-called "pumping".

The **smearing effect** is the phenomenon when a moving object on the screen no longer has any clear definition, but has a "tail" in its wake. This effect is encountered particularly in passive LCDs (STN, DSTN) and is conditioned by the slowness of the liquid crystals used. For

fast, moving pictures with high differences in contrast, the liquid crystals cannot assume the new condition with sufficient speed. In TFT displays and tube monitors similar effects are encountered to a significantly smaller extent. (For picture tubes a compromise always has to be sought between rapid response time and the suppression of flickering by sufficient after-glow of the phosphor.)

The **brightness control** adjusts the black level of the picture. (The contrast adjusts the brightness level). If the brightness level is set too high, black elements appear grey.

Convergence designates the alignment of the three electron beams for red, green and blue. Ideally these beams are directly one on top of the other in order that a clean, white dot can be produced. If this is not the case, lines and characters are displayed with coloured edges and lack definition.

Moiré refers to a wave-like pattern running over the screen, which changes when the picture width and height is altered and only happens with certain resolutions. These rippled images are caused by interferences between the video signal and the monitor mask and occur particularly with patterns of lines or dots. Ironically, the better the focus, the worse the moiré is likely to be. Patterns of lines or dots, extending over the entire screen, are particularly suited for testing purposes.

22.1.2 Types of Monitors

22.1.2.1 CRT Monitors

22.1.2.1.1 Principle of operation:

In **CRT** (**C**athode **R**ay **T**ube) monitors the image displayed on the monitor is produced when an electron beam hits a phosphor coating on the inside of the screen. To create the colours, three phosphor coatings for the primary colours red, green and blue are provided. The colour mixture is based on the so-called additive method, i.e., when a beam hits one of the phosphor coatings, light of the corresponding colour is emitted. Thus, by combining the three primary

colours, other colours of varying brightness can be created.
To ensure that the rays hit the intended phosphor dots, a mask is located in front of the phosphor coating.

At present there are three types of picture tubes on the market, which are differentiated by the kind of mask used:

22.1.2.1.2 Shadow mask, slotted mask and aperture grille

In the **shadow mask** the mask consists in principle of a perforated metal or ceramic sheet with round holes through which the electron beams are directed so as to strike the correct dot on the screen. The shortest distance between two dots of the same colour on the screen provides information about the fineness of the shadow mask and indirectly indicates the number of dots a monitor can display as single elements. The smaller the shadow mask, the greater the definition of the image.
In the **slotted mask**, however, the holes are rectangular. This type of mask is no longer prevalent in the monitor sector, this technology is used only by manufacturers of television sets.
The **aperture grille**, on the other hand, consist only of vertical wires under high tension, through which the electron beam is directed.

The advantages and disadvantages of the different mask types:
The **shadow and slotted mask picture tubes** are less sensitive to vibration due to the perforated sheet, but as yet cannot offer the contrast and vibrant colours of an aperture grille. In addition the enlarged surface means they are more subject to so-called doming, the effect resulting from intensified heating by the electron beam and leading to deformation of the mask. The electron beam is hence unable to traverse the shadow mask without problems, resulting in distorted colours especially in the corners of the screen, and poor distribution of brightness. Shadow mask monitors are suited particularly for use in industrial environments due to being insensitive to vibration.
By virtue of its design, the **aperture grille** converts less of the energy in the electron beams into heat, meaning it is less susceptible to colour distortion and can achieve higher contrast and more vibrant colours. The down side is that it is somewhat more sensitive to vibration and is stabilised by two damper wires perceptible to the trained eye as slight shadow lines.

Another advantage of a Trinitron monitor (Trinitron is a registered trademark of Sony Co.) is that it has a flatter screen surface. The basic form of conventional shadow mask systems is spherical and is difficult to flatten, even for the more modern FSTs (Flat Screen Tubes). In contrast, an aperture grille monitor is shaped like a cylinder. The screen lies like a window in this cylinder. Due to an enlarged radius of the cylinder, the screen surface of the Trinitron monitor is considerably flatter. The flatter screen surface prevents images from being distorted. Straight lines appear straight, thus making it suitable for modern CAD/CAM and graphics applications. Furthermore, the black tinted glass of a Black Trinitron monitor absorbs a great deal of the extraneous light such as sun rays and lamp light.

22.1.2.2 LCD Monitors

22.1.2.2.1 Principle of operation

This type of display makes use of the special properties of a group of chemical elements, the so-called **l**iquid **c**rystals, which are transparent and have twisted molecules. The twist of the molecules alters the polarisation of light passing through. When an electric field is applied, the crystals are aligned accordingly. These properties are used in LCDs to control the passage of light through the displays. The light is produced by light sources (backlight) at the rear of the display and is aligned by polarisation filters.

As long as no electric field is applied, the polarisation is changed along the twisted molecules. This light then impinges a second polarisation filter arranged at right angle to the first one and can pass through due to the twisting by the liquid crystals. If an electric field is applied, the twisting angle of the crystals and hence of the light is changed and, accordingly, only part of the light can pass through the display.

In this way the brightness can be controlled to obtain the necessary number of grey levels for a high-quality display. The screen is subdivided into pixels composing the picture as a whole. To obtain a colour display, three pixels per dot are used in the colours of red, green or blue produced by colour filters. The image is generated by a matrix of pixels.

There are two types of LC displays:

22.1.2.2.2 Active and passive matrix LCDs

The essential difference between the two available types of LCDs is the brightness control of each single pixel, known as addressing.

Contrary to the passive liquid crystal displays, the addressing in the **active matrix** TFT displays is active. This signifies that for every single dot, three electronic switches, one each for the colours red, green and blue, are used to control the cell. As a rule these switches are thin film transistors (TFT), which also protect the pixels against being affected by adjacent pixels and prevent crosstalk. The control by way of transistors has the advantage that the cells maintain their condition until they are re-addressed, allowing faster liquid crystals to be used and avoiding the smearing effect. These thin film transistors are produced in a process similar to the production of semiconductor chips. Due to the high reject rates, the production costs and hence the unit price are still rather high (in order to produce a display with a resolution of

1024 by 768, for example, 1024 x 768 x 3 = ~2.36 million transistors are required).
In consequence of the technology employed, the digital LCD displays are of course free from any errors such as geometric distortion, misalignment in convergence or lack of definition, because the arrangement of every dot consisting of three red, green and blue pixels is fixed. Instead, the following restrictions have to be expected, depending on the technology used:
- Missing or dead pixels
- Smearing with moving images
- Uneven illumination due to the backlight.

Passive matrix (STN, DSTN): In this type of liquid crystal display, the individual lines are addressed line by line, as in a monitor. The crystals are aligned by briefly applying an electric field and they then "lose" this alignment again. To reduce this effect, slow liquid crystals are used, which have the drawback of smearing effects in moving images as a result of the slowness of the crystals.

22.1.3 Monitor Requirements

22.1.3.1 Screen diagonal and application

The type of applications for which a monitor is used is important for determining the **screen diagonal** needed. There is a clear trend away from conventional tube monitors to flat panel displays, which are getting cheaper and better all the time. Today's applications, with their numerous menus and buttons, demand a minimum 1074 x 768 pixel resolution when using a TFT 15-inch (38.1 cm) monitor. However, when several applications or windows are to be opened at the same time, the user may reach the limits of a 15-inch flat panel display. 17-inch flat panel displays today having a standard resolution of 1280 x 1024 pixels are suitable in cases such as these. In addition to 18-inch flat panel displays, larger 19-inch and 20-inch displays are becoming ever more prevalent in a normal office environment for space-intensive applications. For the 20-inch displays a resolution of 1600 x 1200 dpi is already standard. Flat panel displays of this size, achieving a higher resolution through extrapolation, are by all means suitable for applications of short duration, however every extrapolation does involve loss in quality. It is therefore advisable to buy a monitor with such a resolution, which physically supports 1600 x 1200 dpi. **Inter-Extrapolation**: Method of enlarging or reducing images by adding or eliminating pixels. The quality of the interpolation method is decisive for the resultant picture quality.

It is advisable to go for a CRT monitor with aperture grille instead of flat panel displays particularly in graphics-oriented applications, in CAD/CAM and for applications in which

absolutely true colour and brilliance are called for.

In an industrial environment where shocks and vibrations cannot be avoided, CRT monitors with shadow masks should be used. (Cf. advantages and drawbacks of the various mask types). For CAD, graphic design and DTP applications, 19-inch (48/50 cm), 20-inch (50.8 cm) or 21-inch (53/55 cm) monitors with a resolution of 1280 x 1024 pixels are appropriate. For more complex applications in this area there should be 1600 x 1200 pixels.

22.1.3.2 Ergonomics

Ergonomic requirements for monitors:
Minimum requirement: MPR II (tested at TÜV Rheinland), environmental Blue Angel label and TCO seals.
Decisive factors are: Size and picture quality, flickerfree, sufficient contrast, good definition, dynamic focus, adjustable convergence control, sufficient character size, free from interference and reflection,tilt-and-swivel facility.
Workplace set-up: The right workplace set-up also plays an important part in permitting relaxed work and preserving health. Also instrumental are the seating adjustment, working posture and viewing conditions. Further information on this is available from transtec.

22.1.3.3 Radiation

Another important factor when selecting a monitor is the intensity of radiation it emits. Guidelines have therefore been defined by different institutes, with regards to radiation levels. An important standard is the **MPR-II**, a recommendation concerning the radiation levels with respect to alternating electromagnetic fields, alternating electric fields and electrostatic charging, which was specified by the Swedish inspection institute in 1990. Monitors with radiation below the values stated by this recommendation are usually referred to as low emission.

The current guideline is **TCO99**, based on the predecessors TCO92 and TCO95, has been in force since January 1999. To be TCO99 certified, monitors must meet the new stringent guidelines and requirements in terms of quality and ergonomics. For the first time, a minimum refresh rate has been determined depending on the resolution and size of the device (e.g. a 17-inch monitor with 1024 x 768 resolution, min. 85 Hz). Further innovations include: a lower power consumption of 15 watts for the energy saving mode (previously 30 watts), a ban on substances containing bromine and chlorine, the avoidance of heavy metals and mandatory recycling regulations in the relevant countries.

TCO-03, accepted at the end of 2002/beginning of 2003, is new and still almost unheard of. It is based on the TCO99 standard and makes more stringent demands on the two areas of tube monitors and flat screen displays, defining, for example, vertical adjustability, cable, housing and recycling requirements.

22.1.4 Types of Monitors

In the PC area, different graphics standards have been developed over time to ensure that graphics cards, monitors and application programmes are all compatible with one another. Today, the most important ones are the VGA standard (Video Graphics Array) and the Super **VGA** standards. The resolution of the VGA standard is 640 x 480 pixels; 16 out of a palette of 256 colours can be displayed at the same time. In an additional mode 256 colours can be displayed, but only with 320 x 200 pixels.

Super-VGA (**SVGA**) and **XGA** stand for higher resolutions of up to 1024 x 768 pixels. In order to display 16 out of 256 colours, a graphics card memory of 512 KB is required. With a 1 MB memory expansion, a total of 256 colours can be displayed from a range of 16.7 million colours.
For workstations, the graphics standards of the PC area are of little importance. The graphics cards available for the Compaq/Digital, HP, IBM and SGI environments allow resolutions from 1024 x 768 and 1280 x 1024 up to a maximum of 1600 x 1280 pixels at different refresh rates. The graphics cards used with SPARC Stations usually support a resolution of 1152 x 864 pixels with a refresh rate of 66 or 76 Hz. Graphics cards that support the VGA standard or reach resolutions from 1024 x 768 up to 1600 x 1280 are also available in the SPARC area.

22.2 LCD Projectors

For presentations to larger groups, the screen display may need to be shown to all participants with the use of a projector. While presentations using an overhead projector are time-consuming to prepare (printing transparencies, making copies, etc.), the use of LCD projectors saves a significant amount of time and changes can be made right up to the last minute, even while the presentation is underway. What is more, animation can be added, making the presentations interactive and giving extra interest.
The development of the present-day generation of projectors is going in the direction of handier and quieter devices for use in conference rooms, for mobile use and at events on all scales. Sound levels of less than 30 dB, ultramodern lamps with a service life of up to 6000 hours and a weight of less than 2 kg are already common. The constantly increasing ANSI Lumen figures permit use in rooms of almost daylight conditions. The projectors from Philips,

for example, operate with the so-called Limesco technology (**Li**ne **Me**mory **S**can **Co**nverter), including a chip permitting screen-filling display of higher graphics standards through the projector's own graphics resolution.

22.2.1 The technologies

The most widespread technology is the **LCD technology**. Light is split into the primary colours of red, green and blue, is conducted through one LC display each as a light valve and at the end of this process is united again, hence each individual dot and each one of the three primary colours has an LC display of its own.
The **DLP technology** is based on multiple, miniscule, square-shaped mirrors applied to a DMD chip. The light is reflected to the screen only via tilted mirrors. The human eye cannot detect this high frequency in which the mirrors are switched on and off, permitting the display of gradations. The colours are generated by a colour wheel filtering the light into the three primary colours.

22.2.2 Advantages and disadvantages of the two technologies

A high degree of saturation of the colours and uniform illumination are the advantages of the dependable LCD technology. Devices incorporating LCD technology also have the advantage of economical purchasing costs.
True colours and the absence of any pixel errors are clear advantages of the DLP devices which, while being slightly more expensive, outshine alternatives by bringing more light to screen and providing very good picture quality.

22.3 Printers

When talking about different printer technologies, inkjet printers and laser printers take the forefront, while matrix printers and thermo printers are being pushed into the background.

22.3.1 The Technologies

22.3.1.1 Inkjet printers

Various technologies have been developed for inkjet printers. Distinction is made between solid ink printers and liquid ink printers, as well as between drop on demand technology and continuous flow technology.

The **solid ink printers** are used for colour printing. Ink, which is gel-like at room temperature,

is heated and in the melted condition is squirted onto the paper where it immediately cools again. Due to the change in condition of the ink (solid - liquid - solid), these printers are also called phase-change inkjet printers. The print image is very good, but the purchase price and maintenance of these printers are relatively expensive. It is the technology which makes them so costly. A separate melting chamber is provided for each colour in which it liquefies and is kept ready in an ink reservoir for printing. The ink cools again after the printer is switched off, entailing time-consuming cleaning of the printing system. Long waiting times are incurred when the printer is switched on again until the operating temperature has been reached and the respective ink is ready in a melted state. Similarly as with the thermo transfer method, the suitability of the printed results for daily use is limited. They are delicate and therefore have to be handled very carefully.

Continous flow inkjet printers are becoming ever rarer on the market. Only a few printers operate using this technology in which ink flows continuously in an electrostatic field from a reservoir into a collection vessel. During printing, individual droplets are deflected from this ink stream onto the paper. This technology achieves very high quality, especially in colour printing, but it is very expensive and the equipment is not compact enough to meet normal printing requirements. They are special answers to specific needs.

Thermal inkjet printers

This inkjet technology from HP has established itself as the mainstream technology. Printers operating according to this principle are fast, provide high print quality and the purchasing and maintenance costs are low. These printers embody a very robust system in which the printer mechanism and the high-precision print heads are kept as separate units. The entire print head is accommodated in replaceable ink cartridges produced using the same production technology as for microchips. The micro-lithographic production technology provides for economical mass production attended by constant quality. A round, mono-crystalline silicon wafer is the starting material which is etched to provide the structures necessary for the ink transport, ink nozzles and, of course, for the electrical connections and components. When an ink cartridge is empty, the new cartridge automatically comes with a new print head, ensuring years of high print quality at a comparatively low price.

In this printing method, the print head travels on a carriage horizontally over the paper. A strip of the image is printed, then the paper moves on ready for the next strip. The printing operation is of non-contact type. Wherever a dot needs printing, a tiny ink droplet is fired from one of the nozzles in the print head. That is why the technology is also known as drop on demand, as opposed to the continuous flow technology. It ensures that only as much ink is used as is actually required for printing. In the meantime there are highly complex operations going on inside.

The nozzle firing the ink is a miniscule chamber with a resistor at the bottom, which becomes

hot when power is supplied. The ink exits the reservoir through capillary force and passes through channels into the nozzle chamber where it is heated to temperatures in excess of 300°C in fractions of a second. The vapour bubble forming at the bottom of the chamber presses the remaining ink in the chamber through the nozzle shaft and out of the opening. At the nozzle opening, which is thinner than half a human hair, the vapour bubble collapses liberating shock waves shooting out the ink droplets at a speed of some 100 km/h. Once the ink droplet has left the nozzle, the resulting vacuum in the nozzle chamber draws ink from the reservoir to replace the ink that was ejected. The process can repeat itself several thousands of times a second. Nowadays sophisticated computer simulation is used for analysing the droplet formation and ejection, enabling on-going further development of the operation. This is important, for instance, with a view to the dynamics of the process, since the print head moves on continuously as printing takes place. The droplet does not fly along a linear path, but describes an accurately calculated, arcuate path towards the paper. These technical parameters are incorporated in the printer control and the entire operation is optimised accordingly.
The developers' attention is also levelled to the size of the ink droplets. Although already in the range of 10 picolitres, the droplets are getting ever smaller. This allows the ink volume to be dosed better by printing several droplets on top of one another. In practice, the following benefits arise: More colours can be represented per dot, the print quality is very high even on normal paper and the print speed is increased.

22.3.1.2 Laser Printers

Whereas the above-stated printer technologies are based on the so-called line principle - a print file generated in the computer - laser technology is based on the page principle. Laser printers use an electro-photographic process, a physical printing process in which no chemical processes take place and hence no liquid solvents are required. The first step in the printing process is that the transfer drum is electrically charged. Then a laser beam guided by a system of mirrors transfers the image line-by-line to the drum. For every line dot that is not to be set, the printer switches the laser beam off. At the impact point of the laser the electric charge on the transfer drum is then neutralised. In the next step of the printing process, the drum takes
up the toner on these neutralised points. The toner is then transferred onto the paper and fixed by applying heat and high pressure.
This method applies both for monochrome and colour laser printers. The only difference in the available colour laser printers is that the complete operation is carried out once for every mix colour (C, Y, M and K). Rather than moving the paper through the machine four times, the image to be printed is accumulated on a special belt. The charged drum then transfers the image in one step onto paper or transparencies.

Several manufacturers have developed methods of improving print quality through controllable dot size or variable dot positioning. HP, for instance, refers to this method as **R**esolution **E**nhancement **T**echnology (**RET**). This built-in feature can produce rounder curves, sharper lines and edges and smoother transitions by optimising the size and position of dots in relation to the adjoining points, irrespective of software.
The benefits offered by laser printers are their high print quality and graphics capability as well as the low noise level during printing. Since the purchase price of laser printers is higher than that of inkjet printers, they are recommended primarily for high print volumes.

22.3.1.3 Matrix Printers

A matrix printer was to be found in the vicinity of almost every PC even a few years ago. For the first time they allowed text and graphics to be printed on one and the same page. Apart from the fact that the print image they produce composed of individual dots is not particularly attractive, matrix printers are comparatively noisy. Some matrix printers can also produce coloured prints using a three or four-colour ribbon instead of the black one. Apart from the extremely slow operating speed, the running costs increase because, during printing, dark colours often get onto the Y strip of the ribbon and smear it. All the same, matrix printers are today indispensable in some areas, particularly in logistics firms and other areas in which multi-page forms (carbon copies) have to be printed.

22.3.1.4 Thermo Printers

We will look at two technologies subsumed under the term thermo printers: Thermo transfer and thermo sublimation printers. The few thermo printers which render dyes incorporated in the paper visible by heat (similarly to thermo fax machines) are no longer of any great significance on the market.
For high-quality colour prints, **thermo sublimation printers** are offered. They heat coloured wax situated on a carrier film and print it on paper. The wax becomes so hot that it vapourises and penetrates the paper. Thermo sublimation printers also permit fine colour gradations to be printed in a quality sufficient for proof prints in graphic art. It is of interest in this connection that these printers achieve their high-quality image output with a comparatively low resolution of even less than 200 dpi.
This is further proof that colour print quality is defined not by the resolution but by the way in which colour is mixed. The purchase costs and operating costs (colour films, special paper) are so high, that thermo sublimation printers do not come into question for the office and home computing sector. The quality of the economical versions offered especially for the home computing sector falls so short of that provided by the professional models that their benefit is doubtful.

The **thermo transfer printer** is a simplified type. It also operates with the use of wax films, but the wax is heated only to the point at which it melts and is applied to the paper as a coloured dot The colour rendition is sufficiently good but the prints are susceptible to creases and scores. Only special paper and films can be used and the wax films needed are expensive. Even though some printers are offered in the lower price segment, the high costs of the consumables is a factor discouraging their general use. They are used particularly by professional commercial artists, for whom they represent a genuine alternative, particularly an economical one, to thermo sublimation printing.

22.3.2 Printer Protocols

Page description language is often used as a synonym for printer protocol.
Most laser printers manage the **HP PCL** and/or **PostScript** protocols. PostScript is a printer-independent page description language that uses mathematical curves (Bezier curves) to define a font. PostScript is used in many sectors due to its high performance and flexible command structure.
HP PCL (**P**rinter **C**ommunication **L**anguage) is a page description language developed by HP for the control of printers, especially plotters, which is already used by numerous other manufacturers. HP PCL 6 is a further development of HP PCL 5 which is downward compatible to the earlier version. The innovations in PCL 6 include accelerated printing, especially of complex graphics, improved print quality and higher network throughput resulting from shorter and more compact commands. Significant increases in performance are to be expected particularly under Microsoft Windows, because the PCL 6 commands support the Graphical Direct Interface (GDI).
Some units also manage the HP-GL (HP **G**raphics **L**anguage) plotter language. The HP-GL/2 procotol integrated in HP PCL is a further development of HP-GL. HP-GL is a vector-based language, which ensures top print quality.
Implementing the HP-GL or HP-GL/2 plotter protocol on a laser printer enables plotter commands to be processed by the laser printer and to be transformed into the corresponding graphics output. Thus, exchanging a plotter for a laser printer supporting HP-GL- or HP-GL/2 is a simple task that can be performed without having to adjust drivers. Graphics output which, for technical reasons, is a time-consuming process with a plotter can be performed within seconds with a laser printer.

23. Multimedia

Multimedia is no longer just a widespread catchword, it has actually been implemented in many applications.

In general, multimedia is a term used to refer to the **combination of various media, such as still pictures (e.g. those produced by digital cameras), audio, motion pictures, texts and graphics**. Many computers perform additional tasks. The user can listen to conventional audio CDs, send e-mail messages with embedded sounds or have the computer serve as a telephone system to forward or answer calls.

Video conference software is becoming increasingly important, since it can do a lot more than just transmit normal video pictures. Conferences can be held between various connected partners and documents can be exchanged rapidly and cost-effectively over great distances. In the context of multimedia, the CD, DVD and jukebox media are mentioned as well. This is because the data quantities are usually very large, requiring fast and cost-effective high-capacity storage media.

23.1 Digital Cameras

The number of multimedia applications continues to grow and therefore digital cameras become more important. In order to process pictures from a conventional camera by computer, the film must be developed first and the slide or print must be scanned. Digital cameras produce a digitised image stored in an internal memory which can be transferred to the computer via an interface, a floppy disk, or a CD/DVD.
The user can enlarge or reduce the pictures, incorporate them into documents, and send them using online services. Desktop publishing, presentations and relatively simple marketing brochures can be illustrated quickly and easily without the need for a scanner. The cycle **"photograph - view - edit - print"** can be completed if a suitable colour printer, e.g. from HP is also available.

23.2 Voice Applications

Like music, voice is an analogue signal which can be converted by digitisation into a computer processable format, and subsequently reconverted into an analogue signal using a sound card. In order to input, output and process sound information, the computer needs an audio interface card, called a sound card in the computer industry.
The audio card samples several thousand times per second the analogue audio signal input from a microphone or patch cable (sampling). The sampling quality depends on the resolution

of the digital-analogue converter (e.g. 8, 10 or 16 bit) as well as on the sampling frequency (number of readings per second). This process is called **P**ulse **C**ode **M**odulation (**PCM**). To transfer voice signals over long distances, for example via e-mail, the quantity of data must be kept small. For that purpose, the signals are compressed before transmission using **ADPCM** (**A**daptive **D**ifferential **PCM**) and decompressed again by the recipient.

23.3 Video Processing

Video cards are applied in a wide range of applications today. Apart from their use for creating digital video clips for training purposes or business presentations, the production of live videos or video-on-demand is arousing interest. **Live video** allows the user to follow news broadcasts and run applications at the same time.

Another field of application is the real-time transmission of compressed video images via networks. The images can be stored on a video server equipped with an appropriate card. By means of **video-on demand**, remote users with network access to the video server can watch the stored video films on their workstations or save them locally for further processing.
To create **presentations**, images can be read into the workstation from a suitably connected video recorder or video camera. Thus, either entire video sequences or freeze frames can be manipulated, stored or incorporated into computer demonstrations as desired.
However, a video card is needed before video clips can be created. A video card stores video sequences or freeze frames from a connected video recorder, camera or television set on the hard disk. Video software then offers the possibility of manipulating these images. Since the video data can be reproduced and edited without noticeable loss of quality, video editing is reduced to file copy and delete commands. Furthermore, the computer processing of video images offers the advantage that each individual section of images from the digitised material can be accessed within seconds, eliminating the time and effort involved in rewinding.
A variety of video cards with different performance features, described below, are available for transferring individual video images or entire video sequences to the PC or workstation.
In order to display real-time videos or television pictures on the monitor without the complexity of digitising the video data, an **overlay card** is needed. Overlay cards usually only have a limited, or no grabbing function at all, and therefore can only be used as television receivers or for surveillance cameras. The video signal is transferred from the overlay card to the graphics card so that it can be displayed without wasting computing time. The video image is displayed in a separate window on the monitor.

In addition to the video and overlay functions, some video cards incorporate a **graphics card**. This means that a separate graphics card is not required and only one slot in the PC rack is needed for both functions. Video cards with a **TV tuner** go one step further. In addition

to live videos, they allow any live TV broadcast to be displayed in a window on the screen.

To display and digitise entire video sequences on-screen, **real-time digitisers** are required. They convert the video signal into digital picture sequences.
The main problem in storing video sequences is the enormous amount of data generated by the digitisation of pictures and videos. Digitising one video image with a 24-bit colour resolution requires about 1.2 MB. Thus a PAL signal of 25 images per second generates a data volume of 30 MB per second. Therefore, to store video sequences, a video card with **hardware compression** should be used. Data compression reduces the memory space required, and at the same time increases the speed at which the video images can be processed. Besides, such compression reduces the bandwidth required for data transmission over the network.

Compression without degrading quality and without taking too much computing time is possible only up to a ratio of 1:2. Practically all solutions for video compression must however have far higher compression ratios. Software solutions use algorithms that have compression ratios of up to 1:20, with only a slight degradation of image quality. Given the limitation of this compression ratio, many video digitisers use the **JPEG** compression standard. Special integrated circuits on the video card allow images to be compressed and decompressed by factors of up to 1:10 in real time, i.e., at 25 frames per second or more. However, when compressing using this algorithm, there is a risk of losing some information, especially at higher compression ratios, i.e., not only is there a difference between the original image and the compressed one, but also the former cannot be reconstructed from the latter. So-called cube effects occur and the components of the image become extremely fuzzy. The degree of compression can be adjusted. In practise, a compromise has to be made here, since high quality means little compression and therefore large quantities of data.
Motion JPEG only works with single images; in effect, a file is created containing nothing more than JPEG-compressed single images.
MPEG goes one step further in that it takes into account that there is usually only a small difference from one image to the next. This method eliminates the similarities between consecutive images. As few as two images per second contain the full independent information of a single image. All the other intermediate images contain only the changes to the preceding image. It requires an enormous amount of computing power to determine which parts of the image can be used in the next image. This accounts for the high costs of video digitisation cards with MPEG motion compression.
To ensure an appropriate **output quality** of the digitised video sequences, it is recommended to use only high-quality starting material. Problems with image quality soon become apparent when digitising VHS videos. S-VHS or, even better, professional Betacam are much more suitable for this task. The fact that image quality is lost due to the compression does not

mean that the input quality is unimportant. Faults such as drop outs or noise become doubly apparent after digitising. Compression methods such as MPEG are very sensitive, particularly to signal noise, due to the fact that a great deal of information changes from one image to the next.
Once the video images are stored on the hard disk, they can be edited and manipulated as desired using the video-processing software. Thus, by means of dissolves and other special effects it is possible to create ones own video films.
With suitable libraries and supplied programme interfaces, users are able to write their own **programmes** for image processing on the workstation or PC. Thus, for training and business presentation purposes, for example, the desired image sequences can be arranged and combined with text inserts.
Tutorial programmes can be designed so that the individual user can choose which chapter he wants to display on screen and with which degree of difficulty. The programme then searches for the corresponding pictures.
For product presentations in particular, it is obviously suitable to output the previously digitised video sequences onto a video recorder or television after processing. For this, the video card must have a separate **video output**. The advantage is that a television set is considerably cheaper than a large screen display monitor with an appropriate diagonal size. Furthermore, the videos can also be presented to a larger audience by means of a video projector. The production of video tapes is another interesting option, especially since the data stored on the tape can then be displayed on a TV set by any video recorder.

23.4 Video Conference Systems

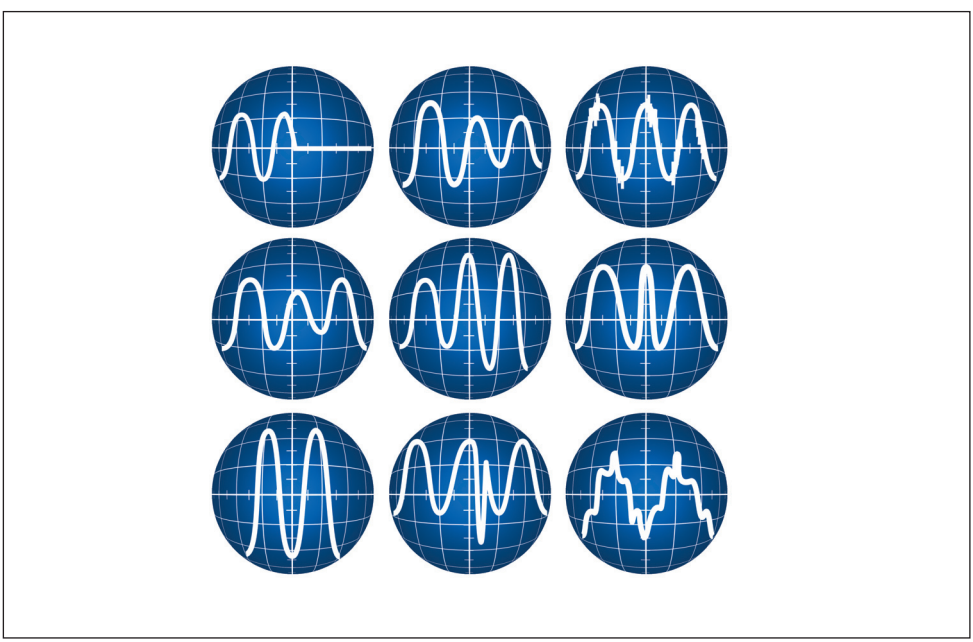
Costs savings is the main reason why companies value video conference systems more and more. Audio/video communication leads to easier and clearer understanding between individual or multiple communicating partners. The extensive spread of ISDN has facilitated the implementation of video conference systems. Most systems available today, in addition to the required software, include an appropriate graphics card, a video camera, a microphone and an ISDN card. With ITU industry standards H.323 (LAN), H.320 (ISDN) and T.120 (common data access), it is possible to communicate with video conference systems from other manufacturers who meet these standards.

24. Uninterruptible Power Supply

Uninterruptible **P**ower **S**upplies provide all electric devices with power from a battery, if the mains power fails. An unexpected network disruption can lead to the loss of data and programmes, and the resulting costs and damage cannot even be calculated. The contents of a disk can become damaged if the power failures occur during a write operation or if the contents of the write-cache memory was not previously transferred to the disk. This is why it is important to shut down the computer in a controlled manner when such disruptions occur. However, power failures are not the only reason for network disruptions, there are a variety of causes for this and and they can occur in many different ways.

24.1 Problems, Causes and Effects

Problem	Cause	Effect
power failure	storms, transformer failures, generator failures, construction work	mains-powered devices fail, loss of data
power disruption	storms, short circuits, power utility switching operations	defective connections, system crashes, loss of data
undervoltage, undervoltage, voltage variations	mains overloading, motor operation, device disconnection, motor cut-out	hardware and software failures, system crashes, undiagnosable fault symptoms
frequency variations	diesel generators	hardware and software failures, system crashes
voltage distortion	harmonics, clocked power supplies, short circuit, mains feedback	hardware and software failures, instabilities (symptoms only partially diagnosable)
switching peaks, transients	lightning protection diodes, DC-AC converters, CCC transformers	hardware and software failures, instabilities (symptoms only partially diagnosable)



24.2 Technologies

24.2.1 Offline Technology

Devices are directly supplied by the network via a filter, no voltage or frequency regulation. The devices are supplied by the battery through the inverter after a switch time of < 10 ms. Offline models function like an emergency generator and are only activated when the power has already failed. The switch times listed are generally sufficient for PCs, this time is usually not long enough to cause disruption. In contrast, even this short switch time can lead to data loss in other systems.

24.2.2 Line Interactive Technology

Just as with offline technology, devices are directly supplied by the network via a filter when using line interactive technology, but voltage regulation is possible. A control circuit connected in parallel to the power supply of the UPS regulates fluctuations of the main power within a range which is tolerable for the user and allows full utilisation of line resources because, in the event of a power interruption, the UPS does not immediately switch to battery power. The result is a longer battery life. If the power fails, the devices are supplied by the battery through the inverter after a switch time of < 10 ms. This improved version of offline technology, however, is not sufficient if the connected devices are sensitive to the phase shift caused by the switch from line to battery power. This is the case with many telecommunications systems.

24.2.3 Online Technology

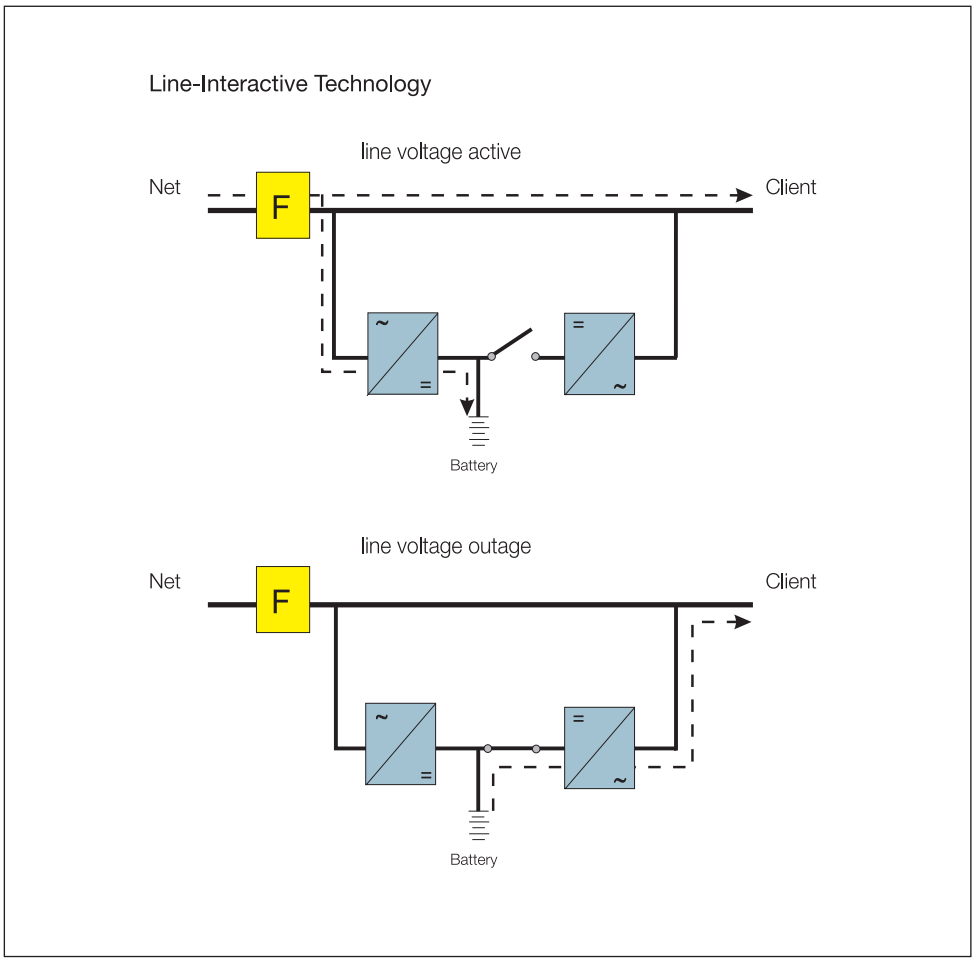
This model provides more data security, but it is generally more expensive than offline or line interactive models, due to its complex technology. The incoming mains power is electrically isolated and rectified. The rectified voltage is smoothed, stabilised and transformed back into a clean alternating current. This technology guarantees that the connected devices will be free of voltage fluctuations, disruptions, noise, or spikes.
The battery is charged during normal operation. Should the power fail completely, the charged battery takes over the power supply to the inverter smoothly and without interruption. This technology continuously produces a completely new voltage signal, independent of that of the mains. So online UPS systems do not operate without activation times, but offer the additional advantage of functioning as voltage stabilisers and line filters. A disadvantage of online UPS is larger energy expenditure due to the continuous supply through the inverter. The high amount of heat that is generated by this necessitates fans, which also lead to more noise. Therefore the optimum area to install this system is in a room that is exclusively used for computers.

Common models also have a bypass function, so that the devices are not just independent from the network, but also from the UPS. In the case that the inverter is defective, the system

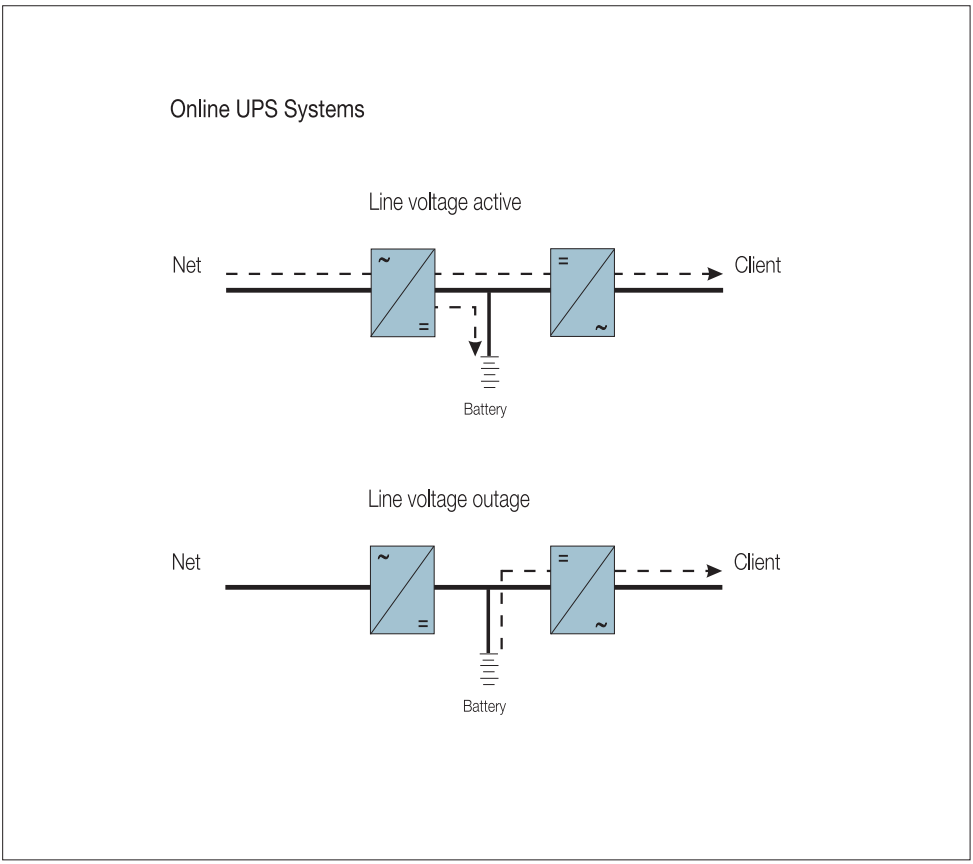
is bridged and the devices are immediately supplied with energy from the network. Additionally, the systems have a manual bypass function, a simple flip of a switch allows you to exchange the batteries and perform maintenance work while the system is running without influencing the devices.

Features	Offline (or standby)	Line-interactive (or delta conversion)	Online (or dual conversion)
Bridging time	a few minutes	a few hours	several hours
Voltage regulation	only battery-operated	limited	unlimited
Filtering	limited	limited	unlimited
Frequency control	no	no	yes
Efficiency	> 98%	> 98%	92% to 97%
Noise level	none	none or very quiet	fan-cooled
Typical applications		network installations	network installations, critical environments

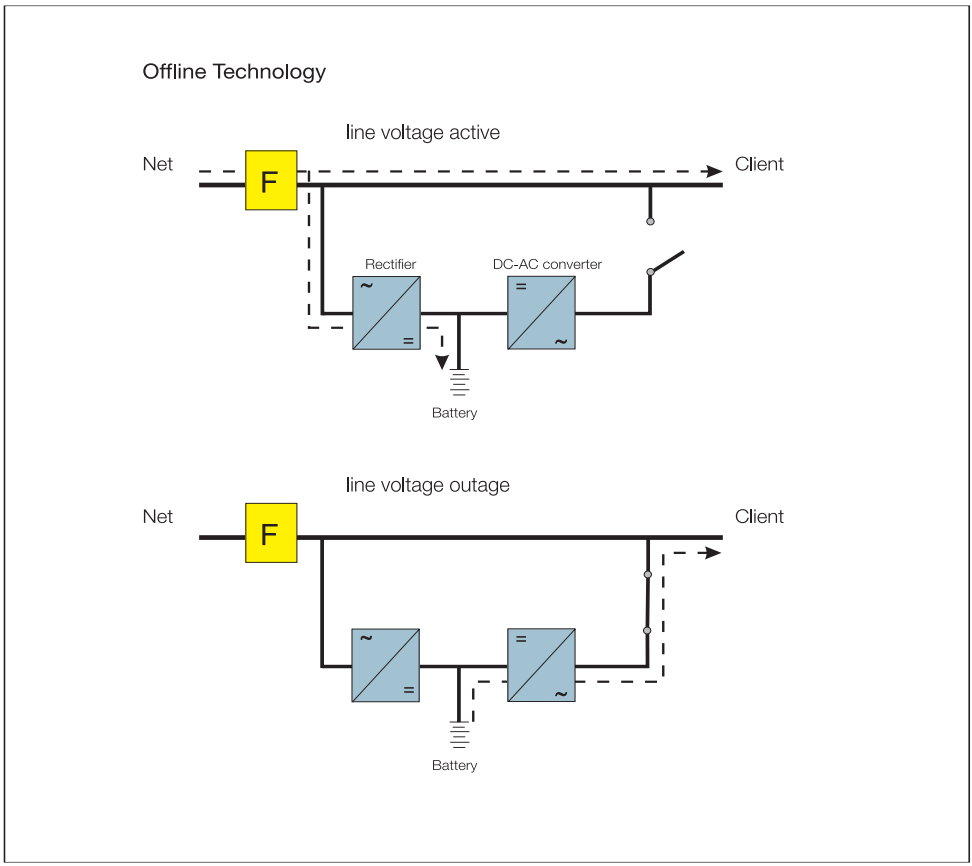
A comparative overview:



Line-Interactive Technology



Online Technology



Offline Technology

24.3 Dimensioning

Proper dimensioning is based either on the rated output in volt amperes (VA) specified on the nameplates, or by calculating the apparent power for non-linear loads. If the rating in VA is indicated on the label on the back of the system, the ratings of the individual devices can simply be added together. These include, among others, monitors, terminals, external data storage and peripherals. If the rating is given in watts, the apparent power must be calculated by multiplying the rating in watt by a constant. For typical switched-mode power supplies, a factor of 1.4 may be used.

Calculation example: Rating on nameplate of 150 W
The corresponding value in VA is given by 150 * 1.4 = 210

If system expansions are planned or foreseeable, the additional components should be included when calculating the capacity of the UPS. It is generally recommended to buy a unit with 20% more power than currently needed. If the power consumption of the connected units increases due to a later extension and thus exceeds the power performance of the UPS, an efficient shutdown of the system can no longer be ensured. The UPS systems contain a self-protection mechanism that simply shuts the system down if it is overloaded, before it can be damaged.

We recommend UPS function testing on a regular basis, including all of the connected devices.

Since a UPS safeguards a system temporarily, the connected computers and peripheral devices can only be supplied for a short time with battery power. This **stored energy time** can be selected and adjusted via battery expansion sets so that data and programmes can always be properly saved in the event of a power failure. Systems are available with bridging times that span several hours for use with critical applications that have to run constantly.

The batteries are an essential part of every UPS. Their average service life is between 3 and 5 years, but this is greatly influenced by the battery temperature. Tests carried out by manufacturers have shown that increasing the ambient temperature by 5 degrees Celsius reduces the battery lifetime by up to 25%.

24.4 Alarm, Interfaces, Network-Wide Communication

It should be possible to sound alarms about critical states such as a power outage or a failure of the UPS itself via network, email, pager, fax or GSM mobile telephones. If a message comes from the UPS, pre-defined instructions or scripts (*.exe, *.bat, *.com, *.cmd, etc.) should be guaranteed to execute automatically. A variety of communication options make direct access (with user-dependent restrictions) possible even over the Internet.
It is possible to manage a UPS system from every workstation in the network. If several UPSs are installed in a network, they can be managed from every authorised workstation throughout the network.

Messages that have been triggered by the UPS management software can be sent at a specific time to pre-defined network users. These messages can be edited and alterations, which are made for every workstation, are activated immediately. This means that the shutdown software or even the computer does not have to be rebooted. Functions such as changing

settings, rebooting the network etc. are protected by a password.
If several network components (servers, workstations, gateways, hubs, routers, etc.) of the same network operating system are connected to a UPS, then a group can be created. This means that one computer is connected serially to the UPS and functions as a group controller. This makes the UPS data available on the network. The shutdown software is installed on all other network components; these components are defined as group members. Their behaviour, however, can be specified independent from the controller. This ensures, for example, a sequential shutdown of all of the network components.
It should also be possible for a user at a workstation to initialise the test of all of the UPS systems in the network.

All incidents are saved in a log file. Keeping a separate battery log file is absolutely necessary, due to the importance of the battery as the energy storage medium in the UPS system. This file will make it possible to correctly judge the status of the battery.

UPS devices are generally equipped with a RS232 interface and usually also have a USB interface, which provides the connection to the computer. All USPs usually come with a management and shutdown software for all the common operating systems. Updates can be downloaded from the homepage of the relevant manufacturer free of charge. Thanks to the user guide, the UPS is easy to install.

24.5 Shutdown

A shutdown triggered by the UPS management software includes the following functions:
1. Storing the cache memory contents on the hard disk
2. Storing the current state of all open files (Windows)
3. Correct termination of all applications
4. Executing any entered commands (e.g. do a backup, save database etc.)
5. Performing a proper system closure and
6. Switching off the UPS

It is also possible to define an exact plan to switch the UPS on or off.