CHAPTER 7

Computer Networking

Lajos Bálint
HUNGARNET-NIIFI

Tamás Máray
Technical University of Budapest

1. INTRODUCTION

Computer networks serve today as crucial components of the background infrastructure for virtually all kinds of human activities. This is the case with Industrial Engineering and any related activities.

The importance of these computer networks stems from their role in communication and information exchange between humans (as actors in any kind of activities) and/or organizations of these human actors.

In contrast to the situation several decades ago, a number of key attributes of the working environment today serve as basic new supportive elements of working conditions. Some of the most important such elements are:

- Computer technology has moved to the desktop, making possible direct access to advanced techniques of high-speed information processing and high-volume information storage from individual workstations.
- Interconnection of computers has been developing at an exponentially increasing rate, and today high-speed global connectivity is a reality, providing fast and reliable accessibility of high-volume distant information at any connected site worldwide.
- Global addressing schemes, structured information content, and multimedia information handling have become an efficiently functioning framework for computer-to-computer (and thus human-to-human) communication and information exchange, as well as for storing and delivering large amounts of highly sophisticated information at millions of sites, accessible by any other authorized sites anywhere within the global network of networks, the Internet.

Computer networks and services themselves, as well as related applications, are briefly investigated in this chapter. The topic is very broad and there are thousands of sources in books and other literature providing detailed information about all the aspects of computer networking introduced here. This chapter therefore provides a condensed overview of selected key subtopics. However, emphasis is given to every important element of what the network infrastructure looks like, what the main characteristics of the services and applications are, and how the information content is being built and exploited. Some basic principles and methods of computer networking are repeatedly explained, using different approaches, in order to make these important issues clear from different aspects.

This chapter starts with an introductory description of the role of computer networking in information transfer. Then, after a short historical overview, the networking infrastructure is investigated in some detail, including the basic technical elements of how networking operates in practice. Later sections deal with services and applications based on networked computer systems. Separate subsections are devoted to the World Wide Web, which is probably the most important tool at the start of the third millennium for communication, information access, and collaboration. Some questions of content generation and provision are then investigated, after which aspects of future development in computer networking are briefly dealt with. Other subsections deal with a few practical issues, including some related to industrial engineering.

The following major issues are dealt with in the following references at the end of the chapter:

- *Networking in general:* Derfler and Freed (1998); Hallberg (1999); Keshav (1997); Kosiur (1998); Lynch and Rose (1993); Marcus (1999); Taylor (1999)
- *Internet in general:* Minoli and Schmidt (1999); Quercia (1997); Smythe (1995); Young (1999)
- *History of networking:* Salus (1995)
- *Intranets:* Ambegaonkar (1999); Bernard (1997); Hills (1996); Minoli (1996)
- *Extranets:* Baker (1997); Bort and Felix (1997)
- *World Wide Web:* Stout (1999)

- *Multimedia:* Agnew and Kellerman (1996); Buford (1994); Wesel (1998)
- *Virtual reality:* Singhal and Zyda (1999)
- *Quality of service:* Croll and Packman (1999)
- *Security:* Cheswick and Bellovin (1994); Goncalves (1999); Smith (1997)
- *Applications:* Angell and Heslop (1995); Hannam (1997); Kalakota and Whinston (1997); McMahon and Browne (1998); Treese and Stewart (1998)
- *Practical issues:* Derfler (1998); Dowd (1996); Guengerich et al. (1996); Murhammer et al. (1999); Ptak et al. (1998); Schulman and Smith (1997); Ward (1999)
- *Future trends:* Conner-Sax and Krol (1999); Foster and Kesselman (1998); Huitema (1998); Mambretti and Schmidt (1999)

2. THE ROLE OF COMPUTER NETWORKING

Computer networks (see Figure 1) are frequently considered as analogous to the nervous system in the human body. Indeed, like the nerves, which connect the more or less autonomous components in the body, the links of a computer network provide connections between the otherwise stand-alone units in a computer system. And just like the nerves, the links of the network primarily serve as communication channels between the computers. This is also the basis of the global network of networks, the Internet.

The main task in both cases is information transfer. Connected and thus communicating elements of the system can solve their tasks more efficiently. Therefore, system performance is not just the sum of the performance of the components but much more, thanks to the effect of the information exchange between the components.

Computer networks consist of computers (the basic components of the system under consideration), connecting links between the computers, and such additional things as devices making the information transfer as fast, intelligent, reliable, and cheap as possible.

- Speed depends on the capacity of the transmission lines and the processing speed of the additional devices, such as modems, multiplexing tools, switches, and routers. (A short description of these devices is given in Section 4, together with more details about transmission line capacity.)
- Intelligence of the network depends on processing capabilities as well as stored knowledge of these active devices of the network. (It is worth mentioning that network intelligence is different from the intelligence of the interconnected computers.)
- Reliability stems from, first of all, the decreased risk of losing, misinterpreting, or omitting necessary information, as well as from the well preprocessed character of the information available within the system.

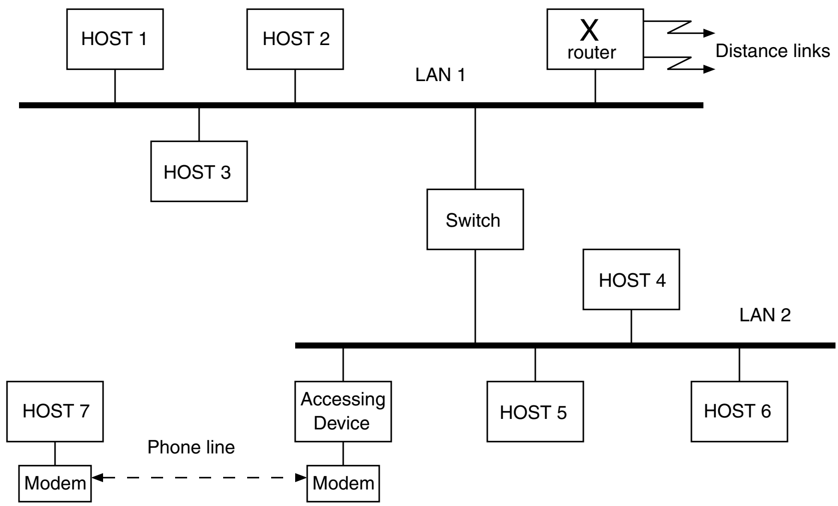

Figure 1 Networks and Network Components.

- Cost is associated with a number of things. The most important direct costs are purchase costs, installation costs, and operational costs. Indirect costs cover a wide range, from those stemming from faulty system operation to those associated with loss of business opportunity. In the planning of the system, the cost/performance ratio must always be taken into consideration.

It is important not to forget that adequate system operation requires adequate operation of all components.

Obviously, adequate operation of each component in a system requires intensive interaction among them. However, direct information transfer between these components makes the interaction much faster and more reliable than traditional means of interaction. This means highly elevated adequacy in the behavior of each component and thus of the system itself.

Another important capability of a system consisting of connected and communicating components is that the intelligence of the individual components is also connected and thus each component can rely not only on its own intelligence but also on the knowledge and problem solving capability of the other components in the system. This means that the connections between the computers in the network result in a qualitatively new level of complexity in the system and its components. The network connections allow each system component to:

- Provide information to its partners
- Access the information available from the other system components
- Communicate while operating (thus, to continuously test, correct, and adapt its own operation)
- Utilize highly sophisticated distant collaboration among the system components
- Exploit the availability of handling virtual environments*

These features are described briefly below. As will be shown below, these basic features pave the way for:

- The systematic construction of network infrastructures
- The continuous evolution of network services
- The rapid proliferation of network-based applications
- The exponentially increasing amount of information content accumulated in the network of interconnected computers

2.1. Information Access

Information access is one of the key elements of what networking provides for the users of a computer connected to other computers. The adequate operation of a computer in a system of interconnected computers naturally depends on the availability of the information determining what, when, and how to perform with it. Traditionally, the necessary information was collected through "manual" methods: through conversation and through traditional information carriers: first paper, and then, in the computer era, magnetic storage devices.

These traditional methods were very inefficient, slow, and unreliable. The picture has now totally changed through computer networking. By the use of the data communication links connecting computers and the additional devices making possible the transfer of information, all the information preprocessed at the partner sites can be easily, efficiently, immediately, and reliably downloaded from the partner computers to the local machine.

The model above works well until individual links are supposed to serve only the neighbor computers in accessing information at each other's sites. But this was the case only at the very beginning of the evolution of networking. Nowadays, millions of computers can potentially access information stored in all of the other ones. Establishing individual links between particular pairs of these computers would be technically and economically impossible.

The solution to the problem is a hierarchical model (see Figure 2). Very high-speed backbones (core network segments) interconnect major nodes located at different geographical regions, and somewhat lower-speed access lines transfer the information from/to these nodes to/from those servers within the region, which then take care of forwarding/collecting the information traffic in their local

*Handling virtual environments means that the individual components don't communicate continuously with the outside world but systematically use a mirrored synthetic picture about their embedding. This is done by utilizing the information accumulated by their neighbors and thus constructing and using a virtual reality model about the outside world. Of course, direct access to this outside world may provide additional information by which such a virtual reality model can be completed so that the result is a so-called augmented reality model.

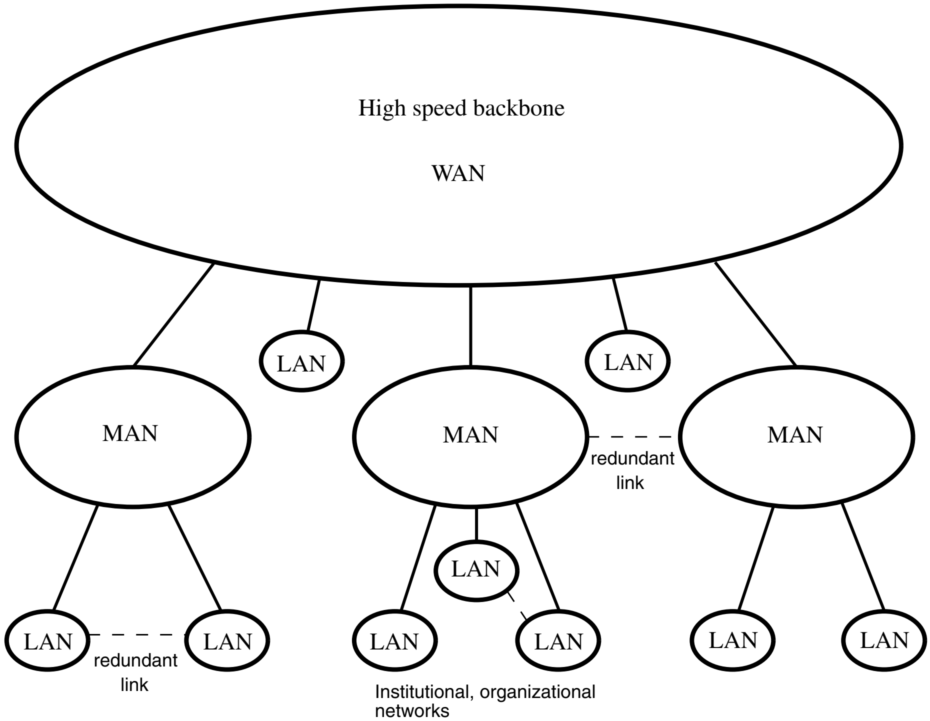

Figure 2 Hierarchical Network Model.

area. Local area networks (LANs) arise from this way of solving the information traffic problem. The hierarchy can be further fragmented, and thus metropolitan area networks (MANs) and wide area networks (WANs) enter the picture. Although this model basically supposes tree-like network configurations (topologies), for sake of reliability and routing efficiency, cross-connections (redundant routes) are also established within this infrastructure.

If a network user wishes to access information at a networked computer somewhere else, his or her request is sent through the chain of special networking devices (routers and switches; see Section 4) and links. Normally this message will be sliced into packets and the individually transferred packets will join possibly thousands of packets from other messages during their way until the relevant packets, well separated, arrive at their destination, recombine, and initiate downloading of the requested information from that distant site to the local machine. The requested information will then arrive to the initiator by a similar process. If there is no extraordinary delay anywhere in the routes, the full process may take just a few seconds or less. This is the most important strength of accessing information through the network.

Note that the described method of packet-based transmission is a special application of time division multiplexing (TDM). The TDM technique utilizes a single physical transmission medium (cable, radiated wave, etc.) for transmitting multiple data streams coming from different sources. In this application of TDM, all the data streams are sliced into packets of specified length of time, and these packets, together with additional information about their destination, are inserted into time frames of the assembled new, higher-bit rate data stream. This new data stream is transmitted through a well-defined segment of the network. At the borders of such network segments, the packets may be extracted from the assembled data stream and a new combination of packets can be assembled by a similar way. By this technique, packets from a source node may even arrive to their destination node through different chains of network segments. More details about the technique will be given in Sections 6.2 and 6.3.

The next subsection describes the basic prerequisites at the site providing the information. First a basic property of the data communication links will be introduced here.

The capability of transmitting/processing high-volume traffic by the connecting lines in the network depends on their speed/bandwidth/capacity. Note that these three characteristic properties are equivalent in the sense that high bandwidth means high speed and high speed means high capacity. Although all three measures might be characterized by well-defined units, the only practically applied unit for them is the one belonging to the speed of the information transfer, bits per second (bps). However, because of the practical speed values, multiples of that elementary unit are used, namely Kbps, Mbps, Gbps, and more recently, Tbps (kilobits, megabits, gigabits, and terabits per second, respectively, which means thousands, millions, billions and thousand billions of bits per second).

2.2. Information Provision

Access to information is not possible if the information is not available. To make information accessible to other network users, information must be provided. Provision of information by a user (individual or corporate or public) is one of the key contributors to the exponential development of networking and worldwide network usage. How the information is provided, whether by passive methods (by accessible databases or, the most common way today, the World Wide Web on the Internet) or by active distribution (by direct e-mail distribution or by using newsgroups for regular or irregular delivery), is a secondary question, at least if ease of accessing is not the crucial aspect.

If the information provided is freely accessible by any user on the network, the information is public. If accessibility is restricted, the information is private. Private information is accessible either to a closed group of users or by virtually any user if that user fulfils the requirements posed by the owner of the information. Most frequently, these requirements are only payment for the access. An obvious exception is government security information.

All information provided for access has a certain value, which is, in principle, determined by the users' interest in accessing it. The higher the value, the more important is protection of the related information. Providers of valuable information should take care of protecting that information, partly to save ownership but also in order to keep the information undistorted. Faulty, outdated, misleading, irrelevant information not only lacks value but may also cause problems to those accessing it in the belief that they are getting access to reliable content.

Thus, information provision is important. If the provider makes mistakes, whether consciously or unconsciously, trust will be lost. Providing valuable (relevant, reliable, and tested) information and taking care of continuous updating of the provided information is critical if the provider wants to maintain the trustworthiness of his information.

The global network (the millions of sites connected by the network) offers an enormous amount of information. The problem nowadays is not simply to find information about any topic but rather to find the best (most relevant and most reliable) sources of the required information. Users interested in accessing proper information should either turn to reliable and trusted sources (providers) or take the information they collect through the Internet and test it themselves.

That is why today the provision of valuable information is separated from ownership. Information brokerage plays an increasingly important role in where to access adequate information. Users looking for information about any topic should access either those sources they know and trust or the sites of brokers who take over the task of testing the relevance and reliability of the related kinds of information. Information brokerage is therefore an important new kind of network-based service.

However, thoroughly testing the validity, correctness, and relevance of the information is difficult or even impossible, especially because the amount of the stored and accessible information increases extremely fast. The final answer to the question of how to control the content (i.e., the information available worldwide) is still not known. One of the possibilities (classifying, rating, and filtering) is dealt with in more detail in a later section.

Besides the difficulty of searching for adequate information, there is a further important aspect associated with the increasing interest worldwide in accessing information. It relates to the volume of information traffic generated by the enormous number of attempts to download large amounts of information.

Traffic volume can be decreased drastically by using the caching technique (see Figure 3). In the case of frequently accessed information sources, an enormous amount of traffic can be eliminated by storing the frequently requested information in dedicated servers that are much closer to the user than the original information provider. The user asking for these kinds of information doesn't even know if the requested information arrives at the requesting site from such a cache server or from the primary source of the information. Hierarchical cache server systems help to avoid traffic congestion on network links carrying intensive traffic.

2.3. Electronic Communication

Providing and accessing information, although extremely important, is just one element of information traffic through the network. Electronic communication in general, as far as the basic function of the network operation is concerned, involves a lot more.

The basic function of computer networks is to make it possible to overcome distance, timing, and capability problems.

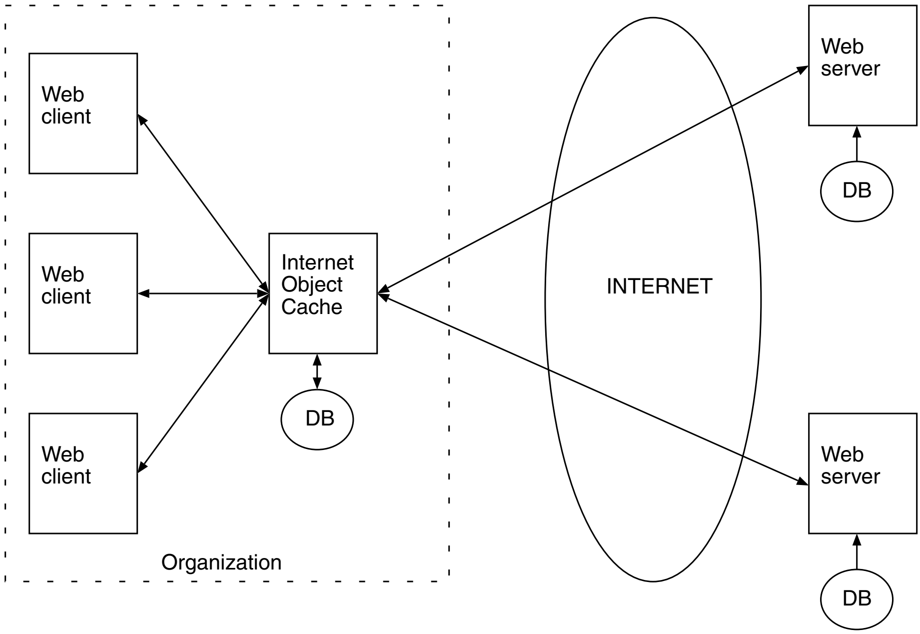

Figure 3 Caching.

- Distance is handled by the extremely fast delivery of any messages or other kinds of transmitted information at any connected site all over the world. This means that users of the network (again individual, corporate, or public) can do their jobs or perform their activities so that distances virtually disappear. Users or even groups of users can talk to each other just as if they were in direct, traditional contact. This makes, for example, remote collaboration a reality by allowing the formation of virtual teams.
- Timing relates to the proper usage of the information transmitted. In the most common case, information arrives at its destination in accordance with the best efforts method of transmission provided by basic Internet operation. This means that information is carried to the target site as soon as possible but without any guarantee of immediate delivery. In most cases this is not a problem because delays are usually only on the order of seconds or minutes. Longer delays usually occur for reasons other than simple transmission characteristics. On the other hand, e-mail messages are most often read by the addressees from their mail servers much later than their arrival. Although obviously this should not be considered the basis of network design, the fact that the recipient doesn't need to be at his or her machine when the message arrives is important in overriding timing problems. However, some information can be urgently needed, or real-time delivery may even be necessary, as when there are concurrent and interacting processes at distant sites. In these cases elevated-quality services should be provided. Highest quality of service (QoS) here means a guarantee of undisturbed real-time communication.
- Capability requirements arise from limited processing speed, memory, or software availability at sites having to solve complex problems. Networked communication helps to overcome these kinds of problems by allowing distributed processing as well as quasisimultaneous access to distributed databases or knowledge bases. Distributed processing is a typical application where intensive (high-volume) real-time (high-speed) traffic is assumed, requiring high-speed connection through all the routes involved in the communication.

The latter case already leads us to one of the more recent and most promising applications of the network. Although in some special areas, including highly intensive scientific computations, distributed processing has been used for some time, more revolutionary applications are taking place in networked remote collaboration. This new role of computer networking is briefly reviewed in the following section.

2.4. Networked Collaboration

Computer networks open new ways of increasing efficiency, elevating convenience, and improving cost and performance of collaboration. The basis of this overall improvement is provided by the properties of computer networks mentioned above, their ability to help overcome distance, timing, and capability problems. The workplace can thus be dislocated or even mobile, the number of collaborators can be theoretically infinite, the cooperating individuals or groups can be involved in joint activities and in shifted time periods (thus even geographical time zone problems can be virtually eliminated), and the collaborating partners can input their information without having to move high-volume materials (books and other documents) to a common workplace. These are all important issues in utilizing the opportunities that networks provide for collaborating when communities are geographically dispersed.

However, the key attribute of networked collaboration is not simply the possibility of avoiding having to collect all the collaborating partners at a well-defined venue and time for a fixed time period. The fact that the network interconnects computers, rather than human users, is much more useful. The reason is that networked computers are not just nodes in a network but also intelligent devices. They not only help exchange information between the collaborating individuals or groups, but also provide some important additional features.

- The information processing capability of the computers at the network nodes of the collaborating parties may also solve complex problems.
- The intelligence of each computer connected and participating in the joint activities is shareable among the collaborating partners.
- The intelligence in these distant computers can be joined to solve otherwise unsolvable problems by utilizing an enormously elevated computing power.

However, appropriate conditions for efficient and successful collaboration assume transmission of complex information. This information involves not only the data belonging to the joint task but also the multimedia information stemming from the networked character of the collaboration. This may be a simple exchange of messages, such as e-mail conversation among the participants, but may also be a true videoconference connection between the cooperating sites. Thus, in many cases networked collaboration requires extremely high bandwidths. Real-time multimedia transmission, as the basis of these applications, supposes multi-Mbps connectivity throughout. Depending on the quality demands, the acceptable transmission speed varies between one or two Mbps and hundreds of megabits per second. The latter speeds have become a reality with the advent of gigabit networking.

If the above conditions are all met, networked collaboration provides multiple benefits:

- Traditional integration of human contributions
- Improvement in work conditions
- Integration of the processing power of a great many machines
- High-quality collaboration attributes through high-speed multimedia transmission

The highest level of networked collaboration is attained by integrated environments allowing truly integrated activities. This concept will be dealt with later in the chapter. The role of the World Wide Web in this integration process, the way of building and using intranets for, among others, networked collaboration, and the role of the different security aspects will also be outlined later.

2.5. Virtual Environments

Network-based collaboration, by using highly complex tools (computers and peripherals) and high-speed networking, allows the participating members of the collaborating team to feel and behave just as if they were physically together, even though the collaboration is taking place at several distant sites. The more the attributes of the environment at one site of the collaboration are communicated to the other sites, the more this feeling and behavior can be supported. The environment itself cannot be transported, but its attributes help to reconstruct a virtual copy of the very environment at a different site. This reconstruction is called a virtual environment.

This process is extremely demanding as far as information processing at the related sites and information transfer between distant sites are concerned. Although the available processing power and transmission speed increase exponentially with time, the environment around us is so complex that full copying remains impossible. Thus, virtual environments are only approximated in practice.

Approximation is also dictated by some related practical reasons:

- Practical feasibility (even if a certain level of copying the target environment is theoretically possible, it would require the implementation of prohibitively complex practical solutions)
- Cost effectiveness (although the practical solution is feasible, the cost is so high that the application doesn't allow the very level of emulation)
- Intended approximation (although a certain level of true copying would be possible and cost effective, the application requires a model of the virtual copy rather than the true original because of experimental applications, trials regarding modifying the environment, etc.)
- Inaccessible environments (the environment to be copied is inaccessible either because of being outside the connected sites or because of transient lack of some prerequisites)

In such cases, virtual environments are substituted for by virtual reality models. These models provide virtual reality by mimicking the supposed environment but not copying it. However, in some applications a mix of virtual copies and virtual models is needed. In *augmented reality,* transmitted information about the attributes of the considered real environment is used together with model parameters about a model environment to build up a virtual construct for the supposed environment.

Virtual constructs are essential tools in a number of applications utilizing the concept of telepresence and/or teleimmersion by users of networked computers. Teleinstruction, telemedicine, and different types of teleworking are examples of how the development of networked communication and collaboration leads to an increasing number of applications using virtual and augmented reality.

3. A SHORT HISTORY OF COMPUTER NETWORKING

Computer networking started in the United States in the late 1960s. Through the rapid development of computing and telecommunications, in 1969 the ARPANet program was launched.

The first voice transmission over cables (1876) and by radio waves (1896), the first radio transfer of video images (1936), the first electronic computer (1944), and the first transistor (1948) preceded this milestone of technical history, but the first microprocessor (1971) and the first IBM personal computer (1981), as well as the first digital radio transmission of voice (1988) and image (1991), came later.

The ARPANet program started with the aim of connecting just a few sites, but the first networking experiments immediately proved that ''Internetting'' was a feasible idea. The program led to the development of the TCP/IP protocol suite (see Section 6), having served as the basis of Internet technology for more than 30 years. These developments resulted in a proliferation in activity, and the number of Internet hosts (networked computers serving several users by Internet services) rapidly increased. As a result of the annual doubling of host numbers, as of the end of the 1990s, dozens of millions of hosts served a worldwide community of an estimated more than 200 million users.

The Internet was established early in the 1980s as a solid infrastructure of networking, mainly in the United States but also with connected subnetworks in Europe and the Asia Pacific region. The organizational, coordinating, and mainly informal controlling frameworks of the Internet were also established by that time. A conscious standardization process also started. These activities resulted in a solid background for an efficiently operating global system of interconnected computers.

Perhaps the most important property of the TCP/IP protocol suite is that it makes possible the interconnection and interworking of subnetworks. The introduction and wide application of the concept of such subnetworks played a fundamental role in the rapid and successful proliferation of the Internet technology. The value of integrating diverse subnetworks into one cohesive Internet is what makes the Internet such a great network tool.

After more than 10 years of exclusively special (defense, and later scientific) applications of the emerging network, commercial use began to evolve. Parallel activities related to leading-edge academic and research applications and everyday commodity usage appeared. This parallelism was more or less maintained during more recent periods, as well, serving as a basis of a fast and continuous development of the technology. A breakthrough occurred when, in the mid-1980s, the commercial availability of system components and services was established. Since then the supply of devices and services has evolved continuously, enabling any potential individual, corporate, or public user simply to buy or hire what it needed for starting network-based activities.

While the leading edge in the early 1980s was still just 56 Kbps (leased line), later in the decade the 1.5 Mbps (T1) transmission speed was achieved in connecting the emerging subnetworks.

Late in the 1980s the basics of electronic mail service were also in use. This was probably the most important step in the application of networking to drive innovation prior to the introduction of the World Wide Web technology.

Early in the 1990s the backbone speed reached the 45 Mbps level, while the development and spread of a wide range of commercial information services determined the directions of further evolution in computer networking.

The 1990s brought at least three new developments.

1. The development of a new-generation Internet was initiated in order to overcome the weaknesses of the traditional Internet technology in handling the exponential growth of the user community, the fast change in quality, reliability, and security requirements, and the increased complexity of the transmitted information. (Strictly speaking, security is outside of the Internet itself. The need for security became a hot topic with the advent of readily accessible information via the Internet. The present TCP/IP protocol suite does not provide security itself.)
2. The speed of transmission (or the speed of the network) and the coverage of the global infrastructure (including worldwide proliferation as well as dense regional network population) reached new records. This development is still far from saturation. As of the turn of the century, the dozens of millions of Internet hosts had, not homogeneously but steadily, spread out through all continents of the world. At the same time, the high-level backbones operated, at least in the more developed areas, at Gbps speed, and the continents were connected by thousands of optical fibers capable of keeping an extremely high volume of traffic.
3. World Wide Web technology has become a standard all over the world, and by the end of the 1990s an enormous amount of information provision and information access had taken place through the Web.

Through these revolutionary advancements, networking has begun to play an important role worldwide in reforming a great many human activities. Parallel to the developments in speed, quality, reliability, manageability, and cost/performance, many international organizations, national governments, and civil initiatives have launched joint projects with the aim of establishing a new and unavoidable worldwide "Information Society." New forms of cooperation among governments, industry, telecom, and network service providers, as well as civil organizations, have emerged, and further development is inevitable.

4. THE NETWORKING INFRASTRUCTURE

As mentioned in Section 2.1, the background of the information transfer is provided by the networking infrastructure. As mentioned, the solution of the problem of how to handle millions of transfers of data simultaneously over the global network in very short time frames is based on a hierarchical model. The infrastructure consists of links and nodes whose role depends on their status in the hierarchy. At the highest level are the largest information exchange centers, which are connected with very high-speed lines. The very high-speed networks of these large exchange nodes and the high-capacity links connecting them are called backbones. The main nodes of the backbones communicate among each other and also collect and distribute the relevant traffic from and to the regional nodes around them. The lines connecting the large exchange nodes to the regional nodes also work at very high speeds and are therefore able to transmit very high-volume traffic.

Lower in the hierarchy, around the regional nodes, somewhat lower-speed access lines transfer the information from and to the area servers (connected to routers and/or switches, i.e., special machines taking care of the direction of the transmitted information packets in the network) within the region. These active devices serve either LANs, MANs, or WANs, depending on the topology and the further hierarchical structure of the network around these area servers. Their task is to collect/distribute the information traffic in their area.

The hierarchical structure might mean a simple (but of course highly segmented and in this respect very complex) tree-like topology. However, as already mentioned, for the purpose of reliability and routing efficiency, cross-connections (redundant routes) are also established within the infrastructure. Obviously, the difficulty of handling the traffic is in relation with the complexity of the network structure, but this is the price for satisfactory global network operation.

The key elements, besides the communication links themselves, in this technique of information transfer through the network are the routers and switches, which take care of directing the information packets through the nodes in this hierarchy so that they finally arrive at their destination.

The speed of the information transfer depends on the bandwidth of the links and the processing speed of the routers and switches in the infrastructure. Taking into account the requirements stemming from the characteristics of the transmitted information and the traffic volume, that is, from the estimated average and peak number of requested transmission transactions, allows the speed of the different sections of the infrastructure to be determined. Some basic figures are as follows:

- Simple alphanumeric messages require a speed that is not too high at the last mile sections, that is, at the lines closest to the transmitters/receivers of the information. Several Kbps is considered satisfactory in such cases of message delivery without special requirements regarding the delivery time.

- Real-time transmission of similarly simple alphanumeric information requires either similar speed, as above, or, in case of higher volume of these kinds of data, a certain multiple of that speed.
- Real-time transmission of sampled, digitized, coded, and possibly compressed information, stemming from high-fidelity voice, high-resolution still video, or high-quality real video signals do require much higher speed. The last-mile speed in these cases ranges from dozens of Kbps to several and even hundreds of Mbps (HDTV). This information is quite often cached by the receiver to overcome short interruptions in the network's delivery of the information at high speed, providing the receiver a smooth uninterrupted flow of information.

This means that although until recently the last-mile line speeds were in the range of several dozens of Kbps, the area access lines were in the range of several Mbps, and the top-level backbones were approaching the Gbps speed, lately, during the 1990s, these figures have increased to Mbps, several hundreds of Mbps, and several Gbps level, respectively. The near future will be characterized by multi-Gbps speed at the backbone and area access level, and the goal is to reach several hundreds of Mbps even at the last-mile sections. Of course, the development of the infrastructure means coexistence of these leading-edge figures with the more established lower-speed solutions. And obviously, in the longer term, the figures may increase further.

As far as the organization of the operation of the infrastructure, the physical and the network infrastructure should be distinguished. Providing physical infrastructure means little more than making available the copper cables or fibers and the basic active devices of the transmission network, while operating the network infrastructure also means managing of the data traffic. Both organizational levels of the infrastructure are equally important, and the complexity of the related tasks has, in an increasing number of cases, resulted in the jobs being shared between large companies specializing in providing either the physical or the network infrastructure.

The first step from physical connectivity to network operation is made with the introduction of network protocols. These protocols are special complex software systems establishing and controlling appropriate network operation. The most important and best-known such protocol is the Internet Protocol.

The elementary services are built up on top of the infrastructure outlined above. Thus, the next group of companies working for the benefit of the end users is the Internet service providers (ISPs), which take care of providing network access points.

Although network services are discussed separately in Section 7, it should be mentioned here that one of the basic tasks of these services is to take care of the global addressing system. Unique addressing is perhaps the most important element in the operation of the global network. Network addresses are associated with all nodes in the global network so that the destination of the transmitted information can always be specified. Without such a unique, explicit, and unambiguous addressing system, it would not be possible even to reach a target site through the network. That is why addressing (as well as naming, i.e., associating unique symbolic alphanumeric names with the numeric addresses) is a crucial component of network operation. The task is solved again by a hierarchy of services performed by the domain name service (DNS) providers. This issue is dealt with in more detail later. Note that it is possible, as a common practice, to proxy the end-node computer through an access point to share the use of a unique address on the Internet. The internal subnetwork address may in fact be used by some other similarly isolated node in a distinctly separate subnetwork (separate intranet).

Everything briefly described above relates to the global public network. However, with the more sophisticated, serious, and sometimes extremely sensitive usage of the network, the need to establish closed subsets of the networked computers has emerged. Although virtual private networks (subnetworks that are based on public services but that keep traffic separated by the use of special hardware and software solutions) solve the task by simply separating the general traffic and the traffic within a closed community, some requirements, especially those related to security, can only be met by more strictly separating the related traffic. The need for this kind of separation has led to the establishment of intranets (special network segments devoted to a dedicated user community, most often a company) and extranets (bunches of geographically distant but organizationally and/or cooperatively connected intranets, using the public network to connect the intranets but exploiting special techniques for keeping the required security guarantees). Although building exclusively private networks, even wide area ones, is possible, these are gradually disappearing, and relying on public services is becoming common practice even for large organizations.

5. INTERNET, INTRANETS, EXTRANETS

The Internet is the world's largest computer network. It is made up of thousands of independently operated (not necessarily local) networks collaborating with each other. This is why it is sometimes called the network of networks. Today it extends to most of the countries in the world and connects dozens of millions of computers, allowing their users to exchange e-mails and data, use online services, communicate, listen to or watch broadcast programs, and so on, in a very fast and cost-effective way. Millions of people are using the Internet today in their daily work and life. It has become part of the basic infrastructure of modern human life.

As was already mentioned above, the predecessor of the Internet was ARPANet, the first wide area packet switched data network. The ARPANET was created within a scientific research project initiated by the U.S. Department of Defense in the late 1960s. Several universities and research institutes participated in the project, including the University of Utah, the University of California at Los Angeles, the University of California at Santa Barbara, and the Stanford Research Institute. The goal of the project was to develop a new kind of network technology that would make it possible to build reliable, effective LANs and WANs by using different kinds of communication channels and connecting different kinds of computers. By 1974, the basics of the new technology had been developed and the research groups led by Vint Cerf and Bob Kahn had published a description of the first version of the TCP/IP (Transmission Control Protocol and Internet Protocol) suite.

The experimental TCP/IP-based ARPANet, which in its early years carried both research and military traffic, was later split into the Internet, for academic purposes, and the MILNet, for military purposes. The Internet grew continuously and exponentially, extending step by step all over the world. Early in the 1990s, the Internet Society (see www.isoc.org) was established, and it has served since then as the driving force in the evolution of the Internet, especially in the technological development and standardization processes.

The invention of the World Wide Web in 1990 has given new impetus to the process of evolution and prompted the commercialization of the Internet. Today the Internet is a general-purpose public network open to anyone who wishes to be connected.

Intranets are usually TCP/IP-based private networks. They may in fact gateway through a TCP/IP node, but this is not common. They operate separately from the worldwide Internet, providing only restricted and controlled accessibility. Although an intranet uses the same technology as the Internet, with the same kinds of services and applications, it principally serves only those users belonging to the organization that owns and operates it. An intranet and its internal services are closed to the rest of the world. Often this separation is accomplished by using network addresses reserved for such purposes, the so-called ten net addresses. These addresses are not routed in the Internet, and a gateway must proxy them with a normal address to the Internet.

Connecting intranets at different geographical locations via the public Internet, results in extranets. If an organization is operating at different locations and wants to interconnect its TCP/IP-based LANs (intranets), it can use the inexpensive public Internet to establish secure channels between these intranets rather than build very expensive, large-scale, wide area private networks. This way, corporate-wide extranets can be formed, allowing internal users to access any part of this closed network as if it were a local area network.

6. TECHNICAL BACKGROUND

To understand the possibilities of the TCP/IP-based networks fully, it is important to know how they work and what kinds of technical solutions lie behind their services.

6.1. Architecture of the Internet

As mentioned in the previous section, the Internet can be described as a network of local area networks (see Figure 4). Therefore, perhaps the most important component of the Internet is LANs. LANs connect computers or, in other words, hosts, and are connected to other LANs by gateways and communication lines. Thus, the four basic logical components of the Internet (or of intranets) are:

1. Hosts
2. Local area networks
3. Gateways
4. Communications lines

The main role of gateways is to provide connections between the communication lines and the local area networks and to route the data packets toward their destinations. This is why they are often called routers. Gateways also play an important role in security by protecting the local area network against external attacks and other illegal or unauthorized access. Quite often, a security zone, referred to as the DMZ, is set up where network traffic is passed through a gateway into a firewall (see Section 15.2). The firewall then provides application-level security to network traffic before sending the information through an internal router to pass to end systems in the intranet.

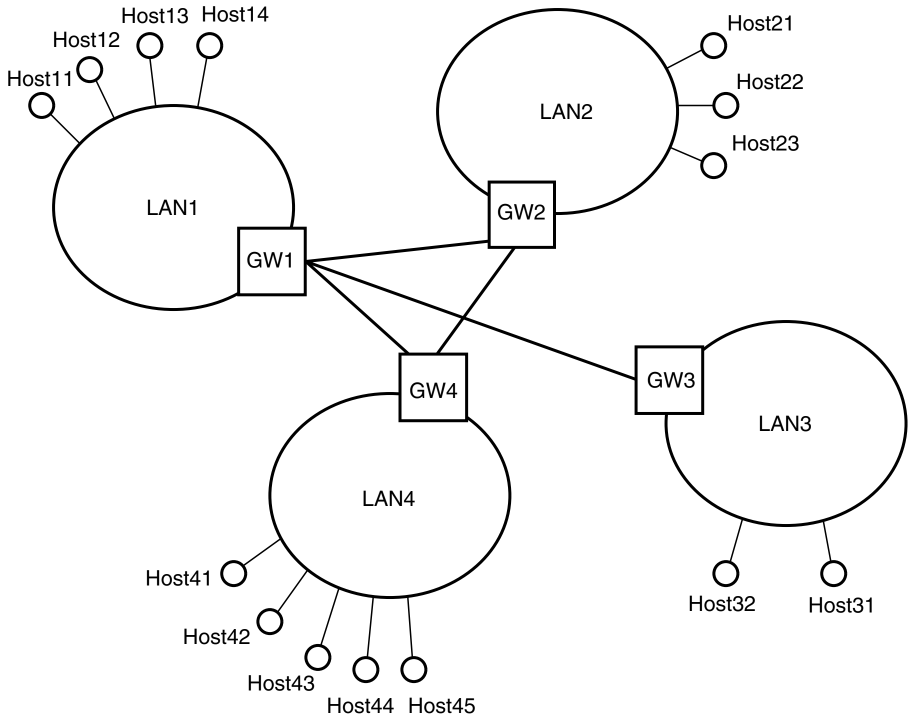

Figure 4 Internet Architecture.

In most cases, communication lines establish steady, 24-hour connection between their endpoints. In order to increase reliability, and sometimes also for traffic load balancing reasons, alternative routes can be built between the local area networks. If any of the lines is broken, the gateways can automatically adapt to the new, changing topology. Thus, the packets are continuously delivered by the alternative routes. Adaptive routing is possible at gateways, particularly for supporting 24×7 operation. It commonly occurs in the backbone of the Internet.

6.2. Packet Switching

Packet switching is one of the key attributes of the TCP/IP technology. The data stream belonging to a specific communication session is split into small data pieces, called packets. The packets are delivered independently at the target host. The separated packets of the same communication session may follow different routes to their destination. In contrast to line-switching communication technologies, in packet switched networks there is no need to set up connections between the communicating units before the start of the requested data transmission. Each packet contains all of the necessary information to route it to its destination. This means that packets are complete from a network perspective.

A good example of line-switched technology is the traditional public phone service. In contrast to packet switching, line switching assumes a preliminary connection setup procedure being performed before a conversation starts. After the connection is set up, an individual communication channel (called a circuit) is provided for the data transmission of that communication session. When the data transmission is over, the connection should be closed.

6.3. Most Important Protocols

A network protocol is a set of rules that determines the way of communication on the network. All the attached hosts must obey these rules. Otherwise they won't be able to communicate with the other hosts or might even disturb the communication of the others. For proper high-level communication, many protocols are needed.

The system of protocols has a layered structure. High-level protocols are placed in the upper layers and low-level protocols in the lower layers. Each protocol layer has a well-defined standard interface by which it can communicate with the other layers, up and down. TCP/IP itself is a protocol group consisting of several protocols on four different layers (see Figure 5). TCP and IP are the two most important protocols of this protocol group.

IP is placed in the internetworking layer. It controls the host-to-host communication on the network. The main attributes of IP are that it is connectionless, unreliable, robust, and fast. The most astounding of these is the second attribute, unreliability. This means that packets may be lost, damaged, and/or multiplied and may also arrive in mixed order. IP doesn't guarantee anything about the safe transmission of the packets, but it is robust and fast. Because of its unreliability, IP cannot satisfy the needs of those applications requiring high reliability and guaranteed QoS.

Built upon the services of the unreliable IP, the transport layer protocol, TCP, provides reliable communication for the applications. Reliability is guaranteed by positive acknowledgement and automatic retransmission. TCP performs process-to-process communication. It also checks the integrity of the content. TCP is connection oriented, although the TCP connections are virtual, which means that there is no real circuit setup, only a virtual one. A TCP connection is a reliable, stream-oriented, full-duplex, bidirectional communication channel with built-in flow control and synchronization mechanisms.

There are also routing and discovery protocols that play an important role in affecting network reliability, availability, accessibility, and cost of service.

6.4. Client–Server Mechanism

Any communication on a TCP/IP network is performed under the client–server scheme (see Figure 6). When two hosts are communicating with each other, one of them is the client and the other is the server. The same host can be both client and server, even at the same time, for different communication sessions, depending on the role it plays in a particular communication session. The client sends requests to the server and the server replies to these requests. Such servers include file servers, WWW servers, DNS servers, telnet servers, and database servers. Clients include ftp (file transfer) programs, navigator programs (web browsers), and telnet programs (for remote access). Note that the term *server* does not imply special hardware requirements (e.g. high speed, or large capacity, or continuous operation).

The typical mode of operation is as follows: The server is up and waits for requests. The client sends a message to the server. The requests arrive at particular ports, depending on the type of the expected service. Ports simply address information that is acted upon by the serving computer. Clients provide port information so that server responses return to the correct application. Well-known ports are assigned to well-known services. Communication is always initiated by the client with its first message.

Because multiple requests may arrive at the server at the same time from different clients, the server should be prepared to serve multiple communication sessions. Typically, one communication

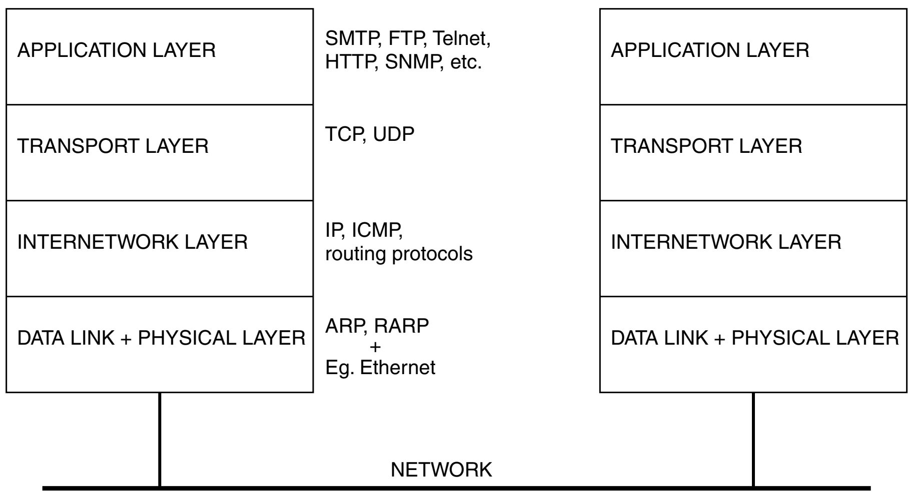

Figure 5 The Layered TCP/IP Protocol Stack.

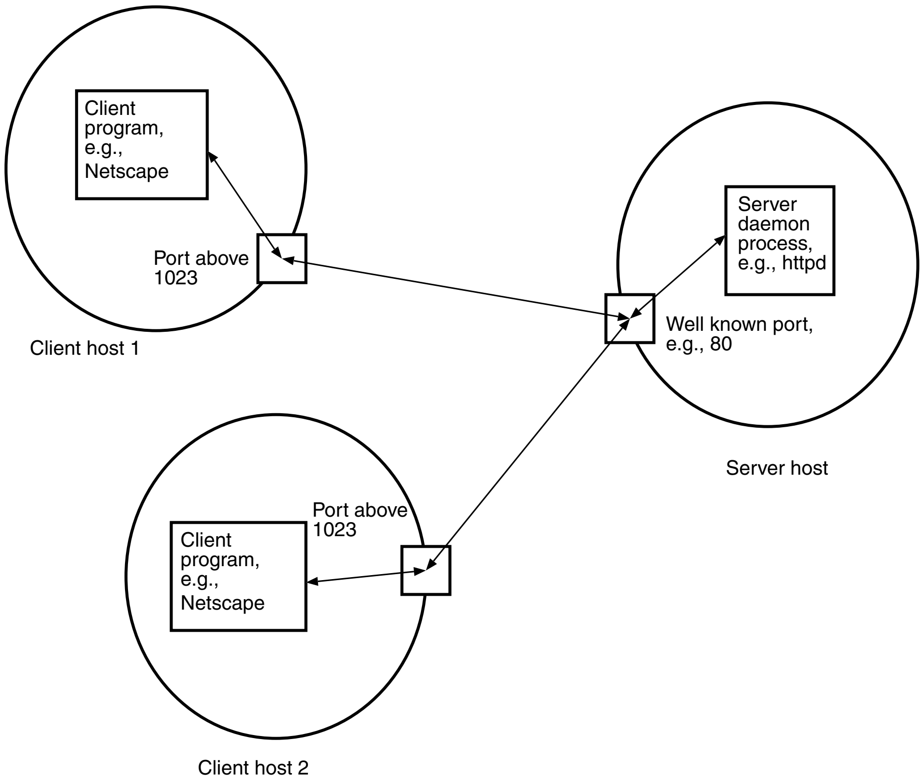

Figure 6 Client/Server Scheme.

session is served by one process or task. Therefore, servers are most often implemented on multi-tasking operating systems. There is no such requirement at the client side, where only one process at a time may be acting as a client program.

One of the main features in the new concept of computing grids is the substitution of the widely used client–server mechanism by realizations of a more general distributed metacomputing principle in future computer networking applications (see Section 12.3).

6.5. Addressing and Naming

In order to be unambiguously identifiable, every host in a TCP/IP network must have at least one unique address, its IP address. IP addresses play an essential role in routing traffic in the network. The IP address contains information about the location of the host in the network: the associated number (address) of the particular LAN and the associated number (address) of the host in the LAN.

Currently, addresses in the version 4 Internet protocol are 32 bits long (IPv4 addresses) and are classified into five groups: A, B, C, D, and E classes. Class A has been created for very large networks. Class A networks are rare. They may contain up to 2^{24} hosts. Class B numbers (addresses) are given to medium-sized networks (up to 65,534 hosts), while class C network numbers are assigned to small (up to 254 hosts) LANs.

The 32-bit-long IP addressing allows about 4 billion different combinations. This means that in principle, about 4 billion hosts can be attached to the worldwide Internet, by using IPv4 addressing. However, intranets, because they are not connected to the Internet or connected through firewalls, may use IP addresses being used by other networks too.

Because the number of Internet hosts at the start of the third millennium is still less than 100 million, it seems that the available IPv4 address range is wide enough to satisfy all the needs. However, due to the present address classification scheme and other reasons, there are very serious limitations in some address classes (especially in class B). The low efficiency of the applied address distribution scheme has led to difficult problems. Because of the high number of medium-sized networks, there are more claims for class B network numbers than the available free numbers in this class.

Although many different suggestions have been made to solve this situation, the final solution will be brought simply by implementing the new version of the Internet protocol, IPv6, intended to be introduced early in the 2000s. It will use 128-bit-long IP addresses. This space will be large enough to satisfy any future address claims even if the growth rate of the Internet remains exponential.

There are three different addressing modes in the current IPv4 protocol:

- Unicast (one to one)
- Multicast (one to many)
- Broadcast (one to all)

The most important is unicast. Here, each host (gateway, etc.) must have at least one class A, class B, or class C address, depending on the network it is connected to. These classes provide unicast addresses.

Class D is used for multicasting. Multicast applications, such as radio broadcasting or video conferencing, assign additional D class addresses to the participating hosts.

Class E addresses have been reserved for future use.

Class A, B, and C addresses consist of three pieces of information: the class prefix, the network number, and the host number. The network number, embedded into the IP address, is the basic information for routing decisions. If the host number part of the address contains only nonzero bits (1s), the address is a broadcast address to that specific network. If all the bits in the host number part are zeroes (0s), the address refers to the network itself.

In order to make them easier to manage, IP addresses are usually described by the four bytes they contain, all specified in decimal notation, and separated by single dots. Examples of IP addresses are:

- Class A: 16.1.0.250
- Class B: 154.66.240.5
- Class C: 192.84.225.2

The human interface with the network would be very inconvenient and unfriendly if users had to use IP addresses when referring to computers they would like access. IP addresses are long numbers, making them inconvenient and difficult to remember and describe, and there is always a considerable risk of misspelling them. They don't even express the type, name, location, and so on of the related computer or the organization operating that computer. It is much more convenient and reliable to associate descriptive names with the computers, the organizations operating the computers, and/or the related LANs. The naming system of the TCP/IP networks is called the domain name system (DNS). In this system there are host names and domains.

Domains are multilevel and hierarchical (see Figure 7). They mirror the organizational/administrative hierarchy of the global network.

At present, top-level domains (TLDs) fall into two categories:

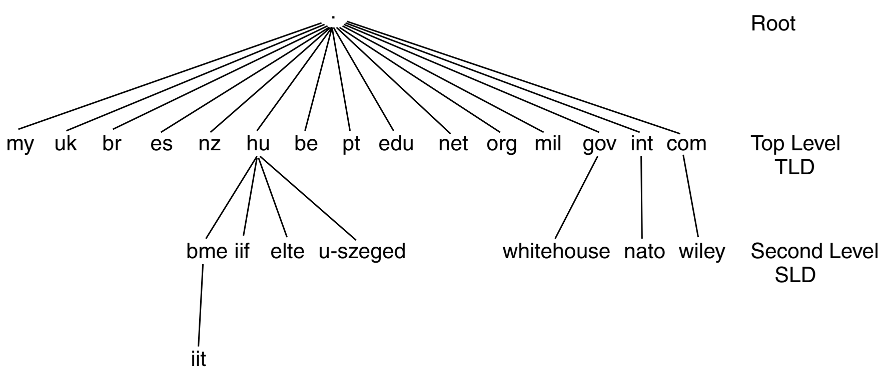

Figure 7 DNS Hierarchy.

1. Geographical (two-letter country codes by ISO)
2. Organizational (three-letter abbreviations reflecting the types of the organizations: edu, gov, mil, net, int, org, com)

Second-level domains (SLDs) usually express the name or the category of the organization, in a particular country. Lower-level subdomains reflect the hierarchy within the organization itself.

The tags of the domain names are separated by dots. The first tag in a domain name is the hostname. The tags are ordered from most specific to least specific addressing.

The domain name service is provided by a network of connected, independently operated domain name servers. A domain name server is a database containing information (IP address, name, etc.) about hosts, domains, and networks under the authority of the server. This is called the domain zone.

Domain name servers can communicate with each other by passing information retrieved from their databases. If a host wishes to resolve a name, it sends a query to the nearest domain name server, which will provide the answer by using its local database, called a cache. However, if necessary, it forwards the query to other servers. The Internet domain name service system is the largest distributed database in the world.

7. OVERVIEW OF NETWORKING SERVICES

Thirty years ago, at the beginning of the evolution of computer networking, the principal goal was to create a communication system that would provide some or all of the following functional services:

- Exchanging (electronic) messages between computers (or their users)—today the popular name of this service is e-mail
- Transferring files between computers
- Providing remote access to computers
- Sharing resources between computers

E-mail is perhaps the most popular and most widely used service provided by wide area computer networks. E-mail provides easy-to-use, fast, cheap, and reliable offline communication for the users of the network. It allows not only one-to-one communication but one-to-many. Thus, mailing lists can be created as comfortable and effective communication platforms for a group of users interested in the same topics.

Of course, file transfer is essential. Computer files may contain different kinds of digitized information, not only texts and programs but formatted documents, graphics, images, sounds, movies, and so on. File transfer allows this information to be moved between hosts.

With the remote access service, one can use a distant computer just as if one were sitting next to the machine. Although not all operating systems allow remote access, most of the multiuser operating systems (UNIX, Linux, etc.) do.

Sharing resources makes it possible for other machines to use the resources (such as disk space, file system, peripheral devices, or even the CPU) of a computer. Well-known examples of resource sharing are the file and printer servers in computer networks.

Besides the above-listed basic functions, computer networks today can provide numerous other services. Operating distributed databases (the best example is the World Wide Web), controlling devices, data acquisition, online communication (oral and written), teleconferencing, radio and video broadcasting, online transaction management, and e-commerce are among the various possibilities.

8. OVERVIEW OF NETWORK-BASED APPLICATIONS

As mentioned at the beginning of this chapter, computer networking has gradually entered all possible fields of applications. Providing an exhaustive overview of the many applications would be practically impossible, so only a brief overview will be given here of what is (and what still will be) available. Classification itself is subjective in this case, so the aspects mentioned below provide only an approach to systematic handling of the wide spectrum of network-based applications.

First, applications can be categorized by what kind of community is interested in them. The two main groups are:

- Public applications, available for any individual or group/organization of individuals
- Private applications, available only to a certain subset of individuals or groups/organizations

Although the realizations of the applications within the two distinct groups are not different in principle from each other, there are at least two basic differences in the background attributes. On one hand, the infrastructure behind an application closely reflects the width of its user community (wide area public network, in contrast to intranet/extranet-type solutions using private or virtual private networks). On the other hand, security issues are definitely different: on top of the general security and privacy requirements, special security and authorization aspects require special firewall techniques in the case of private applications.

Second, applications can be distinguished by how they are utilized. The three main classes are:

- Personal applications (individual and family use)
- Administrative applications (use by governments, municipalities, etc., in many cases with overall, or at least partly authorized, citizen access)
- Industrial applications (practically any other use, from agriculture to manufacturing to commercial services)

The major differences among these three classes are the required levels of availability, geographical distribution, and ease of access. As far as networking is concerned, the classes differ in the required quality of service level/grade.

Third, there is a well-structured hierarchy of the networked applications based on their complexity. This hierarchy more or less follows the hierarchy of those single-computer applications that are enhanced by the introduction of network-based (distributed, dislocated) operation:

- Low-level applications in this hierarchy include distributed word processing, database handling, document generation, computer numerical control, etc.
- Medium-complexity applications include network-integrated editing and publishing, networked logistics and inventory control, CAD (computer aided design), CAM (computer aided manufacturing), etc.
- High-complexity applications include integrated operations management systems, such as integrated management and control of a publishing/printing company, integrated engineering/management of a CAD-CAM-CAT system, etc.

Note that high-complexity applications include lower-complexity applications, while low-complexity applications can be integrated to build up higher-level applications in the hierarchy. Computer networking plays an important role not only in turning from single-computer applications to network-based/distributed applications but also in the integration/separation process of moving up and down in the hierarchy.

Finally, network-based applications can be classified by the availability of the related application tools (hardware and software). The possible classes (custom, semicustom, and off-the-shelf solutions) will be considered later in Section 13.

9. THE WORLD WIDE WEB

Next to e-mail, the World Wide Web is perhaps the most popular application of the Internet. It is a worldwide connected system of databases containing structured, hypertext, multimedia information that can be retrieved from any computer connected to the Internet.

9.1. History

The idea and basic concept of the World Wide Web was invented in Switzerland. In the late 1980s, two engineers, Tim Berners-Lee from the United Kingdom and Robert Cailliau from Belgium, working at CERN (the European Laboratory of Particle Physics) in Geneva, decided to develop a new communication system for Internet-based international collaboration for physicists. The idea became a reality in 1990, when the first version of HTTP (the Hypertext Transfer Protocol), HTML (the Hypertext Markup Language), and URL (the Universal Resource Locator) addressing scheme were introduced and implemented.

For the first three years of its life, the growth of the World Wide Web was quite slow. Only a few dozen web servers were set up, and the Web was used only for what it was originally invented for: scientific communication.

The growth rate of the Web began to accelerate significantly in the second half of 1993 after the introduction of the first portable, graphical web browser, X Mosaic. This was the first browser facilitating the new Internet application, the Web, to spread rapidly and penetrate all over the world. The Web has become a general-purpose technology, today serving not only scientific but all kinds of communication and information presentation needs.

The commercialization of the Internet also started with the birth of the World Wide Web. Companies not previously interested in using the Internet suddenly discovered the new possibilities and became the most active driving force in the development of networking. The first version of Netscape, another important web browser, was released in October 1994. The World Wide Web (W3) Consortium, formed in 1994, drives and leads the technological development and standardization processes of Web technology.

9.2. Main Features and Architecture

As mentioned above, the World Wide Web is a hypertext-based database system of multimedia content. It can be used for building either local or global information systems. It is interactive, which means that the information flow can be bidirectional (from the database to the user and from the user to the database). It integrates all the other traditional information system technologies, such as ftp, gopher (a predecessor of the WWW in organizing and displaying files on dedicated servers), and news.

The content of a Web database can be static or dynamic. Static information changes rarely (or perhaps never), while dynamic content can be generated directly preceding (or even parallel to) actual downloading.

The Web content is made available by web servers, and the information is retrieved through the use of web browsers. The more or less structured information is stored in web pages. These pages may refer to multimedia objects and other web pages, stored either locally or on other Web servers. Thus, the network of the Internet web servers forms a coherent, global information system. Nevertheless, because the different segments of this global system have been created and are operated and maintained independently, there is no unified structure and layout. This fact may result in considerable difficulties in navigating the Web.

The Universal Resource Locator is a general addressing scheme that allows unambiguous reference to any public object (or service), of any type, available on the Internet. It was originally invented for the Word Wide Web, but today URL addressing is used in many other Internet applications. Examples of different URLs include:

- http://www.bme.hu/index.html
- ftp://ftp.funet.fi/
- mailto:maray@fsz.bme.hu

9.3. HTTP and HTML

Hypertext Transfer Protocol (HTTP) is the application-level network protocol for transferring Web content between servers and clients. (Hypertext is a special database system of multimedia objects linked to each other.) HTTP uses the services of the TCP transport protocol. It transfers independently the objects (text, graphics, images, etc.), that build up a page. Together with each object, a special header is also passed. The header contains information about the type of the object, its size, its last modification time, and so on. HTTP also supports the use of proxy and cache servers. HTTP proxies are implemented in firewalls to pass HTTP traffic between the protected network and the Internet. Object caches are used to save bandwidth on overloaded lines and to speed up the retrieval process.

Hypertext Markup Language (HTML) is a document description language that was designed to create formatted, structured documents using hypertext technique (where links are associated to content pieces of the document so that these links can be easily applied as references to other related documents).

As originally conceived, HTML focused on the structure of the document. Later it was improved to provide more formatting features. In spite of this evolution, HTML is still quite simple. An HTML text contains tags that serve as instructions for structuring, formatting, linking, and so on. Input forms and questionnaires can also be coded in HTML. A simple example of an HTML text is as follows:

```
<HTML>
<HEAD>
<TITLE>Example</TITLE>
</HEAD>
<BODY>
<CENTER>
<H1>Example</H1>
</CENTER>
This is just a short <B>HTML</B> example, containing almost nothing.
<P>
<HR>
</BODY>
</HTML>
```

Details about the syntax can be found in the titles related to multimedia and the World Wide Web in the References. The above example will be displayed by the browser like this:

Example

This is just a short **HTML** example, containing almost nothing.

9.4. Multimedia Elements

The World Wide Web is sometimes called a multimedia information system because Web databases may contain also multimedia objects. It is obvious that in a well-designed and formatted document page graphic information and images can be used in addition to text. Web technology allows the inclusion of sound and moving pictures as well. Web pages may also contain embedded programs (e.g., Javascript, a language for designing interactive WWW sites) and thus can have some built-in intelligence.

10. THE ROLE OF THE WORLD WIDE WEB IN COMMUNICATION, INTEGRATION, AND COLLABORATION

The following subsections provide an insight into the role of the World Wide Web in information access, communication, exchanging information, collaboration, and integration.

10.1. The World Wide Web as a Means for Universal Information Access

In the past few years the World Wide Web has become an essential tool of common application in many areas of human life. Today it is a universal, global, and widely used information system that makes it possible for every user of the Internet to access a vast amount of information from every segment of human activity, including science, business, education, and entertainment.

The World Wide Web is universal because it can easily be adapted to almost any kind of information publishing needs. It can substitute for traditional paper-based publications like books, newspapers, newsletters, magazines, catalogs, and leaflets. But it is much more. It can be dynamic, multimedia based, and interactive.

The World Wide Web is global because using the transport provided by the Internet, it can be used worldwide. It allows any kind of information to flow freely to any destination, regardless of physical distance.

The available human interfaces (editors, browsers, etc.) of the Web are friendly and easy to use, making it open for everyone. The Web technology supports different national languages and character sets and thus helps to eliminate language barriers.

10.2. The World Wide Web as a Tool for Communicating and Exchanging Information

The basic role of the Web is to allow users to share (exchange) information. Accessing a particular set of web pages can be free to anybody or restricted only to a group of users.

However, the Web is also a technology for communication. There are Web-based applications for online and offline communication between users, for teleconferencing, for telephoning (Voice over IP [VoIP]), for using collective whiteboards, and so on.

Because the Web integrates almost all other Internet applications (including e-mail), it is an excellent tool for personal and group communication.

10.3. Collaboration and Integration Supported by the World Wide Web

An increasing amount of Web-based applications are available in the market for supporting collaborative and collective work (see the titles related to the World Wide Web and applications in the References). This evolution is based on the fact that the Web technology (including formats and protocols) is standardized, widespread, inexpensive, and platform independent.

Any network user can have access to the Web, regardless of the type of machine he or she works with or the way that machine is connected to the network. Thus, distributing information among collaborating partners is made easier by using the Web.

The Web can also be used for cooperative preparation of documents and/or software by collaborating network users, even if they are far away from each other. That is why Web technology is becoming increasingly popular in intranets and extranets as well. Companies and organizations can thus increase the efficiency of collective work within their staff by using Web-based tools.

Web technology can also be integrated in many different applications or products where there is a need for or a function of sharing or distributing information and/or communicating between different parties and users.

11. CONTENT GENERATION AND CONTENT PROVISION

Generation, provision, distant accessibility, and remote processing of information content can be considered the ultimate goal in introducing computer networks into an organization.

As many systems of information content can be postulated as there are organizations and applications within these organizations. Successful integration of the applications within an organization requires the integration of these separate sets of information content as well. Moreover, the organizations' outside relations can be made most effective if the information content within the organization is correlated to the relevant information content accessible in communicating with partner organizations.

As far as the structure of the correlated individual systems of information content is concerned, widely accepted distributed database techniques, and especially the worldwide-disseminated World Wide Web technology, provide a solid background. However, the structure and the above-mentioned techniques and technologies provide only the common framework for communicating and cooperating with regard to the related systems of information content. Determining what information and knowledge should be involved and how to build up the entire content is the real art in utilizing computer networks. There are no general recipes for this part of the task of introducing and integrating computer and network technologies into the activities of an organization.

The generation and accessibility (provision) of information content are briefly investigated in the following subsections. Also provided is a short overview of the classification of the information content, along with a subsection on some aspects of content rating and filtering.

11.1. Electronic I/O, Processing, Storage, and Retrieval of Multimedia Information

Generation and provision of information content involves a set of different steps that generally take place one after the other, but sometimes with some overlapping and possibly also some iterative refinement. In principle, these steps are independent of the type of the information. However, in practice, multimedia information requires the most demanding methods of handling because of its complexity. ("Multimedia" means that the information is a mix of data, text, graphics, audio, still video, and full video components.)

Obviously, before starting with the steps, some elementary questions should answered: what information is be to put into the stored content, how, and when is this information to be involved? These questions can be answered more or less independently of the computer network.

If the answers to these questions above are known, the next phase is to start with that part of the task related to the network itself. The basic chain of steps consists of the following elements:

1. Inputting the information electronically into the content-providing server
2. Processing the input in order to build up the content by using well-prepared data
3. Storing the appropriate information in the appropriate form
4. Providing "anytime accessibility" of the information

The first step is more or less straightforward using the available input devices (keyboard, scanner, microphone, camera, etc.)

The next steps (processing and storing) are a bit more complicated. However, this kind of processing should be built into the software belonging to the application system. This kind of software is independent of the information itself and normally is related exclusively to the computer network system being under consideration.

More important is how to find the necessary information when it is needed, whether the information is internal or external to the organization.

The task is relatively simple with internal information: the users within an organization know the system or at least have direct access to those who can supply the key to solving any problems in accessing it.

However, the task of looking for external information will not be possible without the most recent general tools. These tools include different kinds of distributed database handling techniques. More important, they include a number of efficient and convenient World Wide Web browsers, search engines, directories, general sites, portals, topical news servers, and custom information services. With these new ways of accessing the required information and a good infrastructure as a basis, there is virtually no search task that cannot be solved quickly and efficiently, provided that the requested information is recognizable.

11.2. Classification of Electronically Accessible Information Content

The more application areas there are (see Section 8 above), the more content classes can be identified. Thus, we can distinguish between:

- Public and private contents (depending on who can access them)
- Personal, administrative, and industrial contents (depending on where they were generated and who accesses them)
- Contents of varying complexity (depending on volume of information, structure of storage and accessibility, kind of information, distribution of storage, etc.)
- Contents of varying commercial availability (off-the-shelf, semi-custom, or custom generated).

These classifications determine the requirements for the contents that are under consideration. The main aspects of these requirements are:

- Availability (where and how the content can be reached)
- Accessibility (what authorization is needed for accessing the content)
- Reliability (what levels of reliability, validity, and completeness are required)
- Adequacy (what level of matching the content to the required information is expected)
- Updating (what the importance is of providing the most recent related information)
- Ease of use (what level of difficulty in accessing and using the information is allowed)
- Rating (what user groups are advised to, discouraged from, prevented from accessing the content)

Some of these aspects are closely related to the networking background (availability, accessibility), while others are more connected to the applications themselves. However, because accessing is always going on when the network is used, users may associate their possibly disappointing experiences with the network itself. Thus, reliability, adequacy, updating, ease of use, and rating (content qualification) are considered special aspects in generating and providing network accessible contents, especially because the vast majority of these contents are available nowadays by using the same techniques, from the same source, the World Wide Web. Therefore, special tools are available (and more are under development) for supporting the care taken of these special aspects.

Networking techniques can't help too much with content reliability. Here the source (provider or broker) is the most important factor in how the user may rely on the accessed content. The case is similar with adequacy and updating, but here specific issues are brought in by the browsers, search techniques/engines, general sites, portal servers (those responsible for supporting the search for adequate information), and cache servers (mirroring Web content by taking care of, for example, correct updating).

Some specific questions related to rating and filtering are dealt with in the next subsection.

11.3. Rating and Filtering in Content Generation, Provision, and Access

Although the amount of worldwide-accessible information content is increasing exponentially, the sophistication, coverage, and efficiency of the tools and services mentioned in the previous two subsections are increasing as well. However, one problem remains: the user can never exactly know the relevance, reliability, or validity of the accessed information. This problem is well known, and many efforts are being made to find ways to select, filter, and, if possible, rate the different available sources and sites (information providers) as well as the information itself. An additional issue should be mentioned: some information content may be hurtful or damaging (to minorities, for example) or corrupting, especially to children not well prepared to interpret and deal with such content when they accidentally (or consciously) access it.

There is no good, easily applicable method available today or anticipated in the near future for solving the problem. The only option at the moment is to use flags as elementary information for establishing a certain level of rating and filtering. This way, special labels can be associated with content segments so that testing the associated labels before accessing the target content can provide information about the content, itself. These flags can provide important facts about the topics, the value, the depth, the age, and so on, of the related content and about the target visitors of the site or target audience of the content.

The flags can be associated with the content by the content provider, the content broker, the service provider, or even the visitors to the related sites. However, because appropriate flags can be determined only if the content is known, the most feasible way is to rely on the content provider. The problem is that if the provider consciously and intentionally wishes to hide the truth about the content or even to mislead the potential visitors/readers, there is practically no way of preventing such behavior. This is a real problem in the case of the above-mentioned hurtful, damaging, or corrupting contents, but also with any other contents as well, as far as their characteristics (validity, value, adequacy, age, etc.) are concerned.

If flags are present, filtering is not a difficult task. It can be performed either "manually" (by a human decision whether to access or not access) or even by an automated process, inhibiting the access if necessary or advisable.

However, the above procedure assumes standardized techniques and standard labeling principles/rules, as well as fair usage. There is still a lot of work to do before these conditions will generally be met (including codification).

More has to be done with respect to future intelligent machine techniques that could automatically solve the tasks of labeling and filtering. Solving this problem is extremely complex but also extremely important and urgent because of the exponentially increasing amount of content accessible through the Internet.

12. TRENDS AND PERSPECTIVES

Computer networking is a continuously developing technology. With the rapid evolution of the theory and techniques behind them, the worldwide infrastructure, the globally available services, the widening spectrum of applications, the exponentially growing amount of accessible content, and the ability and readiness of hundreds of millions of users are all going through an intensive developmental process. With the progression of the networks (sometimes simply called the Internet), together with the similarly rapid evolution of the related computing, communications, control, and media technologies and their applications, we have reached the first phase of a new historical period, the Information Society. The evolving technologies and applications penetrate every segment of our life, resulting in a major change in how people and organizations, from small communities to global corporations, live their lives and manage their activities and operations.

Behind the outlined process is the development of computer networking in the last 30 years. Global connectivity of cooperating partners, unbounded access to tremendous amounts of information, as well as virtually infinite possibilities of integrating worldwide distributed tools, knowledge, resources, and even human capabilities and talents are becoming a reality as we go rapidly down a road that we have only just started to pave. Although the future of networking is impossible to see today, some important facts can already be recognized.

The most important attributes of the evolution in computer networking are openness, flexibility, and interoperability:

- Openness, in allowing any new technologies, new solutions, new elements in the global, regional, and local networks to be integrated into the existing and continuously developing system of infrastructure, services, and applications
- Flexibility, in being ready to accept any new idea or invention, even if unforeseen, that takes global advancement a step ahead
- Interoperability, in allowing the involvement of any appropriate tool or device so that it can work together with the existing system previously in operation

These attributes stem from:

- The more conscious care taken of the hierarchical and/or distributed architecture in networking technology
- The cautious standardization processes going on worldwide
- The carefulness in taking into consideration the requirements for interfacing between evolving new tools and devices

As a result, the progression in networking continues unbroken and the proliferation of network usage unsaturated, even at the start of the third millennium. Some trends in this development process are briefly outlined in the following subsections.

12.1. Network Infrastructure

Major elements of the trends in the development of the network infrastructure can be summarized, without aiming at an exhaustive listing, as follows:

- Transmission speeds are increasing rapidly. The trend is characterized by the introduction of more fiberoptic cables; the application of wavelength division multiplexing and multiple wavelengths in the transmission technique (utilizing the fact that a single fiber can guide several different frequency waves in parallel and sharing these waves among the information coming from different sources and being transmitted to different destinations); the integration of both guided and radiated waves in data communication by combining terrestrial and satellite technologies in radio transmission; the introduction of new techniques for mobile communication; the application of a more structured hierarchy in the infrastructure; and more care being taken of "last-mile" connectivity (to the access speeds at the final sections of the connections, close to the individual PCs or workstations), so that multiple-Mbps cable modems and digital subscriber lines, or cable TV lines, direct satellite connections, or wireless local loops, are applied close to the user terminals.
- Active devices of the infrastructure, together with the end user hardware and software tools and devices, are improving continually, allowing more intelligence in the network.

- The number, capability, and knowhow of the public network and telecom operators is increasing, so market-oriented competition is continuously evolving. The result is lower prices, which allows positive feedback on the wider dissemination of worldwide network usage.
- Governments and international organizations are supporting intensive development programs to help and even motivate fast development. Tools applied include financial support of leading-edge research and technological development, introduction of special rules and regulations about duties and taxes related to information technologies, and, last but not least, accelerating codification on Internet-related issues.

As a result of these elements, the network infrastructure is developing rapidly, allowing similarly fast development in the service sector.

12.2. Services

The trends in the evolution of networking services, available mostly through Internet service providers (ISPs) and covering domain registration and e-mail and Web access provision, as well as further service types related to providing connectivity and accessibility, can be characterized by continuing differentiation.

Although the kinds of services don't widen considerably, the number of users needing the basic services is growing extremely fast, and the needs themselves are becoming more fragmented. This is because the use of the network itself is becoming differentiated, covering more different types of applications (see Section 12.3) characterized by different levels of bandwidth and quality demand.

Quality of service (QoS) is the crucial issue in many cases. From special (e.g., transmission-intensive critical scientific) applications to real-time high-quality multimedia videoconferencing to bandwidth-demanding entertainment applications and to less demanding e-mail traffic, the different quality-level requirements of the service result in multiple grades of QoS, from best-efforts IP to guaranteed transmission capacity services.

Different technologies, from asynchronous transfer mode (ATM) managed bandwidth services (MBS) to new developments in Internet Protocol (IP) level quality management, support the different needs with respect to QoS grades. Lower quality demand may result in extremely low prices of service.

An important trend is that access to dark fiber (optical cables as physical infrastructure elements rather than leased bandwidths or managed connectivity services) and/or just separate wavelengths (specific segments of the physical fiber capacity) is becoming available and increasingly popular among well-prepared users, besides the traditional, mainly leased-line, telecom services. While this new element in the service market assumes higher levels of technological knowhow from the user, the cost aspects may make this possibility much more attractive in some cases than buying the more widespread traditional services. This also means an important new kind of fragmentation in the market, based on buying fiber or wavelength access and selling managed transmission capacity.

Another trend is also evolving with respect to services, namely in the field of providing and accessing content on the Web. It is foreseeable that content provision services such as hierarchical intelligent caching/mirroring, as well as content-accessing services such as operating general and topical portals or general and topical sites, will proliferate rapidly in the future, together with the above-mentioned trend toward separation of content provision and content brokerage.

These trends in networking services will result in a more attractive background for the increasing amount of new applications.

12.3. Applications

An overview of the development trends in network infrastructure and networking services is not difficult to provide. In contrast, summarizing similar development trends in the far less homogeneous applications is almost impossible because of their very wide spectrum. However, some general trends can be recognized here.

First, the areas of applications do widen. Just to mention some key potential trends:

- The most intensive development is foreseeable in the commercial and business applications (e-commerce, e-business).
- Industrial applications (including the production and service industries) are also emerging rapidly; one of the main issues is teleworking, with its economic as well as social effects (*teleworking* means that an increasing number of companies are allowing or requesting part of their staff to work for them either at home or through teamwork, with team members communicating through the network and thus performing their joint activities at diverse, distant sites).

- Applications in the entertainment sector (including home entertainment) are also becoming more common, but here the trends in the prices of the service industry have a strong influence.
- Applications in governmental administration and the field of wide access to publicly available administrative information seem to be progressing more slowly at the start of the new millennium. The reasons are probably partly financial, partly political (lack of market drive and competition results in lack of elevated intensity in the development of the networking applications too).
- Applications related to science and art will also evolve rapidly, but the speed will be less impressive than in the commercial area, although internal interest and motivation complement specific market drive elements here.
- Some special factors are playing an important role in the fields of telemedicine and teleteaching (teleinstruction). Here, the healthy mix of market-driven competitive environments, governmental commitment, and wide public interest will probably result in an extremely fast evolution.

Second, some general features of the applications are showing general trends as well, independently of the application fields themselves. Two of these trends are:

- Multimedia transmission is gradually entering practically all applications, together with exploitation of the above-mentioned techniques of telepresence and teleimmersion (these techniques are characterized by combining distant access with virtual reality and augmented reality concepts; see Section 2).
- The concept of grids will probably move into a wide range of application areas. In contrast to the well-known client–server scheme, this concept exploits the potential of integrating distributed intelligence (processing capability), distributed knowledge (stored information), and distributed resources (computing power) so that a special set of tools (called middleware) supports the negotiation about and utilization of the intelligence, knowledge, and resource elements by goal-oriented integration of them within a grid. (A grid is a specific infrastructure consisting of a mutually accessible and cooperative set of the joined hosting sites, together with the distributed intelligence, knowledge, and resources and the middleware tools.)

The development trends in the field of the applications obviously deeply influence the trends in the area of content generation and content provision.

12.4. Information Content

Forecasting in the area of content generation and content provision is even more difficult than in the field of network applications. Only some basic issues can be mentioned here.

The first important trend is the emergence of a new branch of industry, the content industry. No longer are content generation and content provision secondary to building the infrastructure, providing services, and/or introducing or keeping applications. The related complex tasks (at least in the case of professional activities of a high standard) require large investments and call for adequate return on these investments. Well-organized joint efforts of highly educated and talented staff are required in such demanding activities.

The second trend is directly linked to the amount of work and money to be invested in content generation. Contents will gain more and more value in accordance with such investments. Moreover, while accessing the infrastructure and hiring services will probably become less expensive in the future, the price for accessing valuable information will probably grow considerably. The price for accessing any information (content) will soon be cost based, depending on the size of the investment and the potential number of paying users.

The third trend is also due to the increasing complexity of the tasks related to content generation and content provision. Task complexities have initiated the separation of the activities related to generating contents from those related to organizing services for content provision to the public. This means that special expertise can be achieved more easily in both of these types of demanding activities.

The fourth trend is related to storing the information. Future contents will be distributed: because transmission will get cheaper and cheaper, merging distant slices of information content will be made much cheaper by maintaining the current distance and accessing the required remote sites as needed instead of transporting these slices into a central site and integrating them on-site. However, in order to exploit this possibility, a certain level of distributed intelligence is also required, as well as inexpensive access to the connectivity services. This trend is closely related to those regarding portals and grids discussed in the previous subsection.

XML is an emerging protocol of note. Applied for exchange of structured information, it is increasingly used in e-commerce applications.

13. NETWORKING IN PRACTICE

The following subsections go into some detail about the practical issues of networking. A brief overview is provided of the suggested way of starting activities devoted to advance analysis of the environment, planning the applications and their prerequisites, designing the network itself, and implementing the appropriate hardware and software. Classification of networking solutions and implementation of network technology are investigated separately. As in the other parts of this chapter, the reader should look for a more detailed discussion in the specialized literature, some of which is listed in the References.

13.1. Methodology

Designing a good network is a complex task. It is essential that the planning and implementation phase be preceded by a thorough survey in which the following questions, at least, have to be answered:

- What is the principal goal of building the network?
- What is the appropriate type and category of the network?
- Is there a need for integration with existing networks?
- What kinds of computers and other hardware equipment are to be connected?
- What kinds of applications are to be implemented over the network?
- What kind of protocols should be supported?
- How much traffic should be carried by the network? What is the required bandwidth and throughput of the network? What is the highest expected peak load of the network?
- What is the maximum acceptable delay (latency) on the network?
- What level of reliability should be guaranteed?
- What level of security has to be reached?
- What kinds of enhancements and improvements are expected in the future? What level of scalability is desirable?
- What kinds of physical constraints (distance, trace, etc.) should be taken into consideration?

The actual process of implementation depends on the answers to these questions. The result (the adequacy and quality of the network) will greatly depend on how consistently the rules have been followed about building networks.

Designing the most appropriate topology is one of the key issues. Many other things depend on the actual topology. If the topology is not designed carefully, the network may be far from having optimum adequacy, quality, and cost/performance.

Choosing the right physical solutions and low-level protocols is also very important. Low-level protocols (TCP/IP, etc.) have a strong impact on the functionality of the network. Building a multiprotocol network is advisable only if there is an explicit need for such a selection; this may increase the costs and make the management and maintenance tasks difficult and complicated.

13.2. Classification

Classification of networks by type may help significantly in selecting the solution for a specific application. Obviously, different types of networks are designed to serve different types of application needs.

Networks can be classified by a number of attributes. Important classification factors include:

- Purpose (general, industrial, process control, etc.)
- Size (local area network, metropolitan area network, wide area network)
- Technology (ethernet, fast ethernet, token ring, fiber digital data interface [FDDI], asynchronous transfer mode [ATM], etc.)
- Applied low-level protocols (TCP/IP, IPX/SPX, Decnet, AppleTalk, etc.)
- Speed (low, medium, high, ultrahigh)
- Mode of operation (open, private, intranet, extranet, etc.)

Often more than one kind of technology and/or protocol has to be used in the same network. Especially if the size of the network is large, different segments can be built up by using different technologies and/or protocols. Fortunately, modern network equipment allows the use of different technologies and protocols at the same time within the same network. This equipment takes care of the necessary translations when routing the information traffic.

13.3. Implementation Issues

Like the design process, implementation is a complex task that can involve difficult problems. The design team should understand the available technological solutions well in order to select the most appropriate ones for each part of the system.

As an example, ethernet technology is a widespread, efficiently functioning solution in most cases, but because it does not guarantee appropriate response times, it cannot be used in certain real-time environments. Ethernet is a reliable, effective, inexpensive, and very good technology, but the applied CSMA/CD algorithm is not real time. In practice, the process to be controlled by the network may be much slower than the ethernet technology. Therefore, especially if oversized network capacity is implemented, there will be no problem in such cases. But theoretically it is not the best choice. Of course, the choice and the means of implementation are also a financial question, and all the important factors should be taken into consideration in looking for a good compromise.

Decisions about what and how to implement have a serious impact on the reliability and security of the network, too. For building a high-reliability, high-availability network, redundant components must be used. Communication lines and passive and active network components all have to be multiplied in a well-designed way to get a partly or fully fault-tolerant system. Of course, this may result in significant extra costs. Similarly, if extra-high levels of security must be guaranteed, a number of additional hardware and software components may be needed. Again, a good compromise can be found by taking into account all the reliability and security aspects and cost considerations.

14. NETWORKING IN THE PRODUCTION AND SERVICE INDUSTRIES

The application of computer networks in the production and service industries should be fitted to the required complexity and coverage.

As far as complexity of network usage, two main aspects should be taken into consideration:

1. What kind of infrastructure is built within the organization (including internal structure and technology as well as connectivity towards the outside world)? The spectrum goes from connecting a few workplaces (PCs) to each other, and possibly to the Internet, by the simplest LAN techniques and telephone modems, respectively, to using the most complex high-speed intranet and/or extranet applications with high-end workstations, or even large computer centers, together with broadband connections to the global Internet.
2. What kinds of services are applied? In this respect, the simplest solutions make possible only the simple exchange of messages between workplaces, while the other end is characterized by the exchange of the most complex multimedia information by using World Wide Web techniques, accessing large databases, applying distributed information processing, utilizing virtual environments, exploiting network-based long-distance collaboration, and so on.

Coverage of network usage also involves two important aspects:

1. What kinds of applications within the organization are introduced? The low end is simple word processing with possible electronic document handling. The other extreme is characterized by a complex system of network-based planning, computer-based distributed decision making, network-based integrated management, computer aided design, manufacturing, and testing, distributed financing by the use of computer networking, computerized human resource management, networked advertising, promotion, marketing, retailing, online sales transactions, public relations, and so on.
2. What amount and complexity of information content are generated, processed, and stored by the networked system of the organization? Here the levels are very much company specific, but it can be said that companies in both the production and the service industries can start with applications using elementary sets of information and may get ahead until virtually the entire information and knowledge base about the full spectrum of activities within the company (and also about the related outside world) are appropriately stored and processed by a well-established system of networked services and applications.

The basic principle here is that introducing computers and computer networks amplifies the strengths as well as the weaknesses of the organization. Well-organized companies can gain a lot from computerized and networked applications. However, those with considerable organizational problems mustn't look for a miracle: they will rapidly realize that these tools increase rather than decrease the problems of operating the related production or service activities.

The full process (from the first elementary steps until completing the implementation of the network-based system) is a combination of two basic components:

1. Introducing and integrating network technology (infrastructure and services) inside the company
2. Introducing and integrating networked applications into the activities within the company (together with continuously building up the information content).

If careful introduction and integration of the new technologies and methods is performed in a systematic manner and by appropriately restructuring and re-forming all the related activities within the company, the results will be higher efficiency, better performance, and lower costs, provided that good management exploits the opportunities that are made available.

Moreover, properly introduced and appropriately applied computer techniques and network technologies will not only help in getting ahead with enhancing efficiency and performance as well as cost cutting, but also result in elevated competitiveness. This way, survival against the increasing worldwide competition is supported, too.

15. SOME PRACTICAL ASPECTS OF INTRODUCING AND USING COMPUTER NETWORKS IN INDUSTRIAL ENGINEERING

As mentioned above, applying up-to-date computer networks in the production and service industries requires two basic steps:

1. Integrating network technology (Internet connectivity and accessibility) in the related organization or company
2. Integrating networking services and network-based applications (Internet technology) in the operations and activities of the organization or company.

The most important practical issues with regard to these steps are briefly investigated in the following subsections but some preliminary comments should be made.

First, there are, in principle, two distinct possibilities, based on the top-down and the bottom-up approach, respectively:

- In the case of fast top-down introduction of network technology, an advance analysis (feasibility study) should be performed, in order to avoid the potential risks associated with lack of careful preparation for taking a giant step ahead with networking in the company.
- In the case of applying the bottom-up approach, the speed of ''getting networked'' company-wide is much lower, but getting ahead step-by-step makes it possible to correct any mistakes and adapt to all the recognized internal circumstances so that the risks are minimized.

In any case, the top-down approach usually requires outside guidance from a professional company that specializes in introducing and integrating network technology in the organization and activities of industrial companies. This is normally not required with the bottom-up approach, although preliminary analysis and outside advice may help a lot there, too.

Another issue is coverage and depth in introducing networked information technology. The basic questions are where to stop with the bottom-up approach and what goals to define with the top-down approach. Coverage here means organizational (what departments to involve) and geographic (which sites to connect to the corporate network) coverage, while depth relates to the level at which network technology is introduced (which tasks to involve and what level of completion to achieve by network integration).

A further issue is to determine whether buying an off-the-shelf solution, buying a semicustom solution and adapting it to the local circumstances, or starting in-house development of a full custom solution is preferable. For small enterprises, the first alternative will be optimum; for high-end corporations, in many cases, the third. Although today the second alternative is the most common way of integrating network technology in a company, each organization should be considered separately, requiring its own specific handling.

15.1. Internet Connectivity

The introduction of Internet technology in industrial engineering starts with a preliminary definition of the company's business requirements and objectives. From these, the potential network applications can be derived. The planning and implementation phases can be started.

The first steps in the process of establishing Internet connectivity at an industrial company are to define the specifications and launch the necessary procedures:

- Consulting with experts about possibilities and goals
- Identifying what Internet applications will be introduced
- Defining human resource needs and deciding on staff issues related to networking
- Contracting an advisor/consultant to support activities related to company networking
- Estimating what bandwidth the selected applications will require
- Determining what equipment (hardware and software devices and tools) will be necessary
- Deciding about security issues and the extra equipment required
- Selecting the most appropriate Internet service provider (which may be the company itself)
- Deciding about the domain name(s) to be applied by the company
- Negotiating with the ISP about services to be bought (provided that the company is not its own ISP)
- Starting purchasing the necessary equipment

Naturally, the above steps do not follow each other in linear order. An iterative–interactive process is to be assumed, in which earlier decisions may sometimes be changed because of later recognition of specific problems, barriers, or difficulties.

The next steps should be devoted to preparations inside the company (or company sites). This means not just planning and building the in-house infrastructure (cabling, server room, etc.), but also deciding about the internal server structure (domain name server, mail server, web server, etc.). Another issue is starting the purchasing of the components (HW and SW) of the internal infrastructure and services that will be needed for covering the target applications. A further task is to prepare the tools and devices for the applications themselves.

Ideally, the full process may only take several weeks, but it may last several months (e.g., in the case of a complex infrastructure covering several sites, each bringing its own complex networking tasks into the picture).

Behind some of the above decisions and steps are several opportunities and choices (bandwidth of connectivity, level of security, server structure, equipment brands, etc.). Some also involve selecting among different technology variants. These alternatives require careful analysis before decisions are made.

It is not surprising that the Internet connection is to be upgraded later. This is normally not a problem at all and can be performed relatively simply, even if the upgrade also means a technological change.

15.2. LANs, Intranets, WANs, and Extranets

Establishing Internet connectivity is just the first step in the process. It makes communication with the outside world possible, but it doesn't yet solve the task of establishing internal connections within the enterprise.

In the case of a small enterprise and a single company site, the only additional task in establishing connectivity is to implement interconnections between the PCs and workstations used by the company staff. This is done by building the LAN (local area network) of the organization so that the resources (servers as well as PCs and workstations) are connected to the in-house infrastructure, in most cases an ethernet network, of the company. The required speed of the internal network depends on the traffic estimates. New ethernet solutions allow even gigabit per second transmission.

Once the internal LAN is established, the next question is how the users within the LAN will communicate with their partners outside the corporate LAN. Here is where the company intranet enters the picture.

The intranet is a network utilizing the TCP/IP protocols that are the basis of the Internet but belonging exclusively to the company and accessible only to the company staff. This means that outside partners can access the machines within the intranet only if they have appropriate authorization to do so. The intranet is in most cases connected to the global Internet through a firewall, so that although the websites within the intranet look and behave just like other websites outside the intranet, the firewall in front of the Intranet prevents unauthorized access.

In companies with separate, distant sites, all these sites may have their own LANs, firewalls, and intranets. In most cases these are connected to each other through the public MAN (high-speed metropolitan area network of a town or city), or sometimes through the regional WAN (extra-high-speed wide area network of a large region), although private network connections can also be applied in case of specific security or traffic requirements.

However, intranets connected through public MANs or WANs may also be interconnected so that, in spite of the geographically scattered locations of the related company sites, they behave as a single intranet. A common solution is to establish extranets so that the separate LANs are connected by a virtual private network over the public MAN or WAN and authorized access is controlled by a set of firewalls taking care of intranet-like traffic between the authorized users (company staff and authorized outside partners). Thus, extranets may also be considered as distributed Intranets accessible not only to company employees, but partially accessible to authorized users outside the organization or company.

LANs, intranets, and extranets provide a means for organizations or companies to utilize best the possibilities stemming from integrating the Internet in their operations and activities without losing security.

Although LANs, intranets and extranets are quite scalable, it is good to think ahead when planning the infrastructure and services of the company so that extension and upgrading do not occur more frequently than really necessary. The network address schema is impacted by the infrastructure, so changing network numbering can be costly.

15.3. World Wide Web Communication, Integration, and Collaboration

Once the LANs are in place, with their servers, PCs, and workstations, and the intranet/extranet infrastructure is working well, by taking care of secure interconnectivity towards the outside world, and by utilizing appropriate routing and firewall equipment, the details of the applications come into the picture. Here the best solutions can be built up by exploiting the possibilities that modern World Wide Web systems provide.

The Web is the global system of specialized Internet servers (computers delivering web pages to machines asking for them). The Web is also a tool integrating:

- Worldwide addressing (and thus accessibility)
- Hypertext technique (allowing navigation through a multitude of otherwise separate websites by following the links associated to web page content pieces)
- Multimedia capabilities (covering a wide spectrum of formats used by the Web contents, from structured text to graphics and from audio to still and real video)

Interacting with the Web content by initiating content-oriented actions, including also customizing them, is thus supported, too.

The Web is an ideal tool not only for worldwide information access, but also for communication, integration, and collaboration inside an enterprise and among enterprises (eXtensible Markup Language [XML]). The key is the combination of Web technology with the intranet/extranet environment. Practically, this means operating several web servers within the enterprise so that some of them serve only internal purposes and thus do not allow outside access, while others support communication outside the enterprise boundaries. While the latter normally differ from any other web servers only in their special property of being able to take over public content from the internally accessible web servers, the structure, linking, and content of the former are constructed so that they directly serve internal communication, integration, and collaboration.

This type of communication differs from general e-mail usage only in providing a common interface and a well-organized archive of what has been communicated through the system. This makes it possible to build the archive of the company continuously by automatically storing all relevant and recorded documents for easy, fast, and secure retrieval. The management of the organization or company can thus be considerably improved.

More important is that such a (Web-based) store of all relevant documents also take care of integrating all the activities resulting in or stemming from those documents. If a document starts its life in the system, any access to it, modification of it, attachment to it—in general, any event related to it—is uniquely recorded by the system. Thus, a reliable and efficient way of integrating the activities within the company, as well as between the company and its partners, can be achieved. Typical examples are e-commerce and e-business.

Communication and integration within an enterprise by the use of computer networks also mean high-quality support for overall collaboration, independent of where, when, and how the collaborating staff members take part in the joint activities. The only prerequisite is that they all take into consideration the rules of how to access the company websites. (Nevertheless, efficiently functioning systems are aware of these rules, too.) Different departments, from design to manufacturing, from sales to marketing, and from service provision to overall logistics, can collaborate in this way so that the company can maintain reliable and efficient operation.

However, as mentioned above, disorder in the operations of a company is amplified when World Wide Web communication, integration, and collaboration are introduced. Thus, in order to achieve a really reliable and efficient system of operations, careful introduction and well-regulated usage of the network-based tools are a must.

All this would be impossible without computer networking. And this is how industrial engineering can benefit fully from the computer network infrastructure and services within an industrial organization or company.

16. SUMMARY

Computer networking in industrial engineering is an invaluable tool for establishing and maintaining reliable and efficient operation and thus competitiveness. By utilizing the possibilities provided by the Internet and the World Wide Web, industrial enterprises can achieve results that earlier were impossible even to imagine.

However, the present level of networking technology is just the starting phase of an intensive global evolution. The development will not be stopping, or even slowing down, in the foreseeable future. New methods and tools will continue to penetrate into all kinds of human activities, from private entertainment to business life, from government to administrations to industrial engineering.

Getting involved in the mainstream of exploiting the opportunities, by getting connected to the network and introducing the new services and applications based on the Net, means being prepared for the coming developments in computer networking. However, missing these chances means not only losing present potential benefits, but also lacking the ability to join the future network-based innovation processes.

In 1999, the worldwide Internet Society announced its slogan: "The Internet is for everyone." Why not for those of us in the industrial engineering community?

REFERENCES

Ambegaonkar, P., Ed. (1997), *Intranet Resource Kit,* Osborne/McGraw-Hill, New York.

Agnew, P. W., and Kellerman, S. A. (1996), *Distributed Multimedia: Technologies, Applications, and Opportunities in the Digital Information Industry: A Guide for Users and Practitioners,* Addison-Wesley, Reading, MA.

Angell, D., and Heslop, B. (1995), *The Internet Business Companion: Growing Your Business in the Electronic Age,* Addison-Wesley, Reading, MA.

Baker, R. H. (1997), *Extranets: The Complete Sourcebook,* McGraw-Hill, New York.

Bernard, R. (1997), *The Corporate Intranet,* John Wiley & Sons, New York.

Bort, J., and Felix, B. (1997), *Building an Extranet: Connect Your Intranet with Vendors and Customers,* John Wiley & Sons, New York.

Buford, J. F. K. (1994), *Multimedia Systems,* Addison-Wesley, Reading, MA.

Cheswick, W., and Bellovin, S. (1994), *Firewalls and Internet Security,* Addison-Wesley, Reading, MA.

Conner-Sax, K., and Krol, E. (1999), *The Whole Internet: The Next Generation,* O'Reilly & Associates, Sebastopol, CA.

Croll, A. A., and Packman, E. (1999), *Managing Bandwidth: Deploying QoS in Enterprise Networks,* Prentice Hall, Upper Saddle River, NJ.

Derfler, F. (1998), *Using Networks,* Macmillan, New York.

Derfler, F., and Freed, L. (1998), *How Networks Work,* Macmillan, New York.

Dowd, K. (1996), *Getting Connected,* O'Reilly & Associates, Sebastopol, CA.

Foster, I., and Kesselman, K., Eds. (1998), *The Grid: Blueprint for a New Computing Infrastructure,* Morgan Kaufmann, San Francisco.

Goncalves, M., Ed. (1999), *Firewalls: A Complete Guide,* Computing McGraw-Hill, New York.

Guengerich, S., Graham, D., Miller, M., and McDonald, S. (1996), *Building the Corporate Intranet,* John Wiley & Sons, New York.

Hallberg, B. (1999), *Networking: A Beginner's Guide,* McGraw-Hill, New York.

Hannam, R. (1997), *Computer Integrated Manufacturing: From Concepts to Realization,* Addison-Wesley, Reading, MA.

Hills, M. (1996), *Intranet as Groupware,* John Wiley & Sons, New York.

Huitema, C. (1998), *Ipv6, the New Internet Protocol,* Prentice Hall, Upper Saddle River, NJ.

Kalakota, R., and Whinston, A. B. (1997), *Electronic Commerce: A Manager's Guide,* Addison-Wesley, Reading, MA.

Keshav, S. (1997), *An Engineering Approach to Computer Networking,* Addison-Wesley, Reading, MA.

Kosiur, D. (1998), *Virtual Private Networks,* John Wiley & Sons, New York.

Lynch, D. C., and Rose, M. T. (1993), *Internet System Handbook,* Addison-Wesley, Reading, MA.

Mambretti, J., and Schmidt, A. (1999), *Next Generation Internet,* John Wiley & Sons, New York.

Marcus, J. S. (1999), *Designing Wide Area Networks and Internetworks: A Practical Guide,* Addison-Wesley, Reading, MA.

McMahon, C., and Browne, J. (1998), *CAD/CAM: Principles, Practice and Manufacturing Management,* Addison-Wesley, Reading, MA.

Minoli, D. (1996), *Internet and Intranet Engineering,* McGraw-Hill, New York.

Minoli, D., and Schmidt, A. (1999), *Internet Architectures,* John Wiley & Sons, New York.

Murhammer, M. W., Bourne, T. A., Gaidosch, T., Kunzinger, C., Rademacher, L., and Weinfurter, A. (1999), *Guide to Virtual Private Networks,* Prentice Hall, Upper Saddle River, NJ.

Ptak, R. L., Morgenthal, J. P., and Forge, S. (1998), *Manager's Guide to Distributed Environments,* John Wiley & Sons, New York.

Quercia, V. (1997), *Internet in a Nutshell,* O'Reilly & Associates, Sebastopol, CA.

Salus, P. H. (1995), *Casting the Net: From ARPANET to INTERNET and Beyond . . . ,* Addison-Wesley, Reading, MA.

Schulman, M. A., and Smith, R. (1997), *The Internet Strategic Plan: A Step-by-Step Guide to Connecting Your Company,* John Wiley & Sons, New York.

Singhal, S., and Zyda, M. (1999), *Networked Virtual Environments: Design and Implementation,* Addison-Wesley, Reading, MA.

Smith, R. E. (1997), *Internet Cryptography,* Addison-Wesley, Reading, MA.

Smythe, C. (1995), *Internetworking: Designing the Right Architectures,* Addison-Wesley, Reading, MA.

Stout, R. (1999), *The World Wide Web Complete Reference,* Osborne/McGraw-Hill, New York.

Taylor, E. (1999), *Networking Handbook,* McGraw-Hill, New York.

Treese, G. W., and Stewart, L. C. (1998), *Designing Systems for Internet Commerce,* Addison-Wesley, Reading, MA.

Ward, A. F. (1999), *Connecting to the Internet: A Practical Guide about LAN-Internet Connectivity, Computer Networks and Open Systems,* Addison-Wesley, Reading, MA.

Wesel, E. K. (1998), *Wireless Multimedia Communications: Networking Video, Voice and Data,* Addison-Wesley, Reading, MA.

Young, M. L. (1999), *Internet Complete Reference,* Osborne/McGraw-Hill, New York.