NEAR EAST UNIVERSITY

Faculty of Engineering

Department of Computer Engineering

INTERNETWORKING WITH TCP/IP

Graduation Project
COM- 400

Student: Hakan Gönülay (990951)

Supervisor: Assist.Prof.Firudin Muradov

Lefkoşa - 2003

ACKNOWLEDGEMENTS

First I want to thank Assist.Prof.Dr.Firudin Muradov to be my advisor. Under his guidence, I succesfully overcome many difficulties and learn a lot about Internetworking with TCP/IP. In each discussion, he explained my questions patiently, and I felt my quick progress from his advices. He always helps me a lot either in my study or my life. I asked him many questions in my subject and he always answered my questions quickly and in detail.

Special thanks to the Mr.Halil Adahan for his practical advices. And thanks to Faculty of Engineering for having such a good computational and electronical environment.

I also want to thank to my friends in Near East University: Sercan, Murat, Deniz and Cihan. Being with them make my 4 years in NEU full of fun.

Finally, I want to thank my family, especially my parents. Without their endless support and love for me, I would never achieve my current position. I wish my mother and fother lives happily always.

ABSTRACT

The standards collectively known as TCP/IP, first developed to allow exchange of information between computers in the US government, defence and university research communities, appeared in attractive, useful and widely available products for general commercial computing.

With the increasing interest in the use of TCP/IP for general commercial applications, there is a need to know what management and technical difficulties will be encountered.

This project is about the practical problems of installing, configuring and maintaining information systems based on the TCP/IP set of standards, from initial installation to on-going maintenance.

For a system to be successful and to retain that success over a long time, requires frequent revision on initial assumptions; system designers must take account not only of technical, but also of social and organizational problems they will encounter. Successful systems grow and develop; they are used in ways and for purposes that the initial design probably did not predict. Once convenient and reliable operation is achieved, the users abandon and then lose the older, less convenient alternatives. The new system becomes part of the corporate infrastructure and day-to-day life, its value increases and any change in performance and availability can dramatically affect prosperity and well-being.

TABLE OF CONTENTS

INTRODUCTION

The increasing interest in the use of Internet for commercial applications requires to know what management and technical difficulties will be encountered. This project describes the practical problems of installing, configuring and maintaining information systems based on the TCP/IP.

We have introduced the components of TCP/IP which support interconnectivity and interoperability, and discussed how technology affects and is affected by the size, growth rate existing structures, culture and geography of the organization that deploys it.

The layered structure of the TCP/IP protocol suite suggests how we might divide up the management responsibilities, but it must be remembered that the technology must be treated as one integrated system if it is to operate successfully. Changes in one area cannot be made without considering the consequences for performance and costs in another. If long-term success and stability is to be achieved with minimum effort, key managerial decisions must be made on specific technical parameters.

In this project, there are six chapters, that each one is complementary for the previous ones.

Chapter 1 introduces the components of TCP/IP and the need for internetworking. Also in this chapter the reader can see Upper layer, Lower layer protocols and management issues.

Chapter 2 presents the factors that allow the LANs and WANs which support TCP/IP to be planned successfully. The planning process is followed by determining the organization, its size, structure and communications flows

Chapter 3 is fully reserved for Planning and managing IP addresses. Different ways of addressing computers in a TCP/IP system, functions and format of the IP addresses, the need for a unique IP address and IP address registration are the topics discussed in this chapter.

Chapter 4 includes the topics which are the need for subnetwork addressing, the structure and format of the subnetwork mask and supernetworks.

Chapter 5 contains Routing management and configurations for routers.

Finally in chapter 6 TCP/IP upper layers, transport and applications services are described containing UDP, TCP, Winsock, The X Window system, Telnet, FTP, TFTP, SMTP and World Wide Web.

CHAPTER ONE: INTRODUCING TCP/IP

1.1. The need for internetworking

It is some 13 years since the authors first began to work together on communications networks in a large commercial environment. In that early part of our careers, which of course predated readily available sophisticated systems like TCP/IP, we were first involved in the design and implementation of a large corporate terminal-based information network. The aims of corporate network managers were expressed in three memorable phrases that appeared in many management presentations of the time:

- A single terminal on a desk
- One terminal per seated employee
- Total logical interconnectivity

The requirements were simple, if somewhat shrouded in the jargon of the day: employees should be able to access any information or system in the corporation (for which they had authority) from a single terminal on their own desk - they would not see the structure of the underlying communications medium; the terminal should become as important and wide spread and as easy to use as the telephone; and the communication system that supported the terminals should be as reliable and responsive as our private telephone network.

Immediately, we were able to make some decisions based on the communications flows within the corporation rather than on the limitations of the technology. We needed any-to-any communications to move information reliably, we needed switching and we needed conversion systems that would allow access to different computers by any terminal type as the number of terminals increased, it was also clear that we require a structured cabling system which would give an independent connection for each terminal, but which would greatly simplify installations or changes (Figure 1.1). The communications industry coined the term 'local area network' and 'wide area network' to describe different aspects of this technology, but some organization had already seen the requirement for full interconnection of these two technologies as one integrated system. Certainly data network users were not interested in the detail of the technology as long as it delivered the data.

These requirements were thought as just 'networking'. Since then, this level of integration has become known as 'internetworking'. In the late 1970s this was truly visionary for most commercial corporate environments.

The continued endurance of these statements is interesting. Despite, the many advances of the 1980s, managers in Information Technology (IT) or of Management Information Systems (MIS) still has the same goals. But now the tool is the personal computer workstation rather than a 'dumb terminal'. Data rates in common use the 1990s were, in 1979, available only in research labs.

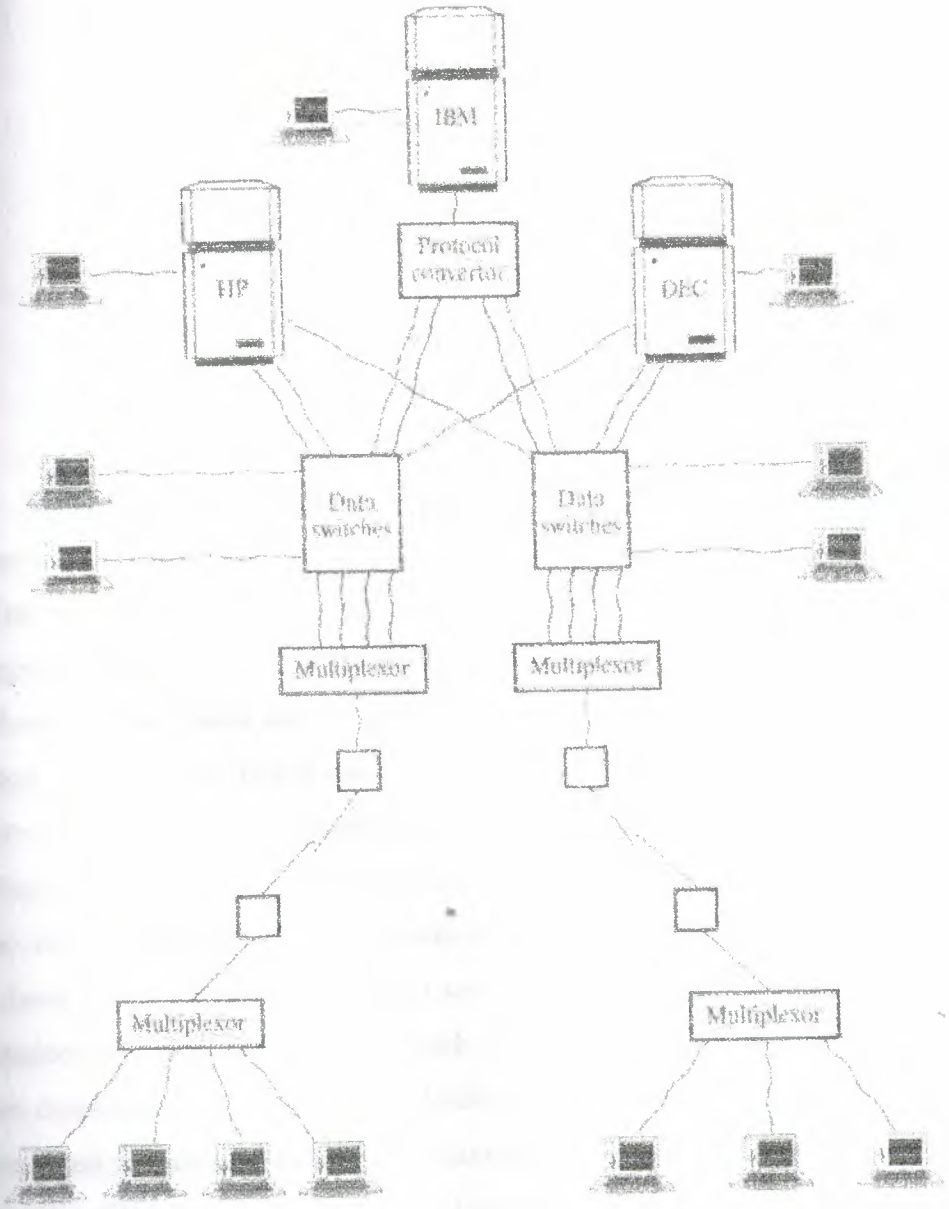

Figure 1.1 Terminal switching network

1.1.1 Achieving the goals

If the goals described are to be achieved, three separate functions are required:

1. The ability to move data anywhere in an organization with chosen reliability, security and performance.
2. The correct interpretation of that data in a manner appropriate to the receiving equipment.

Figure 1.2 Communications transparency

3. Display of the interpreted information in an acceptable form for user consumption.

The requirement for communications transparency is often expressed by showing the communications network as a cloud (Figure 1.2). Network users are not interested in the technology that makes up the cloud, or indeed, where their communicating partner is located, only that the data is delivered reliably, in good time, and at acceptable cost. The network cloud can be surrounded by a second cloud which similarly disguises the technology and 'architecture' of the communicating machines. The two clouds represent the two key components of internetworking - intercommunication and interoperation. When these issues are solved, computer users can focus on using technology to further their businesses, rather than on the details of the machines, their locations or the way they are connected. If these technical details do become visible, system managers have perhaps failed the users. Creating the standards for these two clouds has, since 1977, been an aim of the International Organization for Standardization (ISO) with their Open Systems Interconnection (ISO OSI) initiative.

The third requirement, that of displaying data, is up to individual manufacturers and is one way they differentiate their products. In the early 1980s, OSI activity was perhaps strongest outside the USA. While ISO committees were developing the OSI

protocols, the USA was developing in parallel, but with an interchange of ideas, an alternative set of techniques which became known as TCP/IP. These protocols became a US Department of Defense standard in 1983.

1.1.2 The popularity of TCP/IP

In recent years, knowledge of the capabilities of Transmission Control Protocol/Internet Protocol (TCP/IP) has spread far beyond the USA. IT managers in all types of organizations have begun to research its suitability as an internetworking technology. TCP/IP seems to be a ready made solution to the commercial information systems requirements of intercommunication and interoperation.

Many in the US government and research communities and many UNIX aficionados are already well versed in the vocabulary and configuration issues of this set of protocols as UNIX system administrators. But for the newcomer to TCP/IP, there is less information about the practical problems of implementing TCP/IP from scratch, in, for example, a commercial rather than a technical or research environment; here the skills and constraints may be very different.

The first part of this project explains the commercial, organizational and technical issues that implementing TCP/IP raises. For the newcomer to TCP/IP, we present solutions where they exist. Where they are incomplete or do not exist, we try to offer alternatives or to highlight the limitations; it is to be hoped that equipment purchasers will explain their needs to suppliers in the manner that will best achieve the development of better products! It is inescapable to know the detailed technical aspects of the protocols, showing the implications of the bits, bytes and fields for the technical planners and implementors who must track down the difficult problems of interoperation and compatibility between different implementations.

1.1.3 TCP/IP emerges

When Berkeley Software Distribution released Berkeley UNIX 4.2BSD in September 1983, a comprehensive set of 'ready-made' communications protocols called TCP/IP became much more widely available and well-known than it had been before. This was not a coincidence; its inclusion in this release was funded by the US government. TCP/IP protocols are based on standards originally developed for the US government and US research community. With the release of UNIX 4.2BSD, these communications standards emerged from the confines of the US Department of Defense

and the US university and research networks; TCP/IP became the way to interconnect UNIX systems. Berkeley UNIX 4.2BSD and subsequent releases spread quickly throughout the US university and commercial communities. As UNIX has achieved wide popularity as an 'open system', so the fame of TCP/IP has continued to spread. But TCP/IP is not, and never has been, narrowly confined to UNIX; it was developed to allow free interchange of data among all machines, independent of type, manufacturer, hardware or operating system.

In the late 1980s, TCP/IP received a further boost to its fortunes, when Sun Microsystems published the specification for Open Network Computing (ONC), often called the Network File System (NFS). NFS adds important functions to TCP/IP and is now very widely available and regarded as an integral part of the TCP/IP protocol suite. It is particularly valuable for the commercial implementor because of the simple user interfaces that it provides.

Cost-effective implementations of TCP/IP are now available for all types and sizes of machines from the largest mainframe to personal computers and workstations. This has brought TCP/IP and its capabilities to the attention of a very wide audience far beyond the initial US interest. Computer managers and users in commercial organizations throughout the world have begun to implement TCP/IP as a way of solving the problems of interworking between machines of different manufacture.

TCP/IP provides all the facilities for two computer systems to exchange information (intercommunication), interpret it properly, and present it in a format which can be understood by the local machine and its users (interoperation). NFS gives a simple and locally-familiar representation of a set of remote and possibly unfamiliar computer filing systems; like the original components of the TCP/IP suite, NFS is now available for many different computers.

In the mid-1990s, the explosion in commercial interest in the Internet gives a new market for TCP/IP products.

1.2. OSI and TCP/IP initiatives

In 1977, ISO began to develop a communications architecture which would become an international standard, a set of communications protocols known as open systems interconnection (OSI). This initiative had the same general aims as TCP/IP - intercommunication and interoperation across different manufacturers' computing

architectures - but unlike TCP/IP, in a way that met a published set of 'open' international standards. OSI now comprises many hundreds of standards, each of which has taken years to develop, agree and publish in its final ISO form. Regrettably, the best known aspect of OSI is still the OSI reference model and its seven layers (Figure 1.3); the model itself is only a development aid to allow standards developers to produce the detailed communications standards within a consistent architectural framework.

In the standard which describes the reference model, OSI standards developers state that they will exclude any details which would be implementation dependent; the result is that while the standards have been kept 'pure', many details which would aid development of viable OSI products are excluded from the standards themselves. While some would argue that OSI is more rigorous in its standardization than TCP/IP, the OSI development process seems to have become enmeshed in procedures, weighed down by the difficulties of obtaining consensus in large committees and dogged by supplier politics. By confining OSI standards to abstract definitions in a complex vocabulary defined just for the purpose and then charging considerable sums for copies of those standards, ISO committees have undoubtedly, if unintentionally, slowed the OSI development process and the delivery of useful conforming products.

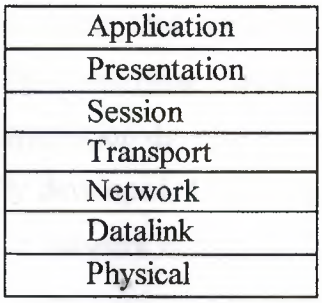

Figure 1.3 ISO OSI reference model

With a more restricted geographic and technical scope, TCP/IP developers adopted a pragmatic approach. TCP/IP standardization was based on the Request for Comments (RFC), a flexible and fast standardization process using electronic mail to publish and exchange comments and ideas, and to update drafts. Developers often outlined parts of a standard in a familiar computer language, usually 'C' which, while not intended to be implemented directly, gave a very good starting point for an initial

implementation.

TCP/IP standards are freely available on-line from a number of computer systems, originally without full drawings or graphics, but today with all the quality of a laser-printed, desktop-published document as PostScript files. If you must resort to paper and the postal service, the charge in the recent past has been a minimal $10 per copy. For manufacturers of communications and computing products, the contrast with OSI could not be more stark; it is just so much easier to obtain TCP/IP information than OSI. Standards were produced more quickly and they are written in a readable and comprehensible form by developers for developers.

The US government demanded TCP/IP for all systems, thereby ensuring every US government computer supplier provided it. They also funded universities to implement the standards. In the USA, such publicly funded work enters the public domain, and, if not of a military nature, is freely available to all citizens. While it may not be used directly for commercial purposes, having a working example in 'C' source code certainly assists future developments by commercial suppliers!

Neither OSI nor TCP/IP has been developed in isolation. There has been a considerable interchange of ideas and techniques, particularly evident in the changes in OSI since the mid-1980s with the development of the connectionless OSI suite. Nor have the OSI standards been ignored by suppliers. As with TCP/IP in the USA, universities have been busy developing OSI implementations and governments have, since the mid-1980s, required OSI-conforming products (particularly European governments through the European Commission directives). But this activity did not create a general market demand and fully developed, OSI computing products, except perhaps with the notable exception ofX.25 network equipment X.400 Mail and X.500 Directory Services.

The 'pump priming' of TCP/IP has been more successful and has ensured that it has moved ahead much faster than OSI. Governments and commercial organizations worldwide have waited patiently for OSI to become available and to reap the benefits of the promised flexibility of an international standard for computer communications and inter-operation. Suffice it to say that the development of OSI has lagged considerably behind TCP/IP, despite support from a number of governments (including since 1985, the US government and Department of Defense).

When it comes to breaking down communications barriers between different computer suppliers, information systems managers in commercial companies now see

TCP/IP as the only fully functional, proven and available option.

1.3. TCP/IP as a communications architecture

As with much of the specialist vocabulary which surrounds computers and telecommunications networks, the term TCP/IP will conjure up different concepts to different readers. TCP/IP is used as shorthand for a large set of standards with many different features and functions. The letters 'TCP/IP' stand for two communications protocols, Transmission Control Protocol (TCP) and Internet Protocol (IP). These were developed during the late 1970s and early 1980s as the key communications protocols for the 'Internet', the collected set of interconnected communications networks of the US government, the Department of Defense, the US military, and university, education, research and commercial organizations throughout the world.

The two protocols, TCP and IP, are but two of the building blocks required for a complete communications 'architecture', but the term 'TCP/IP' is most often used as a shorthand term for the whole communications 'architecture' specified originally by the US Department of Defense. This architecture is a much bigger set of standards than just TCP and IP and is more properly referred to as the Internet Protocol Suite (IPS) (Figure 1.4). We shall carry on the tradition and use 'TCP/IP' to mean the complete architecture, except where this will cause confusion.

Communications architectures have been developed by computer manufacturers since the mid-1970s. An architecture describes three facets of communications in an abstract way which is independent of particular hardware or technology. The three aspects are

(1) Data exchange (intercommunications)

(2) Data interpretation (interoperation)

(3) System management

Like the OSI reference model, communications architectures are described in layers, each layer providing its own functions but using the functions of the layer below. This layering decouples the functions of one layer from another so that layered architectures are flexible; their designers can respond to changes in technology and in application software without a major upheaval for existing users. The implementation and existing installations can be extended, hopefully indefinitely as new, often faster

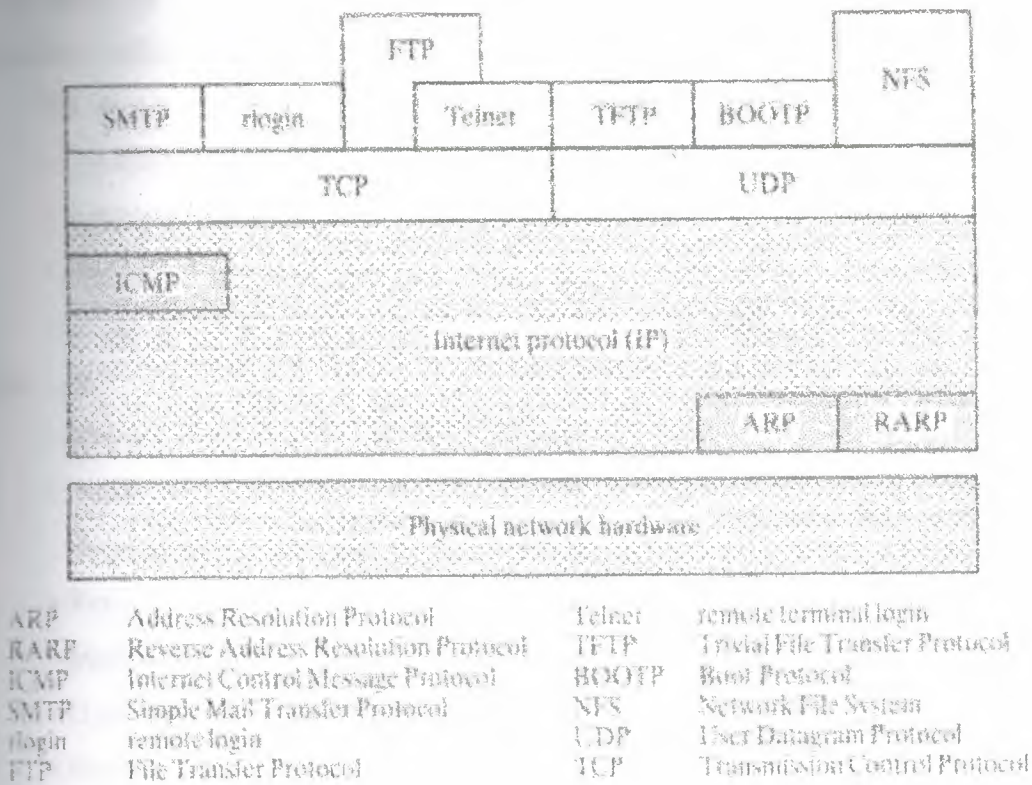

Figure 1.4 TCP/IP architecture

techniques and technologies become available. It is important to realize that the standards do not specify the interfaces seen by computer users. Though suppliers often base their implementations on a competitor's successful product, you must expect that user interfaces will differ in major or minor ways from supplier to supplier.

For TCP/IP, the architectural standards are controlled by the Internet Architecture Board (IAB). The IAB devolves its responsibilities for development, operations and management to a number of subcommittees and working groups which it controls and to other commercial companies specializing in communications and computing research and consultancy.

1.3.1. TCP/IP -The Complete Suite

The two protocols, TCP and IP, describe only the communication aspects, the movement of data across a set of interconnected physical networks.

The complete architecture must include standard mechanisms for interpreting and converting data for the common tasks that users of computers have come to expect. As

these tasks sit on top of and depend on the communications protocols, they are sometimes called the Upper Layer Protocols (ULPs).

Historically, computer users have needed three major functions:

- File transfer (including some simple file management)
- Terminal access (virtual terminal protocols)
- Mail preparation and transfer

But in today's commercial environments other tasks have become equally important

- Resource sharing (files, printers, plotters)
- Diskless workstations
- Computer conferencing
- Transaction processing
- Management
- Directory services
- Security
- Multimedia access

Those familiar with the resource-sharing Local Area Network (LAN) (such as Microsoft LAN Manager and Novell NetWare) will have seen the power of remote or distributed file and disk sharing and of peripheral sharing; for some companies, diskless workstations have a number of advantages.

In a complete, modern architecture, standard protocols are required for all these new distributed systems as well as for the proven minicomputer and mainframe-based architectures.

The TCP/IP protocol suite addresses these issues comprehensively. The standards are not static but are being added to at a steady rate. Recent activities relate to new facilities in information retrieval and display, and in directory services at the application layer and improvements to routing and addressing mechanisms at the lower layers.

1.3.2. Upper layer protocols

Beginning with the well-known application layer protocols, File Transfer Protocol (FTP), Telnet (TCP/IP's virtual terminal protocol), and Simple Mail Transfer Protocol (SMTP), NFS adds a disk/file system resource-sharing capability, LPR deals with printing, the BOOT Protocol, (BOOTP) provides the basis for diskless workstation operation, and simple Network Management Protocol (SNMP) is the Internet standard

for management. As Figure 1.4 shows, TCP is complemented by the User Datagram Protocol (UDP).

The explosion of information sources on the Internet has led to new ways of searching that information with Archie, Gopher, WAIS and the World Wide Web.

1.3.3. Lower layer protocols

IP is not the lowest level of the layered architecture. TCP/IP does not describe new standards for low-level communications but TCP/IP standards include descriptions of how IP operates over the commonly available long-distance and local physical (or link level) communications networks. These descriptions include many proprietary networks as well as ITU-T (formerly CCITT) and other international standard transmission mechanisms.

For long-distance operation over public telecommunications circuits, the standards include point-to-point leased and dial-up, synchronous and asynchronous links, and X.25 connections. For local communications, Ethernet Version 2 (as specified by DEC, Intel and Xerox) is by far the most widely used. ISO/IEEE networks (ISO 8802.3, 8802.4 8802.5) and FDDI (ISO 9314) are also specified as are proprietary networks such as ARCNET. Recent technologies like ISDN frame relay and Asynchronous Transfer Mode (ATM) are supported.

1.4. Management issues and responsibilities

The information systems manager in a commercial organization must take a highly technical product like TCP/IP and turn it into a success for the company's commercial activities, so that the investment in equipment, software and training speeds up, improves accuracy and reduces the costs of business processes. This can be a long road to travel.

Adopting and using any new system presents managers with new issues and 'opportunities' (problems!). These centre on the technology, its standards, the supplier and the organization. Many questions must be asked about the capabilities of the basic technology itself, the capabilities of the supplier and about the organization that will use that technology. Managers attempt to map the features that are available onto the day-to-day operations of the company. Only when they achieve a good fit between the facilities and features of the basic system, the capabilities and support of suppliers and

the requirements and practices of the company will the system add the highest value to the business.

In some business sectors, the application of new information technology is so revolutionary that rather than merely applying it to existing organizational structures, procedures and practices, those structures, procedures and practices are revised and simplified by it in a fundamental way.

Whatever the magnitude of the effect, there are two keys to success:

(1) Making the right choices - technological and organizational.

(2) Providing the correct training for technologists and for commercial users.

1.4.1. Choices

Many standards for TCP/IP specify only the facilities that should be available between two communicating machines. They do not specify (in general) how those facilities are delivered to the screens, keyboards, printers and disk systems of end users. There will often be one way of supporting some features, but the details of that support will differ (in what can be annoying ways). It is this variety that allows suppliers to differentiate their products and attract particular market sectors to their interpretation of the 'best' user interface for a given purpose.

The early 1990s saw a departure from some of the traditional command-line, text-based TCP/IP implementations of UNIX, towards assistance for less computer-literate users through window-based 'point and click' interfaces.

The selection of products involves choosing:

- A product with the best user interfaces for our needs.
- The features of the product to use in practice (not all facilities will be appropriate).
- How to implement and control the user features.
- How to implement and control the internal technical features.

1.5. Flexibility and control

Suppliers sell features and flexibility. Many organizations adopt TCP/IP precisely because it covers such a broad range of function and options. But flexibility is a two-edged sword: it demands a degree of imposed control otherwise freedom becomes licence, flexibility becomes a snare.

Introducing the full functionality of a system as complex as TCP/IP over a short period of time to a large user population would be likely to fail, even if it were physically possible! Dividing up the problems and phasing the introduction across a number of departments is essential. Fortunately, the model presented by Figure 1.4 is in itself the guide to how to divide the problems into manageable pieces. It is this subdivision and structured approach to problem solving that guarantees success, particularly in a large organization where there may be hundreds, or even thousands of computers, each using TCP/IP.

1.6. Separating the management functions

The layering of TCP/IP as a communications architecture allows us to separate different functions, to consider them individually, and to simplify the management of the complex whole. This approach is valid whether we merely wish to describe the tasks to be carried out by one person for a small network or to divide up the management of a large, geographically diverse, corporate network for management by many different departments and individuals. In the latter complex case, we must ensure that all aspects of the technology are addressed, otherwise what should appear as one 'seamless' cooperating system will expose the managerial divisions to the users.

As we have seen, TCP/IP provides two complementary functions:

Interoperation in applications processors through the use of the upper layer protocols (FTP, Telnet, SMTP, NFS) and intercommunication, through the use of the IP protocol operating over a collection of physical communications networks (LANs and point-to-point links). These two functions provide one possible division of responsibility: upper layer protocols operate only in personal workstations and shared mini- and mainframe computers; the internetworking protocols operate in all of these and in the routers/gateways that form the communications network. In a small implementation, one or two people are responsible for all aspects. In the larger network, the communications aspects are often implemented and managed by communications specialists. The applications are implemented and managed by business, computer system and programming specialists.

But this simple division of responsibility is a dangerous illusion. The application requires certain facilities and performance from the underlying communications networks. The communications networks must be planned with the application in mind;

they must evolve with the changing use of applications and the demand that they impose. Consider two simple examples:

- The simplest change in the configuration files of a personal computer can have an unexpected impact on network load in resource-sharing systems such as NFS. An incorrect setting on 200 computers can make the difference between a system which performs well and one which is totally unsatisfactory.
- Application programmers today develop the applications to a standard network interface in the belief that the network will deliver everything in a short time. They often test the application on a lightly loaded, small, local area network and have no concept of how that application will perform in a more general environment. A simple change in the structure of a program can alter its loading on the network by a factor of 10 to 100.

Dividing communications architectures into layers has been very successful for the development of new low-level communications media and systems; it has been too successful at isolating the applications developers from the realities of available networks and the impact that their actions have upon them.

The upper layers of TCP/IP, the applications and their environment, must be managed and controlled not just from their interface to the user, but from their impact on communications. Unfortunately, this impact is often not well understood. When it is, we find that the controls and adjustments that are available require more intervention than is desirable and are not as obvious or instinctive as they should be.

A common aim of systems and network managers is to reduce costs while delivering good service; merely optimizing and reducing the costs of their own components is not likely to provide optimum performance of the whole system. Indeed, as expressed in Figure 1.5, squeezing costs in one area can have unexpected effects in other areas. Each group must have an understanding of how the actions that they take in pursuit of reduced costs impacts on the performance and spending of the other group. Communications and systems managers need to cooperate very closely to achieve an overall successful IT service.

In large installations the problem is even more complex, for the communications and applications management is subdivided further. Management of computer applications is usually divided in one (or more) of four ways:

- by department (a functional subdivision)
- by site (geographic subdivision)

NEAR EAST UNIVERSITY

Faculty of Engineering

Department of Computer Engineering

INTERNETWORKING WITH TCP/IP

Graduation Project
COM- 400

Student: Hakan Gönülay (990951)

Supervisor: Assist.Prof.Firudin Muradov

Lefkoşa - 2003

ACKNOWLEDGEMENTS

First I want to thank Assist.Prof.Dr.Firudin Muradov to be my advisor. Under his guidence, I succesfully overcome many difficulties and learn a lot about Internetworking with TCP/IP. In each discussion, he explained my questions patiently, and I felt my quick progress from his advices. He always helps me a lot either in my study or my life. I asked him many questions in my subject and he always answered my questions quickly and in detail.

Special thanks to the Mr.Halil Adahan for his practical advices. And thanks to Faculty of Engineering for having such a good computational and electronical environment.

I also want to thank to my friends in Near East University: Sercan, Murat, Deniz and Cihan. Being with them make my 4 years in NEU full of fun.

Finally, I want to thank my family, especially my parents. Without their endless support and love for me, I would never achieve my current position. I wish my mother and fother lives happily always.

ABSTRACT

The standards collectively known as TCP/IP, first developed to allow exchange of information between computers in the US government, defence and university research communities, appeared in attractive, useful and widely available products for general commercial computing.

With the increasing interest in the use of TCP/IP for general commercial applications, there is a need to know what management and technical difficulties will be encountered.

This project is about the practical problems of installing, configuring and maintaining information systems based on the TCP/IP set of standards, from initial installation to on-going maintenance.

For a system to be successful and to retain that success over a long time, requires frequent revision on initial assumptions; system designers must take account not only of technical, but also of social and organizational problems they will encounter. Successful systems grow and develop; they are used in ways and for purposes that the initial design probably did not predict. Once convenient and reliable operation is achieved, the users abandon and then lose the older, less convenient alternatives. The new system becomes part of the corporate infrastructure and day-to-day life, its value increases and any change in performance and availability can dramatically affect prosperity and well-being.

TABLE OF CONTENTS

INTRODUCTION

The increasing interest in the use of Internet for commercial applications requires to know what management and technical difficulties will be encountered. This project describes the practical problems of installing, configuring and maintaining information systems based on the TCP/IP.

We have introduced the components of TCP/IP which support interconnectivity and interoperability, and discussed how technology affects and is affected by the size, growth rate existing structures, culture and geography of the organization that deploys it.

The layered structure of the TCP/IP protocol suite suggests how we might divide up the management responsibilities, but it must be remembered that the technology must be treated as one integrated system if it is to operate successfully. Changes in one area cannot be made without considering the consequences for performance and costs in another. If long-term success and stability is to be achieved with minimum effort, key managerial decisions must be made on specific technical parameters.

In this project, there are six chapters, that each one is complementary for the previous ones.

Chapter 1 introduces the components of TCP/IP and the need for internetworking. Also in this chapter the reader can see Upper layer, Lower layer protocols and management issues.

Chapter 2 presents the factors that allow the LANs and WANs which support TCP/IP to be planned successfully. The planning process is followed by determining the organization, its size, structure and communications flows

Chapter 3 is fully reserved for Planning and managing IP addresses. Different ways of addressing computers in a TCP/IP system, functions and format of the IP addresses, the need for a unique IP address and IP address registration are the topics discussed in this chapter.

Chapter 4 includes the topics which are the need for subnetwork addressing, the structure and format of the subnetwork mask and supernetworks.

Chapter 5 contains Routing management and configurations for routers.

Finally in chapter 6 TCP/IP upper layers, transport and applications services are described containing UDP, TCP, Winsock, The X Window system, Telnet, FTP, TFTP, SMTP and World Wide Web.

CHAPTER ONE: INTRODUCING TCP/IP

1.1. The need for internetworking

It is some 13 years since the authors first began to work together on communications networks in a large commercial environment. In that early part of our careers, which of course predated readily available sophisticated systems like TCP/IP, we were first involved in the design and implementation of a large corporate terminal-based information network. The aims of corporate network managers were expressed in three memorable phrases that appeared in many management presentations of the time:

- A single terminal on a desk
- One terminal per seated employee
- Total logical interconnectivity

The requirements were simple, if somewhat shrouded in the jargon of the day: employees should be able to access any information or system in the corporation (for which they had authority) from a single terminal on their own desk - they would not see the structure of the underlying communications medium; the terminal should become as important and wide spread and as easy to use as the telephone; and the communication system that supported the terminals should be as reliable and responsive as our private telephone network.

Immediately, we were able to make some decisions based on the communications flows within the corporation rather than on the limitations of the technology. We needed any-to-any communications to move information reliably, we needed switching and we needed conversion systems that would allow access to different computers by any terminal type as the number of terminals increased, it was also clear that we require a structured cabling system which would give an independent connection for each terminal, but which would greatly simplify installations or changes (Figure 1.1). The communications industry coined the term 'local area network' and 'wide area network' to describe different aspects of this technology, but some organization had already seen the requirement for full interconnection of these two technologies as one integrated system. Certainly data network users were not interested in the detail of the technology as long as it delivered the data.

These requirements were thought as just 'networking'. Since then, this level of integration has become known as 'internetworking'. In the late 1970s this was truly visionary for most commercial corporate environments.

The continued endurance of these statements is interesting. Despite, the many advances of the 1980s, managers in Information Technology (IT) or of Management Information Systems (MIS) still has the same goals. But now the tool is the personal computer workstation rather than a 'dumb terminal'. Data rates in common use the 1990s were, in 1979, available only in research labs.

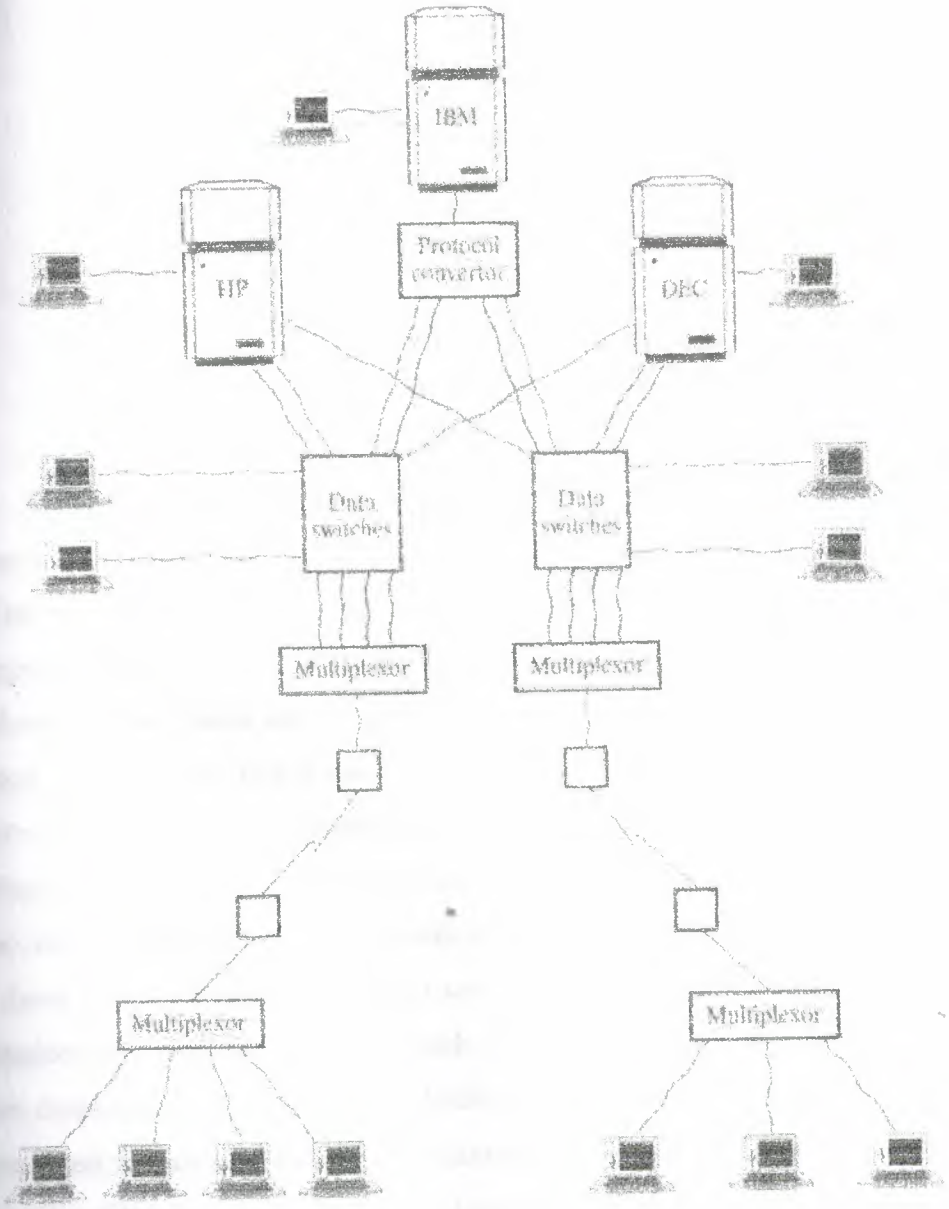

Figure 1.1 Terminal switching network

1.1.1 Achieving the goals

If the goals described are to be achieved, three separate functions are required:

1. The ability to move data anywhere in an organization with chosen reliability, security and performance.
2. The correct interpretation of that data in a manner appropriate to the receiving equipment.

Figure 1.2 Communications transparency

3. Display of the interpreted information in an acceptable form for user consumption.

The requirement for communications transparency is often expressed by showing the communications network as a cloud (Figure 1.2). Network users are not interested in the technology that makes up the cloud, or indeed, where their communicating partner is located, only that the data is delivered reliably, in good time, and at acceptable cost. The network cloud can be surrounded by a second cloud which similarly disguises the technology and 'architecture' of the communicating machines. The two clouds represent the two key components of internetworking - intercommunication and interoperation. When these issues are solved, computer users can focus on using technology to further their businesses, rather than on the details of the machines, their locations or the way they are connected. If these technical details do become visible, system managers have perhaps failed the users. Creating the standards for these two clouds has, since 1977, been an aim of the International Organization for Standardization (ISO) with their Open Systems Interconnection (ISO OSI) initiative.

The third requirement, that of displaying data, is up to individual manufacturers and is one way they differentiate their products. In the early 1980s, OSI activity was perhaps strongest outside the USA. While ISO committees were developing the OSI

protocols, the USA was developing in parallel, but with an interchange of ideas, an alternative set of techniques which became known as TCP/IP. These protocols became a US Department of Defense standard in 1983.

1.1.2 The popularity of TCP/IP

In recent years, knowledge of the capabilities of Transmission Control Protocol/Internet Protocol (TCP/IP) has spread far beyond the USA. IT managers in all types of organizations have begun to research its suitability as an internetworking technology. TCP/IP seems to be a ready made solution to the commercial information systems requirements of intercommunication and interoperation.

Many in the US government and research communities and many UNIX aficionados are already well versed in the vocabulary and configuration issues of this set of protocols as UNIX system administrators. But for the newcomer to TCP/IP, there is less information about the practical problems of implementing TCP/IP from scratch, in, for example, a commercial rather than a technical or research environment; here the skills and constraints may be very different.

The first part of this project explains the commercial, organizational and technical issues that implementing TCP/IP raises. For the newcomer to TCP/IP, we present solutions where they exist. Where they are incomplete or do not exist, we try to offer alternatives or to highlight the limitations; it is to be hoped that equipment purchasers will explain their needs to suppliers in the manner that will best achieve the development of better products! It is inescapable to know the detailed technical aspects of the protocols, showing the implications of the bits, bytes and fields for the technical planners and implementors who must track down the difficult problems of interoperation and compatibility between different implementations.

1.1.3 TCP/IP emerges

When Berkeley Software Distribution released Berkeley UNIX 4.2BSD in September 1983, a comprehensive set of 'ready-made' communications protocols called TCP/IP became much more widely available and well-known than it had been before. This was not a coincidence; its inclusion in this release was funded by the US government. TCP/IP protocols are based on standards originally developed for the US government and US research community. With the release of UNIX 4.2BSD, these communications standards emerged from the confines of the US Department of Defense

and the US university and research networks; TCP/IP became the way to interconnect UNIX systems. Berkeley UNIX 4.2BSD and subsequent releases spread quickly throughout the US university and commercial communities. As UNIX has achieved wide popularity as an 'open system', so the fame of TCP/IP has continued to spread. But TCP/IP is not, and never has been, narrowly confined to UNIX; it was developed to allow free interchange of data among all machines, independent of type, manufacturer, hardware or operating system.

In the late 1980s, TCP/IP received a further boost to its fortunes, when Sun Microsystems published the specification for Open Network Computing (ONC), often called the Network File System (NFS). NFS adds important functions to TCP/IP and is now very widely available and regarded as an integral part of the TCP/IP protocol suite. It is particularly valuable for the commercial implementor because of the simple user interfaces that it provides.

Cost-effective implementations of TCP/IP are now available for all types and sizes of machines from the largest mainframe to personal computers and workstations. This has brought TCP/IP and its capabilities to the attention of a very wide audience far beyond the initial US interest. Computer managers and users in commercial organizations throughout the world have begun to implement TCP/IP as a way of solving the problems of interworking between machines of different manufacture.

TCP/IP provides all the facilities for two computer systems to exchange information (intercommunication), interpret it properly, and present it in a format which can be understood by the local machine and its users (interoperation). NFS gives a simple and locally-familiar representation of a set of remote and possibly unfamiliar computer filing systems; like the original components of the TCP/IP suite, NFS is now available for many different computers.

In the mid-1990s, the explosion in commercial interest in the Internet gives a new market for TCP/IP products.

1.2. OSI and TCP/IP initiatives

In 1977, ISO began to develop a communications architecture which would become an international standard, a set of communications protocols known as open systems interconnection (OSI). This initiative had the same general aims as TCP/IP - intercommunication and interoperation across different manufacturers' computing

architectures - but unlike TCP/IP, in a way that met a published set of 'open' international standards. OSI now comprises many hundreds of standards, each of which has taken years to develop, agree and publish in its final ISO form. Regrettably, the best known aspect of OSI is still the OSI reference model and its seven layers (Figure 1.3); the model itself is only a development aid to allow standards developers to produce the detailed communications standards within a consistent architectural framework.

In the standard which describes the reference model, OSI standards developers state that they will exclude any details which would be implementation dependent; the result is that while the standards have been kept 'pure', many details which would aid development of viable OSI products are excluded from the standards themselves. While some would argue that OSI is more rigorous in its standardization than TCP/IP, the OSI development process seems to have become enmeshed in procedures, weighed down by the difficulties of obtaining consensus in large committees and dogged by supplier politics. By confining OSI standards to abstract definitions in a complex vocabulary defined just for the purpose and then charging considerable sums for copies of those standards, ISO committees have undoubtedly, if unintentionally, slowed the OSI development process and the delivery of useful conforming products.

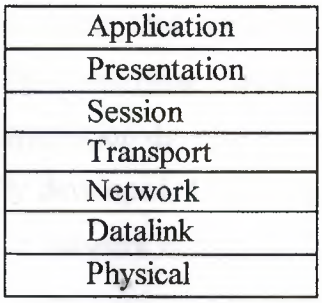

Figure 1.3 ISO OSI reference model

With a more restricted geographic and technical scope, TCP/IP developers adopted a pragmatic approach. TCP/IP standardization was based on the Request for Comments (RFC), a flexible and fast standardization process using electronic mail to publish and exchange comments and ideas, and to update drafts. Developers often outlined parts of a standard in a familiar computer language, usually 'C' which, while not intended to be implemented directly, gave a very good starting point for an initial

implementation.

TCP/IP standards are freely available on-line from a number of computer systems, originally without full drawings or graphics, but today with all the quality of a laser-printed, desktop-published document as PostScript files. If you must resort to paper and the postal service, the charge in the recent past has been a minimal $10 per copy. For manufacturers of communications and computing products, the contrast with OSI could not be more stark; it is just so much easier to obtain TCP/IP information than OSI. Standards were produced more quickly and they are written in a readable and comprehensible form by developers for developers.

The US government demanded TCP/IP for all systems, thereby ensuring every US government computer supplier provided it. They also funded universities to implement the standards. In the USA, such publicly funded work enters the public domain, and, if not of a military nature, is freely available to all citizens. While it may not be used directly for commercial purposes, having a working example in 'C' source code certainly assists future developments by commercial suppliers!

Neither OSI nor TCP/IP has been developed in isolation. There has been a considerable interchange of ideas and techniques, particularly evident in the changes in OSI since the mid-1980s with the development of the connectionless OSI suite. Nor have the OSI standards been ignored by suppliers. As with TCP/IP in the USA, universities have been busy developing OSI implementations and governments have, since the mid-1980s, required OSI-conforming products (particularly European governments through the European Commission directives). But this activity did not create a general market demand and fully developed, OSI computing products, except perhaps with the notable exception ofX.25 network equipment X.400 Mail and X.500 Directory Services.

The 'pump priming' of TCP/IP has been more successful and has ensured that it has moved ahead much faster than OSI. Governments and commercial organizations worldwide have waited patiently for OSI to become available and to reap the benefits of the promised flexibility of an international standard for computer communications and inter-operation. Suffice it to say that the development of OSI has lagged considerably behind TCP/IP, despite support from a number of governments (including since 1985, the US government and Department of Defense).

When it comes to breaking down communications barriers between different computer suppliers, information systems managers in commercial companies now see

TCP/IP as the only fully functional, proven and available option.

1.3. TCP/IP as a communications architecture

As with much of the specialist vocabulary which surrounds computers and telecommunications networks, the term TCP/IP will conjure up different concepts to different readers. TCP/IP is used as shorthand for a large set of standards with many different features and functions. The letters 'TCP/IP' stand for two communications protocols, Transmission Control Protocol (TCP) and Internet Protocol (IP). These were developed during the late 1970s and early 1980s as the key communications protocols for the 'Internet', the collected set of interconnected communications networks of the US government, the Department of Defense, the US military, and university, education, research and commercial organizations throughout the world.

The two protocols, TCP and IP, are but two of the building blocks required for a complete communications 'architecture', but the term 'TCP/IP' is most often used as a shorthand term for the whole communications 'architecture' specified originally by the US Department of Defense. This architecture is a much bigger set of standards than just TCP and IP and is more properly referred to as the Internet Protocol Suite (IPS) (Figure 1.4). We shall carry on the tradition and use 'TCP/IP' to mean the complete architecture, except where this will cause confusion.

Communications architectures have been developed by computer manufacturers since the mid-1970s. An architecture describes three facets of communications in an abstract way which is independent of particular hardware or technology. The three aspects are

(1) Data exchange (intercommunications)

(2) Data interpretation (interoperation)

(3) System management

Like the OSI reference model, communications architectures are described in layers, each layer providing its own functions but using the functions of the layer below. This layering decouples the functions of one layer from another so that layered architectures are flexible; their designers can respond to changes in technology and in application software without a major upheaval for existing users. The implementation and existing installations can be extended, hopefully indefinitely as new, often faster

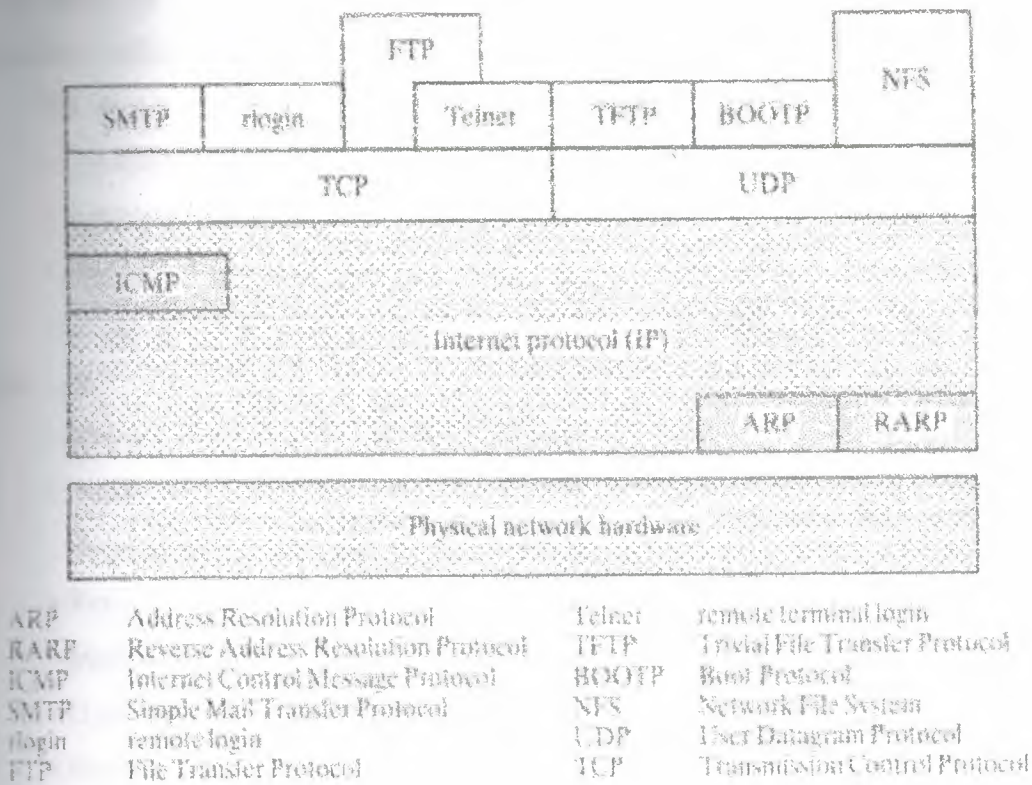

Figure 1.4 TCP/IP architecture

techniques and technologies become available. It is important to realize that the standards do not specify the interfaces seen by computer users. Though suppliers often base their implementations on a competitor's successful product, you must expect that user interfaces will differ in major or minor ways from supplier to supplier.

For TCP/IP, the architectural standards are controlled by the Internet Architecture Board (IAB). The IAB devolves its responsibilities for development, operations and management to a number of subcommittees and working groups which it controls and to other commercial companies specializing in communications and computing research and consultancy.

1.3.1. TCP/IP -The Complete Suite

The two protocols, TCP and IP, describe only the communication aspects, the movement of data across a set of interconnected physical networks.

The complete architecture must include standard mechanisms for interpreting and converting data for the common tasks that users of computers have come to expect. As

these tasks sit on top of and depend on the communications protocols, they are sometimes called the Upper Layer Protocols (ULPs).

Historically, computer users have needed three major functions:

- File transfer (including some simple file management)
- Terminal access (virtual terminal protocols)
- Mail preparation and transfer

But in today's commercial environments other tasks have become equally important

- Resource sharing (files, printers, plotters)
- Diskless workstations
- Computer conferencing
- Transaction processing
- Management
- Directory services
- Security
- Multimedia access

Those familiar with the resource-sharing Local Area Network (LAN) (such as Microsoft LAN Manager and Novell NetWare) will have seen the power of remote or distributed file and disk sharing and of peripheral sharing; for some companies, diskless workstations have a number of advantages.

In a complete, modern architecture, standard protocols are required for all these new distributed systems as well as for the proven minicomputer and mainframe-based architectures.

The TCP/IP protocol suite addresses these issues comprehensively. The standards are not static but are being added to at a steady rate. Recent activities relate to new facilities in information retrieval and display, and in directory services at the application layer and improvements to routing and addressing mechanisms at the lower layers.

1.3.2. Upper layer protocols

Beginning with the well-known application layer protocols, File Transfer Protocol (FTP), Telnet (TCP/IP's virtual terminal protocol), and Simple Mail Transfer Protocol (SMTP), NFS adds a disk/file system resource-sharing capability, LPR deals with printing, the BOOT Protocol, (BOOTP) provides the basis for diskless workstation operation, and simple Network Management Protocol (SNMP) is the Internet standard

for management. As Figure 1.4 shows, TCP is complemented by the User Datagram Protocol (UDP).

The explosion of information sources on the Internet has led to new ways of searching that information with Archie, Gopher, WAIS and the World Wide Web.

1.3.3. Lower layer protocols

IP is not the lowest level of the layered architecture. TCP/IP does not describe new standards for low-level communications but TCP/IP standards include descriptions of how IP operates over the commonly available long-distance and local physical (or link level) communications networks. These descriptions include many proprietary networks as well as ITU-T (formerly CCITT) and other international standard transmission mechanisms.

For long-distance operation over public telecommunications circuits, the standards include point-to-point leased and dial-up, synchronous and asynchronous links, and X.25 connections. For local communications, Ethernet Version 2 (as specified by DEC, Intel and Xerox) is by far the most widely used. ISO/IEEE networks (ISO 8802.3, 8802.4 8802.5) and FDDI (ISO 9314) are also specified as are proprietary networks such as ARCNET. Recent technologies like ISDN frame relay and Asynchronous Transfer Mode (ATM) are supported.

1.4. Management issues and responsibilities

The information systems manager in a commercial organization must take a highly technical product like TCP/IP and turn it into a success for the company's commercial activities, so that the investment in equipment, software and training speeds up, improves accuracy and reduces the costs of business processes. This can be a long road to travel.

Adopting and using any new system presents managers with new issues and 'opportunities' (problems!). These centre on the technology, its standards, the supplier and the organization. Many questions must be asked about the capabilities of the basic technology itself, the capabilities of the supplier and about the organization that will use that technology. Managers attempt to map the features that are available onto the day-to-day operations of the company. Only when they achieve a good fit between the facilities and features of the basic system, the capabilities and support of suppliers and

the requirements and practices of the company will the system add the highest value to the business.

In some business sectors, the application of new information technology is so revolutionary that rather than merely applying it to existing organizational structures, procedures and practices, those structures, procedures and practices are revised and simplified by it in a fundamental way.

Whatever the magnitude of the effect, there are two keys to success:

(1) Making the right choices - technological and organizational.

(2) Providing the correct training for technologists and for commercial users.

1.4.1. Choices

Many standards for TCP/IP specify only the facilities that should be available between two communicating machines. They do not specify (in general) how those facilities are delivered to the screens, keyboards, printers and disk systems of end users. There will often be one way of supporting some features, but the details of that support will differ (in what can be annoying ways). It is this variety that allows suppliers to differentiate their products and attract particular market sectors to their interpretation of the 'best' user interface for a given purpose.

The early 1990s saw a departure from some of the traditional command-line, text-based TCP/IP implementations of UNIX, towards assistance for less computer-literate users through window-based 'point and click' interfaces.

The selection of products involves choosing:

- A product with the best user interfaces for our needs.
- The features of the product to use in practice (not all facilities will be appropriate).
- How to implement and control the user features.
- How to implement and control the internal technical features.

1.5. Flexibility and control

Suppliers sell features and flexibility. Many organizations adopt TCP/IP precisely because it covers such a broad range of function and options. But flexibility is a two-edged sword: it demands a degree of imposed control otherwise freedom becomes licence, flexibility becomes a snare.

Introducing the full functionality of a system as complex as TCP/IP over a short period of time to a large user population would be likely to fail, even if it were physically possible! Dividing up the problems and phasing the introduction across a number of departments is essential. Fortunately, the model presented by Figure 1.4 is in itself the guide to how to divide the problems into manageable pieces. It is this subdivision and structured approach to problem solving that guarantees success, particularly in a large organization where there may be hundreds, or even thousands of computers, each using TCP/IP.

1.6. Separating the management functions

The layering of TCP/IP as a communications architecture allows us to separate different functions, to consider them individually, and to simplify the management of the complex whole. This approach is valid whether we merely wish to describe the tasks to be carried out by one person for a small network or to divide up the management of a large, geographically diverse, corporate network for management by many different departments and individuals. In the latter complex case, we must ensure that all aspects of the technology are addressed, otherwise what should appear as one 'seamless' cooperating system will expose the managerial divisions to the users.

As we have seen, TCP/IP provides two complementary functions:

Interoperation in applications processors through the use of the upper layer protocols (FTP, Telnet, SMTP, NFS) and intercommunication, through the use of the IP protocol operating over a collection of physical communications networks (LANs and point-to-point links). These two functions provide one possible division of responsibility: upper layer protocols operate only in personal workstations and shared mini- and mainframe computers; the internetworking protocols operate in all of these and in the routers/gateways that form the communications network. In a small implementation, one or two people are responsible for all aspects. In the larger network, the communications aspects are often implemented and managed by communications specialists. The applications are implemented and managed by business, computer system and programming specialists.

But this simple division of responsibility is a dangerous illusion. The application requires certain facilities and performance from the underlying communications networks. The communications networks must be planned with the application in mind;

they must evolve with the changing use of applications and the demand that they impose. Consider two simple examples:

- The simplest change in the configuration files of a personal computer can have an unexpected impact on network load in resource-sharing systems such as NFS. An incorrect setting on 200 computers can make the difference between a system which performs well and one which is totally unsatisfactory.
- Application programmers today develop the applications to a standard network interface in the belief that the network will deliver everything in a short time. They often test the application on a lightly loaded, small, local area network and have no concept of how that application will perform in a more general environment. A simple change in the structure of a program can alter its loading on the network by a factor of 10 to 100.

Dividing communications architectures into layers has been very successful for the development of new low-level communications media and systems; it has been too successful at isolating the applications developers from the realities of available networks and the impact that their actions have upon them.

The upper layers of TCP/IP, the applications and their environment, must be managed and controlled not just from their interface to the user, but from their impact on communications. Unfortunately, this impact is often not well understood. When it is, we find that the controls and adjustments that are available require more intervention than is desirable and are not as obvious or instinctive as they should be.

A common aim of systems and network managers is to reduce costs while delivering good service; merely optimizing and reducing the costs of their own components is not likely to provide optimum performance of the whole system. Indeed, as expressed in Figure 1.5, squeezing costs in one area can have unexpected effects in other areas. Each group must have an understanding of how the actions that they take in pursuit of reduced costs impacts on the performance and spending of the other group. Communications and systems managers need to cooperate very closely to achieve an overall successful IT service.

In large installations the problem is even more complex, for the communications and applications management is subdivided further. Management of computer applications is usually divided in one (or more) of four ways:

- by department (a functional subdivision)
- by site (geographic subdivision)

- by manufacturer of computer system (technical subdivision)
- by computing function (FTP, SMTP, Telnet)

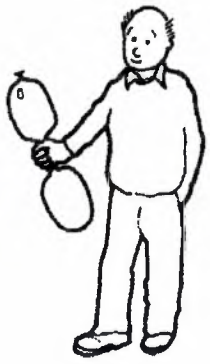

Figure 1.5 Squeezing costs in one area

Communications functions are often split horizontally and vertically, along the TCP/IP architectural model, by function and by technology into:

- Physical layer - building cabling
- Link layer - Ethernet, Token Ring, wide area circuits
- Network layer - Internet protocol (IP)

To this list should be added the transport layer; in TCP/IP the transport layer is TCP and UDP. Immediately there is a conflict of interest, for while these are quite definitely communications protocols and have the major impact on communications performance, they are only present in the host systems of TCP/IP, not in the traditional network components. If TCP/IP systems are to be successful, overcoming this territorial and organizational issue is key.

1.6.1. The design authority

Whatever the size of the organization and its networks, TCP/IP only operates correctly if the same high standards of configuration are used in every operational attached piece of equipment. This is easiest if the major design decisions are imposed by or agreed with a single (central) design authority which is educated to make the correct managerial and technical choices and then ensures their observance throughout the interconnected system. TCP/IP grew from a free culture of cooperating organizations and individuals with a collegiate and research ethos; overall standards and control came from the Internet Architecture Board. The same degree of cooperation and control must be present in an operational commercial TCP/IP network. The control will

normally come from somebody with the necessary technical and organizational authority.

1.7. Technical decisions

What technical decisions must be made? The complexity of the management problem depends on the anticipated size and growth rate of the overall system. Size, growth rate and corporate organization and culture must all be considered when the technical (and organizational) choices are taken during the planning and implementation processes. These choices may be conveniently considered in the layered way that the TCP/IP architecture suggests to us. The important technical decisions to be made centre around the following topics:

- The size of the problem
- Reliability
- Availability
- Budget and costs
- Local and wide area networks to support TCP/IP
- Network addressing scheme (IP address space)
- Use of subnetwork addresses
- Router network and its routing protocols
- Domain name service
- Application layers, including their communications effects
- Overall system management and record keeping

These subjects are interdependent. The structure and geography of the organization determines the underlying communication requirements. This interacts with the layout of LANs. That layout predetermines the flexibility of network addressing, subnetwork addressing and the placement of routers.

Certain combinations of technology can at first sight appear to be compatible; the problems only become evident much later (Figure 1.6).

Subsequent chapters of chapter 1 discuss each of the above as organizational, managerial and user issues without detailed technical description of the bits and bytes involved.

1.7.1. Choosing equipment

In a description of practical issues we frequently find that the standards include many features and facilities which are not generally implemented in available equipment.

Many descriptions of TCP/IP concentrate on the facilities that are available within the protocol and not on the features that can really be exploited by today's equipment. We have frequently found that TCP/IP was implemented for facilities which in practice it cannot provide. Once network managers have determined what facilities they require they should produce a requirements specification for their suppliers, so they can determine what facilities are supported and the level of that support.

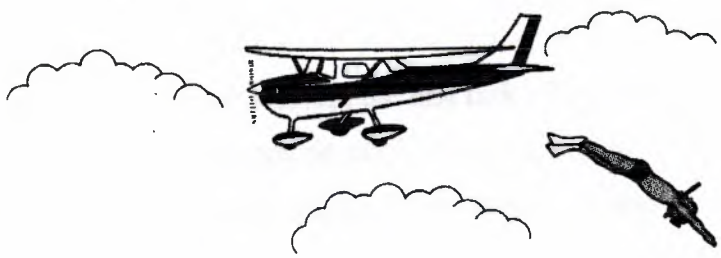

Figure 1.6 Choosing the wrong combinations of technology

Throughout the book we will give an indication of what is generally available, but that can be expected to change quite quickly. Readers must treat such statements as a trigger to ask the right questions of their suppliers to obtain the latest state of play.

CHAPTER TWO: ESTABLISHING THE NETWORK FOUNDATION

Any distributed computer system is composed of two major components:

The processors, which provide applications services to computer users; the physical local and wide area networks and cable systems, which carry information between the sources of data (such as file servers) and the users of that data at workstations or personal computers. We shall discuss application questions later. First, we examine the local and wide area network implications of TCP/IP. This discussion cannot be completely divorced from a consideration of network addressing and of

routing, for the three topics are interdependent. The structure of Local Area Network (LAN) and Wide Area Network (WAN) cables and the presence of bridges and routers change the TCP/IP addressing scheme. Computer systems on the same bridged LAN must be within the same address range. On either side of a router they are in different address ranges. Setting the right standards for the underlying LAN structure allows the design of TCP/IP to be right first time. For reliable and efficient operation, TCP/IP must operate over a stable platform of LAN cabling and components and wide area circuits. In this chapter, we consider how to design and implement that low-level infrastructure.

2.1. Planning the supporting networks

The aim of TCP/IP network planning and design is to provide a communications infrastructure that meets the requirements of the organization with readily available equipment and skills. It must provide the correct level of performance for different functions within the organization, at different places, and at acceptable costs. In setting up a local and wide area structure for TCP/IP, the first task of the designer is to consider what each of these terms means for that situation. Each case is unique but the overall size of the organization and its network requirements will determine the complexity of the problems to be solved.

2.1.1. Size and growth rate

Whatever aspect of human communications you consider, it could be said that a measurement of success is often that the size and usage grows rapidly. As a result of experience in road network planning, this has become known as the 'motorway (or highway) effect'. If a system is useful, people change their travel habits and use the system in new ways that the planners never envisaged (Figure 2.1). Traffic grows to fill the space available. In our organization the network grew at a rate which was totally unexpected and expanded rapidly to cover the company.

When planning a TCP/IP installation it is advisable to consider the initial size of the network and have some indication of maximum size and growth rate. This measure of size has two dimensions:

• The total number of attached ports, that is, computer and router connections to the network

• The geographic coverage

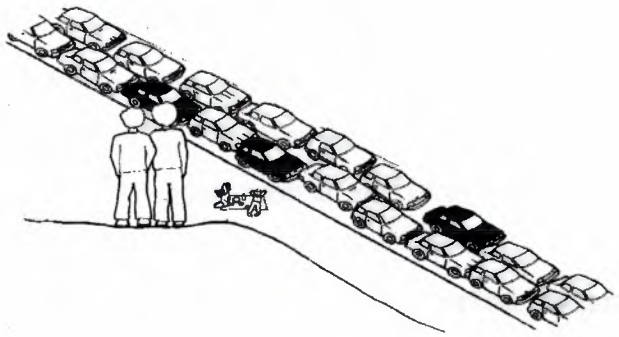

Figure 2.1 The motorway effect

It is also important to consider in what ways the network could grow - by addition of devices on the same network or by connection to other TCP/IP installations, for example by mergers and acquisitions. In the first case, all the issues are under the control of one organization; in the second, 0|6y may not be. The problems to be solved will be different in each case. Today's computer applications and systems are distributed on workstations and PCs throughout the sites and buildings where the computer lasers sit. Some types of change cannot be centrally managed and must be carried out at each machine. These changes are always difficult and labour-intensive, and hence costly. As far as possible, they must be predicted and eliminated.

These considerations can affect the detailed planning process of the LAN and WAN links, the planning of the address space and of the bridges and routers which interconnect LANs and sites. Planners will wish to determine solutions to the following questions. How many cable Interns will there be? How can they be interconnected successfully? Will bridges suffice or are routers required? How should the address plan for the network evolve? Is there a requirement for the Domain Name Service? Is responsibility split or with one person or body? If it is split, is the division by function, by location, or by technology? Is there a need to split application management from communications management? Is there a need (in a large corporation) to split communications management into more than one subfunction?

2.1.2. Existing standards

In an attempt to avoid duplication of effort and perhaps more importantly to avoid

choosing incompatible solutions to the same TCP/IP options and problems, it is quite important to determine if there are any other existing TCP/IP implementations in the organization. Some systems may hide their use of TCP/IP; the Banyan VINES resource-sharing LAN or any communications management system which uses the simple network management protocol are current examples. If other examples do exist and will share the same cables or TCP/IP equipment at any time, the developments must be coordinated.

In some situations there may already be a central authority for TCP/IP implementation which has set local standards for the use of TCP/IP. These standards should relate to:

- Addressing conventions
- Host-naming conventions and domain name system
- Local area network
- Choice of equipment
- Choice of network software
- Configuration and management of the network
- Configuration and management of the applications
- Connection to the Internet
- Security

It is much easier if these standards can be used rather than starting from scratch.

2.1.3. Traffic flows and capacity

In planning the network layout and capacity it is important to know the volume of information that will be carried. At the simplest level, high-performance routes in the network must follow the high traffic flows in the organization. Faults on those major arterial routes will affect more users, so they may need to be protected against failure.

Judging future traffic flows is always one of the most difficult estimates in installing new TCP/IP systems. One of the great successes of modern software development techniques is to provide an effective barrier between the application developer and the underlying network. But it leaves the network planners with a major problem. Few people can relate the use of an application to the traffic it will generate.

Given the continued growth in workstation performance, existing LAN" implementations already seem slow. For many commercial environments, a single physical LAN can support only 20 to 50 workstations and one or two file servers before

the network must be partitioned with bridges or routers. The need for switching hubs, 'private' LANs and 100 Mbps Fibre Distributed Data Interface (FDDI) or even higher network speeds (110 Gbps) is evident, particularly for backbone networks and image systems, but the costs of such a system are not yet within the budgets of many network operators.

In the absence of such speeds, planners must ration the limited resource of normal speed LANs. They must understand how particular applications use the network and what the traffic flows between different user groups and applications processors are at different times of the day and business year. Key questions are:

• What will be the total volume of data on different parts of the network at the busiest time of the day? Is this traffic seasonal?

• How will the traffic grow with time, taking into account the changing and improving user perception of the network services and the increase in number of attached devices and a move towards more demanding applications?

• How is the traffic distributed among adjacent machines and to more remote machines?

Often the only way of obtaining such information is to measure the operation of real applications and extrapolate the results using spreadsheet or other modelling systems. Armed with the traffic information, and its growth, it is time to consider the network structure.

2.2. Network decisions

In a small, compact implementation of a few PCs and file servers in one building, the main issues and decisions centre on the choice of addresses and on configuration and management of the computers. The network will be a few lengths of Unshielded Twisted Pair (UTP) cable or a simple Token Ring installation with, at most, one or two bridges to increase the total traffic capacity of the system. There are unlikely to be any routers. Implementation of the network is usually straightforward. In these small networks, one or two people make most of the decisions; they may also be responsible for carrying out the implementation and configuration.

Larger corporate networks often cover many sites and require long distance PTT-provided (Telco-provided) circuits to interconnect those sites. Rentals for such circuits can form a major portion of the cost which may increase with time Complex

configuration of remote bridges and (In many countries the unit cost of intersite circuits has reduced considerably over the past 10 years, but the increase in capacity required for interconnected LANs has far outstripped the unit cost reductions.) routers is required to ensure that these wide area circuits are used at optimum efficiency.

Figure 2.2 Multiple protocols

Networks are subject to hidden costs. Heavily loaded networks, where the traffic approaches the capacity of the cable, require more management time as systems are reconfigured and moved from cable to cable to reduce loading. Idiosyncrasies and limitations of the overall system and software faults always become apparent at high system loading. A heavily loaded network can also impose a load on application processors, reducing the amount of useful power available for end users.

Management decisions about networks centre on how to provide the right performance with chosen reliability to particular users and at acceptable cost. Options centre on the choice of technology for the cable, the interface cards and the bridges and routers that form the interconnection. These three components form the three lower layers of the OSI model. Cables are part of the OSI physical layer (layer 1); LAN bridges operate medium access control, a part of the OSI data link layer (layer 2);

Routers operate in the network layer (layer 3). TCP/IP performance is not affected by the choice of cable as long as that cable can support the LAN technology at the normal operating speed in an error-free way. We shall not therefore discuss the issues of cabling further.

2.2.1.Link layer options

LAN cables and wide area circuits may be a shared resource, carrying not only

TCP/IP data, but data from other applications and systems as well (Novell NetWare, IBM SNA, DECnet to name but a few examples). Transporting multiple protocols on one infrastructure increases the organizational and technical complexity. In these networks it is certain that the management of network resources, the infrastructure which carries data between application processors, will present as many difficulties as the management of applications themselves. A detailed discussion of these issues is beyond the scope of this project, but for those contemplating using TCP/IP in a multiprotocol environment, it is much easier for protocols to coexist peacefully if the number of options is reduced (Figure 2.2).

The first decision is what option to use for carrying information across the LAN cable that everyone colloquially calls Ethernet. Ethernet (even Version 2 Ethernet) is a different standard from ISO 8802.3+ (IEEE 802.3). While the hardware is compatible for both systems, the frame content is not. Furthermore, Ethernet V2 is a de facto standard; IEEE 802.3 is the Open Systems standard which was adopted as the international standard by ISO.

TCP/IP on LANs predates these international standards, so the most common implementation by far for TCP/IP is Ethernet V2; while most manufacturers of TCP/IP offer IEEE 802.3/802.2 with Sub Network Access Protocol (SNAP), the changeover to IEEE 802.3 encapsulation should not be undertaken lightly; the two standards can coexist on the same Ethernet cable, but they cannot directly interwork. Where both standards are in use, careful planning is required to be sure that all systems can see TCP/IP data in the form that they can recognize. Either the different systems must be segregated by frame type, or some device must selectively retransmit frames in the other format, effectively duplicating that traffic on the same cable and increasing the network loading. On Token Ring (IEEE 802.5) and FDDI (IS9314) networks there is less confusion; ISO standard frames are used exclusively with SNAP encapsulation carrying the Ethernet type field.

2.2.2. Structuring LAN interconnections

Directly attaching all computing resources to one extended 'local area' network of the highest possible speed without intervening bridges and routers gives the best possible performance.

Today's high performance file servers, personal computers and workstations can

a single Ethernet (lOMbps) or Token Ring (4Mbps or 16Mbps) segment. For satisfactory operation, the alternative to replacing the network with one of higher speeds is to partition the overall system into small groups with a common communications interest.

This subdivision was originally done with bridges or routers but increasingly 'switching hubs' are used to connect a user directly to a server for the duration of one LAN frame or message. With the proper structure, the LAN is no longer shared amongst many servers, but each server is effectively given its own high capacity 'pipe' to its groups of users.

Rather than repeat the use of ISO and IEEE equivalent standards for LANs, we shall follow the practice of using the IEEE designations. These have been adopted verbatim for ISO OSI standards.

There are two standards for transmitting IP information on the cable that is commonly referred to as Ethernet. The first is accurately described as Ethernet Version 2; the second requires the lengthy description 'IEEE 802.3 + IEEE 802.2. + SNAP.

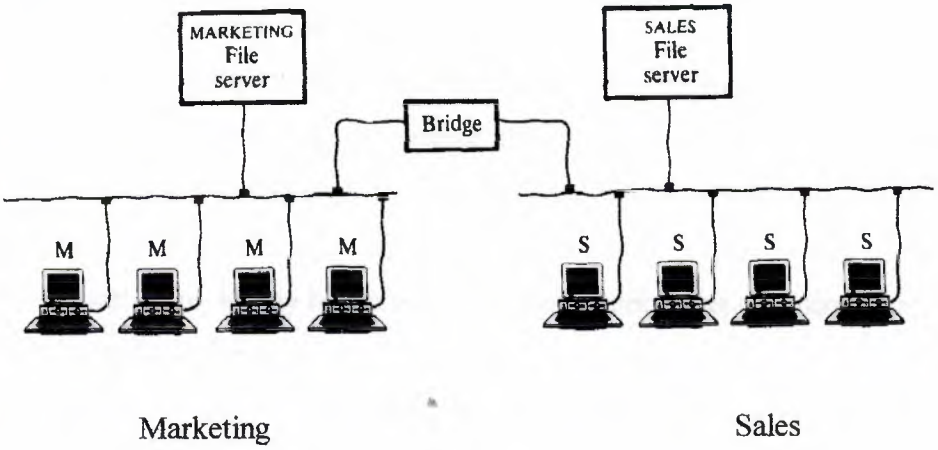

Figure 2.3 Location of servers

Switching hubs are available in a wide variety of formats and sizes; within a building or local group of buildings (a campus), it is becoming common to use cut-through switching rather than bridges. Routers are then used between different technologies and across wide area links. As we shall see, this use of switching or high performance bridging rather than routing has a direct impact on the allocation of

network addresses within TCP/IP systems.

Other than the implications for addresses, TCP/IP is transparent to the use of switching hubs, bridges or routers, or newer technologies such as Asynchronous Transfer Mode (ATM) in the LAN infrastructure.

2.2.3. Laying out LAN cables

Data flows tend to follow organization structure; a single department in an organization may not be located in one building or at one site. If everyone could share the same cable and there was still excess capacity, this would not be an issue. But where the cable is full and must be subdivided, how should we carry out that subdivision?

From some perspectives, LAN and WAN cables should be subdivided and labelled according to the departments that use them, not according to the locations they serve. Bridges and routers should provide the links for the infrequent traffic between departments, not for the bulk of traffic within a department. In organizations where recovery of network costs is important, such decisions are politically important. If one department takes a disproportionate share of the network capacity and drives it to the next phase of expansion, it is much simpler to allocate and justify costs if the network is structured by department rather than by site.

In practice, networks laid out on organizational needs, rather than geography, are not intuitive to manage; training of maintenance staff is more difficult and configuration mistakes are more likely. Most organizations take the intuitive approach and provide an infrastructure firstly for the building and secondly for sites (Figure 2.4a). As traffic grows to fill that structure, it may be necessary to restructure on an organizational basis. But the technical principle remains:

If optimum throughput is the aim, then try to connect directly the various subsections of a department within and between sites, to reduce the number of bridges and routers, and not to minimize the quantity of fibre, copper cable, and separate intersite circuits (Figure 2.4.b).

If optimum reliability and simplified management is the aim, limit the geographic coverage of a single physical LAN and partition the users with bridges and routers to limit the propagation of faults (Figure 2.4.a).

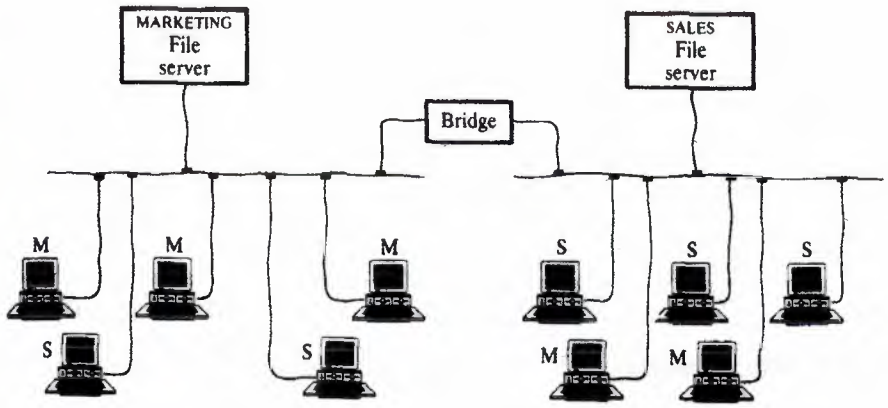

Figure 2.4.a Cabling organized by geography

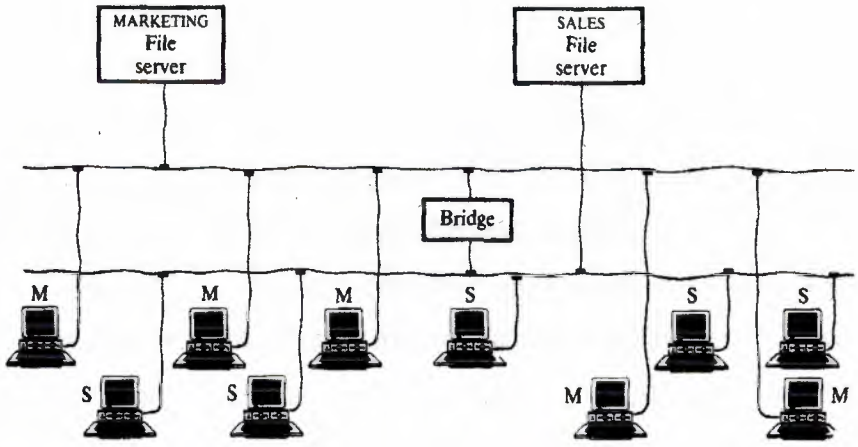

Figure 2.4.b Cabling organized by department

Figure 2.4.c Reference to figure 2.4

The LAN planner needs a flexible set of tools to choose the layout of cabling and which systems are to communicate directly. Developments in Ethernet and Token Ring Switching Hub technology allow the flexible creation of these groups of common interest within floor areas, between floors of a single building and across a single site

using multiple copper or fibre optic backbones. It is worthwhile having a central point, a 'super hub' where the appropriate work areas can be joined into a single physical network with a common backbone fibre. Between geographically separated sites there is little option but to use remote bridges and routers. (Properly designed Token Ring and FDDI can have a degree of resilience to cable breaks.)

2.3. Splitting the network

Bridges and routers partition the network and improve performance for workstations on the same cable. The majority of network traffic should be local to the group of users and a small percentage (15-30%) should cross the bridges onto a backbone network which interconnects different user groups. If this traffic split cannot be achieved, then the bridge is not used in the optimum way.

The prime purpose of a bridge in a LAN is to prevent traffic flowing where it does not need to, rather than to allow traffic to flow where it will. Bridges filter traffic (Figure 2.5). In many networks, it is only a secondary function of the bridge to provide communications where none existed by translating to a different type of cable and extending the cable length limits of the basic LAN specification.

In network partitioning, routers have some advantages and some disadvantages over bridges. They have a more positive role to play. Routers relay only what they are specifically told to relay. They are part of the IP protocol and are designed to interconnect network technologies of different performance. Routers are particularly useful at matching the high speeds of LANs to the much lower speeds of intersite (wide area) communications circuits that are now available. On the other hand they have a higher cost and lower performance and require more careful management than bridges. (Such 'rules of thumb' are not totally satisfactory, for the exact traffic distribution is dependent on the type of the application software and how it was designed to use communications facilities.)

2.4. Using bridges

Bridges filter traffic by Media Access Control (MAC) Address. They must examine each frame received on each port and build a table of frame source addresses with the port on which the frame was received. When the bridge learns that a source and destination are on the same port, it will not pass traffic for that destination to any other cable. The important performance factors of bridges are:

- Filtering rate - the ability to examine frames for possible relaying.
- Forwarding rate - the ability to relay frames which have to be relayed.
- 'Routing' algorithm that allows bridges to be connected redundantly in parallel and in loops but prevents frames circulating round the loops.
- Transit delay - normally related to forwarding rate.
- Filter table size.
- Variations in performance with filter table occupancy.

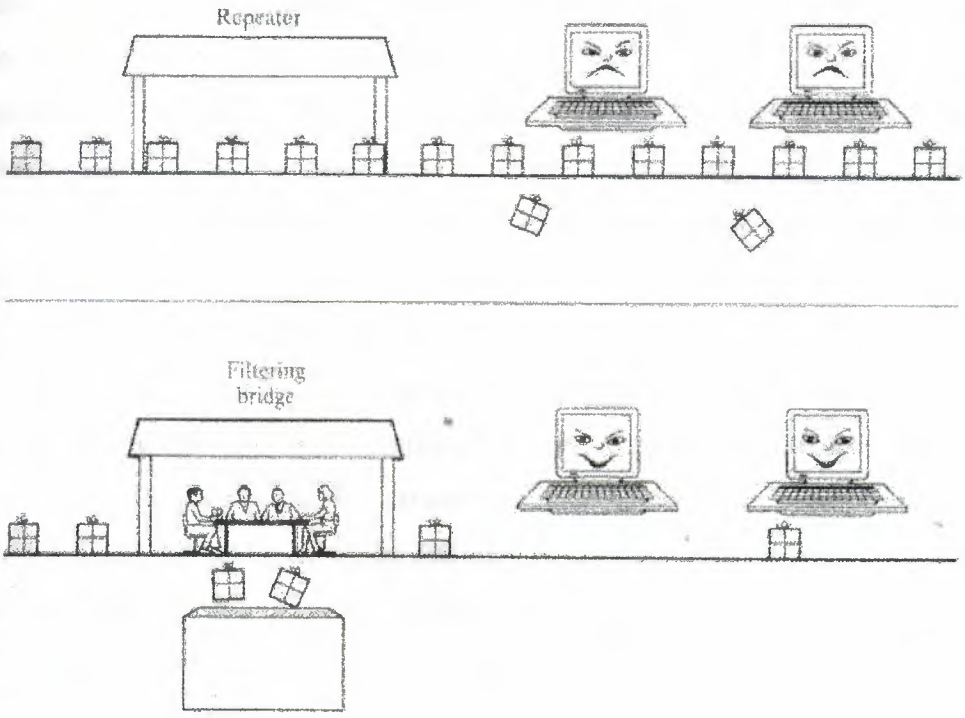

Figure 2.5 Bridges filter traffic.

Bridges now can filter at the theoretical maximum rate for Ethernet of 14880

frames per second. Even in busy operational networks the offered traffic will not be this demanding, for this figure is based on all minimum-sized frames.

2.4.1. Broadcasts, broadcast storms and multicast frames

TCP/IP protocols use two techniques which impact on the performance of the network. They use broadcast and multicast frames for discovering the locations of resources and for communicating between systems which cooperate to provide a common service.

Multicast frames take capacity on the cable but are not processed by every station. They become a problem only on slower wide area bridged links. 'Slow' is a relative term; for today's LANs, 'slow' could be 256 kbps, 128kbps or less.

Broadcast frames are much more detrimental to performance. They are generated by AEP, RARP, BOOTP, and RIP. Every system on an Ethernet or Token Ring network must process every broadcast frame. If the number of broadcast frames increases, the performance of every system will degrade noticeably. A frame takes some time to process, independent of its size. Each broadcast frame can reduce the performance of a PC by almost one normal data frame of 1000 to 1450 bytes. PCs slow down noticeably with 40 to 50 broadcast packets per second.

Bridges relay broadcast frames everywhere. To try to filter broadcasts is to destroy the possibility of any-to-any communication. So networks which are interlinked only with bridges reach a size where the percentage of broadcast traffic from all the devices is unacceptable. Unless the broadcast traffic can be reduced, it becomes impossible to manage the network or to grow it further successfully.

It has been known for in-house application programmers who do not appreciate the potential harmful effects to write applications using broadcasts. On more than one occasion, a complete LAN, including powerful minicomputers, has been brought to a standstill by a broadcast storm of about 100 packets per second. Broadcast storms are sometimes attributed to Ethernet design limitations; it is often not realized that the broadcast storm is usually caused by a combination of bad software design and a poor bridged network structure of any type of LAN and is not limited to Ethernet systems.

2.5. Bridging different technologies

Using bridges to convert from one cable type to another but within a single access

control method introduces no new issues - that is all Token Ring or all CSMA/CD. Bridging between networks with different access control methods from Token Ring to CSMA/CD or to FDDI Token Ring requires more careful planning, as the speeds of operation and hence the frame rate are different (Figure 2.6).

Contrary to popular belief, bridges between IEEE 802.5 Token Ring, IS9314 FDDI and IEEE 802.3 CSMA/CD networks are not transparent to protocol. Bridges must intercept the ARP/RARP protocols of TCP/IP and convert the MAC address representation. Depending on the bridge design this may affect delay and hence throughput. This conversion may limit the maximum number of frames which can be relayed. Do not assume that bridges will achieve the same performance when forwarding between two networks of different types as they do between networks of the same type.

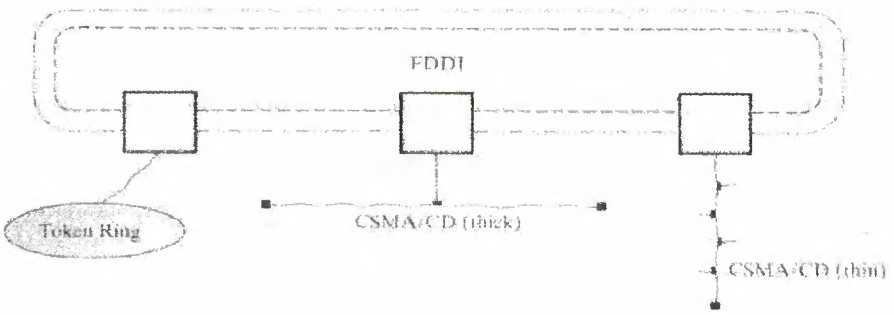

Figure 2.6 Bridging different technologies

2.5.1. Bridging to wide area links

Remote bridges operate in pairs to connect geographically separated sites over long-distance point-to-point links. The main issue is the large difference in speed between LAN technologies and the circuits which are generally available for this type of interconnection (Figure 2.7).

Circuit speeds

The cost of long-distance, high-speed digital services is still regarded by many as prohibitive. In large companies, with many computer users, the cost per user may be acceptable. Even so, few organizations can yet con-aider speeds above 2 Mbps- Outside the USA and UK many organizations find cost and availability of these circuits unacceptable and may be limited to 64 kbps or multiples of 64 kbps perhaps derived

from the newer Integrated Services Digital Network (ISDN) services.

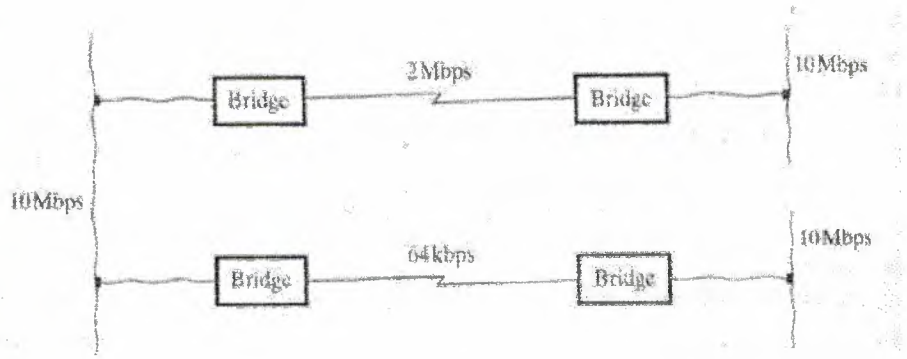

Figure 2.7 Remote bridge

The time taken to send a typical TCP/IP datagram of 1024 bytes at 64 kbps is about 135ms compared to 0.81ms on a lOMbps network. Fortunately, TCP/IP Telnet and FTP were designed to be used in environments with long delays and have time-outs which can adjust to changes in network speed and loading. The performance of these slow network links for file transfers and a system such as NFS, which was designed for the LAN environment, may not be acceptable.

2.6. Routing in bridges

In recent years, the use of the word 'routing' has been extended and applied to any mechanism which can determine and alter paths between end systems (hosts). Sophisticated filtering LAN bridges are often said to have routing features or routing capability. This type of routing does not use the network layer or IP address of TCP/IP, nor does it take account of the features of the 'best path' as requested by the communicating TCP/IP systems.

All IEEE LANs must ensure that all MAC frames arrive in the order they were transmitted and without duplication. Most systems therefore initially only allowed one path between source and destination. This precluded configurations with bridges in loops and in parallel. Such a limitation is unacceptable as a single failure can divide a network in two. The spanning tree and source routing algorithms incorporated in IEEE 802. Id overcome these limitations. For bridge routing, IEEE 802. Id has defined a mechanism called the transparent spanning tree algorithm which is particularly

important for IEEE 802.3 (Ethernet) bridges.

Spanning tree is a disabling technique that suspends the activity of any bridge which provides a redundant path (Figure 2.8). Such bridges carry no traffic but come to life if the current active path fails. The switchover time is of the order of 15-30 seconds if the default configuration of bridges is not changed. This is slower than many computer users would wish but is faster than could be achieved by manual intervention. However, the technique is only fully satisfactory for on-site services. It is often not acceptable to have remote wide area links sitting idle. All the processing for the spanning tree algorithm is placed in the bridge and end systems are unaware of the spanning tree operation. Spanning tree bridges are therefore called transparent bridges.

For Token Ring networks a different mechanism referred to as source routing was adopted by the IEEE 802.5 committee, though it is W>t widely used outside IBM equipment. End systems must be modified to take an active part in setting up a route across redundant source rout-teg bridges. IEEE 802.Id later adopted source routing as part of the IEEE bridge routing standards, though the detail is different from source routing as used by IBM.

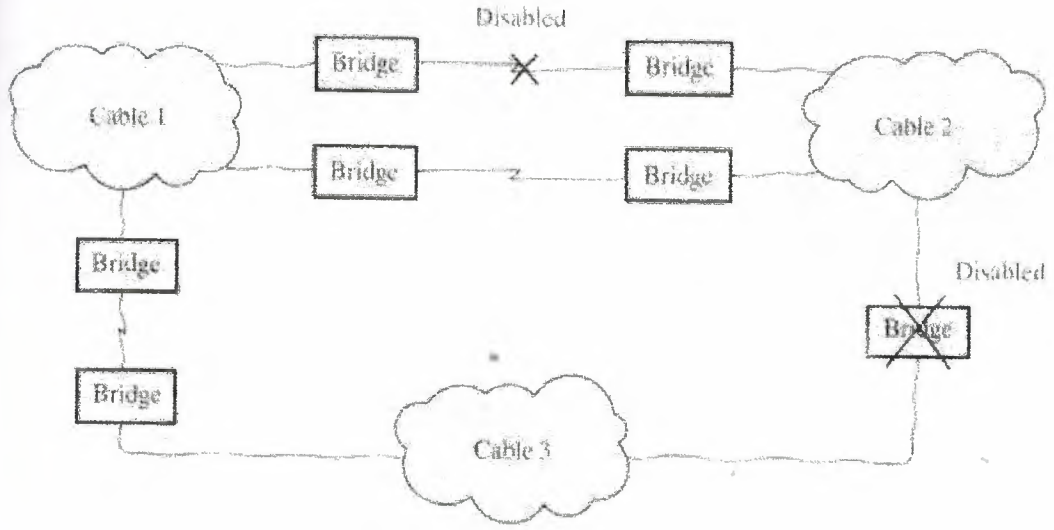

Figure 2.8 Spanning tree operation

An advantage of source routing (Figure 2.9) is that it allows multiple active routes between the source and destination cable systems, one conversation will follow one route for its duration. The main disadvantage is that LAN driver software must be

modified to add and remove source routing fields from the data frames. Attached workstations begin by discovering possible routes between source and destination using either a single route explorer frame or an all routes explorer frame. Transmission of the single route frame relies on the fact that a single stable spanning tree route has been set up either manually by the network managers or automatically by a spanning tree algorithm. Having discovered all available routes, end stations insert a list of bridges which are to be used into each frame they transmit. If the link fails, then an alternate route may be selected instantly from a list maintained in the workstation.

Source routing is only likely to be found with TCP/IP end systems on Token Ring. It is not normally used on Ethernet.

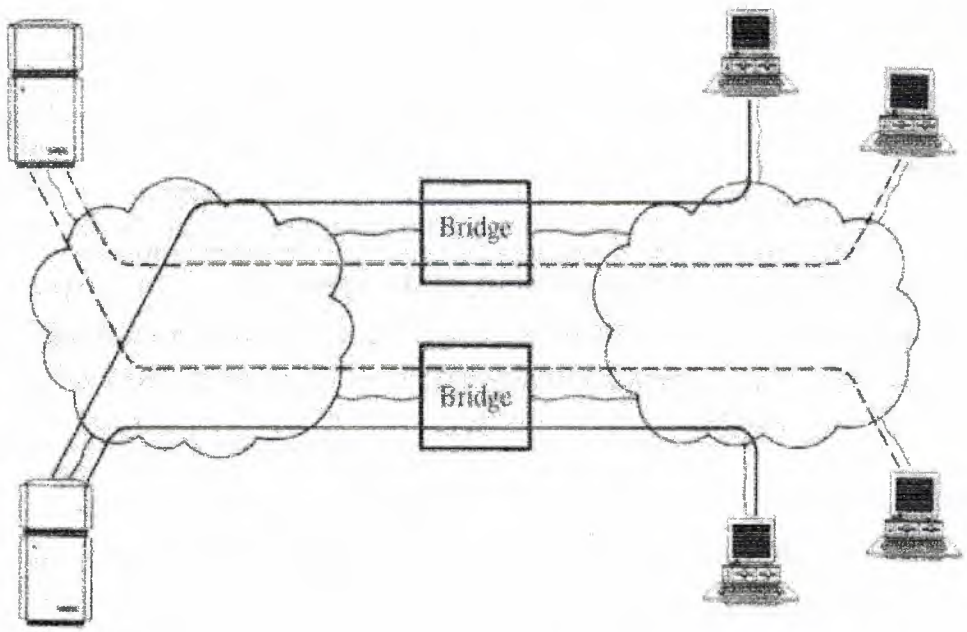

Figure 2.9 Source routing

Before choosing a bridge, managers should check that the implementation is fully IEEE 802. Id compliant, and that interoperation of that implementation of spanning tree has been checked with any other implementation that is used in the network. They should also be aware of how the spanning tree algorithm operates, and how bridges determine if they are to be in the active topology. For effective operation, the spanning tree algorithm must be managed. This means that network managers must change the default configuration of their bridges.

In networks which use bridges between Token Ring and Ethernet, the network designers must have a plan of how the conversion between source routing and spanning tree is to take place in a consistent manner.

2.7. The limits of bridging

While these algorithms can have a place in smaller networks, spanning tree in particular is often not satisfactory for the larger network with WAN interconnections, since spanning tree disables otherwise useful routes. This is one reason why bridged networks can only grow to a limited size (Figure 2.10). Another is that the standards allow for only seven bridges in series.

But there is another limitation which we have already hinted at. Bridged networks, even when well structured, can grow to the point where the volume of broadcast and multicast traffic degrades performance, where widely varying loads on intermediate (backbone) networks affect performance because of packet loss, where spanning tree and source routing algorithms provide less than optimum routing and where the problems of conversion between Token Ring and CSMA/CD technologies are all too evident. The performance of the network in such circumstances is at best poor. Network users may see wide variations in both reliability and response. Uncontrolled congestion prevents the successful transmission of data. For some applications, the network may become unstable and unreliable.

Various authorities have tried to put a figure to the size of network that can be successfully managed with bridges alone. It depends on many factors other than technology. But many organizations find that when their network has a combination of many hundreds of interconnected devices, a few tens of servers and minicomputers, and ten to thirty bridges, the control of broadcast traffic can become impossible and the overall reliability and performance of the network decreases. The problems are made worse by having remote bridges with slow-speed links between sites.

In a poorly performing network, the network planner must carry out a more careful analysis of traffic volumes and flows and then take remedial action to alter network structure and the positioning of bridges to meet the corporate needs. If to these problems is added the possible inappropriate use of the network by badly designed software that propagates faults throughout the system, an alternative interconnection method that gives the network manager more control is required. Bridges do not

normally limit the effects of excessive broadcast frames from poorly designed software.

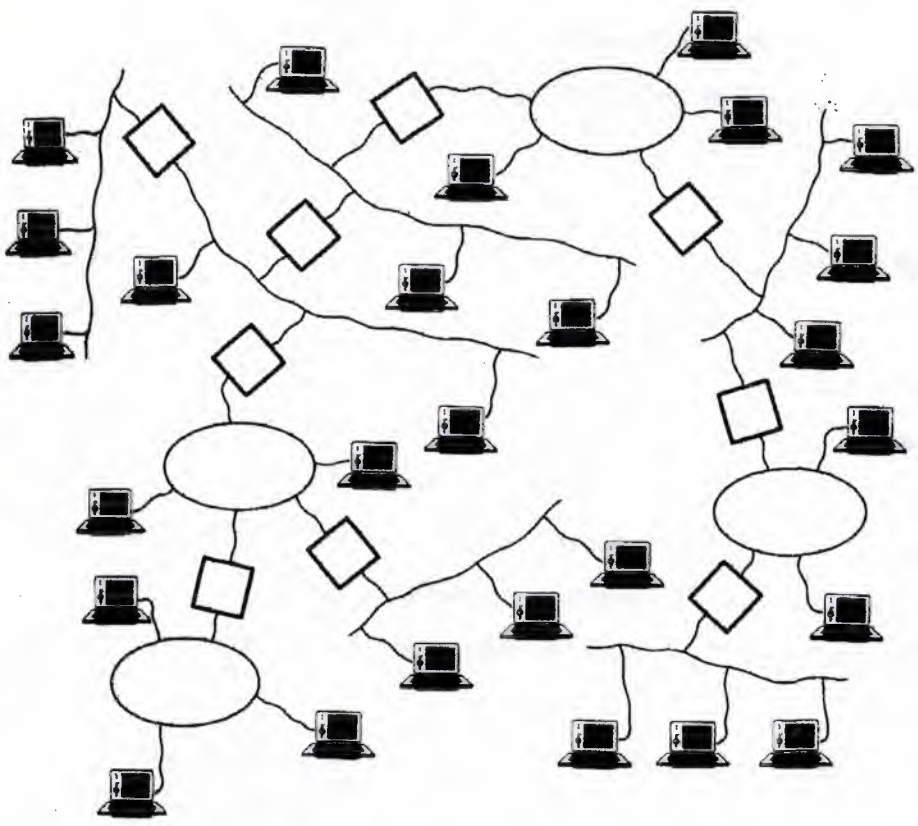

Figure 2.10 The limits of bridging

2.7.1. Using routers

In most communications architectures, routing is a network layer function, not the MAC or data link function performed by bridges. Network layer routing predates bridging; routers were in use in WANs before the link-level LAN bridge was introduced in the early 1980s by companies such as Sytek Inc. and by the Digital Equipment Corporation (which holds patents related to bridge algorithms).

During the mid and late 1980s, routing was eclipsed by an enthusiasm for bridging. It is only as the performance, stability and management limitations of large bridged networks using connectionless protocols have become apparent and the cost and performance of routers has improved, that routing has once again come to the fore

in larger networks.

Routers do not represent a panacea for the problems of the network manager. They can usefully replace some bridges in a bridged internet which is already physically well structured; it is likely to be a waste of time and money to add them to a network which is not well structured, or where the traffic flows are not understood and well managed.

Routers only operate properly if the network addressing structure is correct. We therefore delay a full discussion of router technology until after we have described how to plan an IP network address structure.

CHAPTER THREE:PLANNING AND MANAGING IP ADDRESSES

Careful network planning is key to the long-term success of any, networking technology, particularly for the larger installation. In the past, sufficient consideration has been given to planning the key parameter of IP addresses for TCP/IP. This leads to one of the commonest causes of escalating network support costs.

The focus of this chapter is how to plan and manage IP Version 4 addresses. Implementors of TCP/IP must select, configure and manage IP addresses correctly. Selecting an IP addressing scheme is the first major decision the TCP/IP implementor must make.

An IP address identifies the connection of a machine to an IP network uniquely, and in a way that is independent of the underlying network technology.

Since 1990, the IETF (Internet Engineering Task Force, a technical body under the IAB) has studied methods of overcoming some limitations in IP Version 4. The chosen solution is known as IP Next Generation (IPng) or IP Version 6 (IPv6).

3.1. Identifying a network connection

Connections to a TCP/IP network are often known to each other and to users in three different ways, all of which identify the same connection to the network but at different levels of TCP/IP:

- As names: in the fully qualified host and domain name for example, www.integralis.co.uk

Using the domain name scheme is often not considered early enough in the planning of commercial networks; users of larger networks will find it invaluable.

• As a group of four numbers separated by full stops (periods). This is the IP or network address used by the IP layer of TCP/IP

for example, 193.128.143.1

• On a shared LAN by its network interface card address, the 48-bit MAC address (Medium Access Control address, often called the physical address)

for example, 0060 8C 12 34 56

Why have three levels of address, since they all refer to the same connection to the TCP/IP Internet? The reason is that while at any one time these addresses identify the same location, over a period of time each may change for different reasons. The three types of address allow a degree of flexibility where none would exist if the functions were not separated:

• The host and domain name should remain unchanged for the longest time, for it is completely determined by the system designers and managers, though it will often change if there is a major change in machine location.

3.2. Planning the IP address space

The IP address may change due to network growth and reconfiguration, for example if a machine is moved from one location in a building to another.

The MAC address may change for similar growth or performance reasons, or because the network card is replaced following a card failure. The shortest lived address is likely to be the MAC address.

The fully qualified domain name is for 'human consumption' and consists of a set of identifiers, separated by full stops, which describe a computer in a hierarchical relationship to all others. For example:

-machine.department.site.organization.org_type.country

Introducing flexibility brings with it the possibility of error and communications failure if the flexible functions are not managed and controlled correctly.

To plan an IP addressing scheme you must have or acquire the following information:

• The maximum number of host ports that your organization could ever wish to interconnect. The whole of your organization is not just your particular department, but the whole organizational structure, if necessary, worldwide, which may at any time be interconnected.

• An understanding of the numbers of devices at each location and in each uilding on each site. Think big. Assume one address per member of staff in those uildings and leave some 25% spare.

An understanding of the different departments, their geographic locations, and the ikelihood that they will need to communicate. An estimate of the amount of nformation that is exchanged is useful but it has a very short-lived value.

A knowledge of any other current users of TCP/IP and any standardization ecisions that have been taken already.

.2.1. Internet protocol

The Internet Protocol (IP) may be described as the network layer protocol of CP/IP. IP is operated and interpreted by each intermediate relay in an interconnected et of LANs and WANs which is using the TCP/IP protocol suite to communicate. Such collection is often called as internetwork and the intermediate relays are routers. IP rovides best efforts delivery service based on a technique called datagram ransmission; 'best efforts' and 'datagram' because no attempt is made (by IP) to recover ny errors which may occur in transmission.

IP supports routing and relaying of information between corn municating hosts end devices) according to the Type (or quality) 0 Service (TOS) they require. IP allows outing errors to be trapped an< reported and their effects on system performance to be inimized. II does not make assumptions about the underlying network of physical ables, LAN hardware (Ethernet, Token Ring or FDDI) or the point-to point wide area inks (PTT circuits).

.2.2. IP addressing

The IP address is one component in the Internet protocol. The IP address; is a umber that uniquely identifies the connection of a host computer (01 end system as SI would call it) to a physical network as it communicates with other computers (or nd systems). Hosts with more than one connection have a different IP address for each ne. IP routers also have their own IP addresses for they can be the source and estination of II datagrams.

The purpose of the IP address is to identify each connection to the internetwork in way that is independent of the underlying physical network (LAN or WAN) echnology, and to collect a group of connections together to simplify routing.

Internetwork routers use II addresses to make routing decisions.

3.2.3. The need for address management

IP addresses must be unique in the communicating internetwork. If not communications will fail erratically. It is up to network managers to configure and manage IP addresses correctly by traditional manual method aided by database and network management technology or by using the Dynamic Host Configuration Protocol (DHCP), or BOOT Protocol (BOOTP).

Where TCP/IP systems are connected to a shared cable such as a Ethernet or Token Ring LAN, communication between stations mm take place using MAC addresses, normally built into the network adapter card. Since the IP address is configured independently of the MAC address, it remains unchanged even if the network card fails and is replaced.

Every Ethernet or Token Ring card comes with a preconfigured 48-bit MAC address, its universally managed address. LAN standards also allow locally managed addresses; network cards can be loaded with a 48-bit address chosen by the network manager. Locally managed addresses are not normally used in TCP/IP systems. As IP addresses in TCP/IP must be controlled and managed, there is little value in adding a second layer of management by also configuring and controlling MAC addresses.

But the issue of MAC address management does not disappear completely from the TCP/IP story; some of the facilities of TCP/IP (such as the Boot Protocol, BOOTP, and DHCP) use the MAC address as a fixed reference point for obtaining other information. If you wish to use these features, you must at least have a record of the MAC address of each computer for which the facility will be used. These addresses appear as a reference point in look-up tables in information servers. But, as described above, the MAC address can change if the network card is replaced. In a large network, keeping an up-to-date record of MAC addresses and ensuring that all servers have the correct record can be time-consuming.

Some UTP hubs have management facilities that can record the MAC address attached to each user drop cable (and can deny access if that LAN address changes to an unauthorized value). Such systems make it easier to record the location of particular MAC addresses.

3.3. Characteristic of the IP address

The IP address is a 32-bit number which must be unique in the internetwork. Devices which must have an IP address include every connection of each computer and of each network router. As mentioned previously, computers with more than one network connection have a different IP address for each connection (Figure 3.1). In TCP/IP literature, such computers are referred to as multihomed. These connections are usually on different networks. This is the reason for sizing the network-by-network connections rather than the number of computers, though these numbers will often be very similar.

Unlike some other network addressing schemes, the IP address does not necessarily convey any information about geographic location. Given an IP address you can only deduce a management authority; that authority could manage one corporate network. This has interesting implications for routing between two worldwide networks. Put another way, the IP addressing scheme is not hierarchical (unlike the telephone or telex network or for that matter the CCITT X.121 addressing scheme for the X.25 interface to public packet switched networks).

Network managers can choose to impose a geographic or organizational hierarchy on IP addresses but this hierarchy is not a function of Internet standards.

3.3.1. IP address format and components

The 32-bit IP address has two components: a network number (or network identifier) and a host number or host identifier. Some, 'class bits' can be extracted from the network number, but trying to examine these on their own can lead to confusion. It is safer always to think of the network number as containing the class bits.

It is the network number which identifies the controlling organization and the host number which identifies the particular connection within the authority of that organization. The term 'host number' is historied and this name could cause confusion for, as we have seen, a multihomed host has more than one IP address and hence more than one host number.

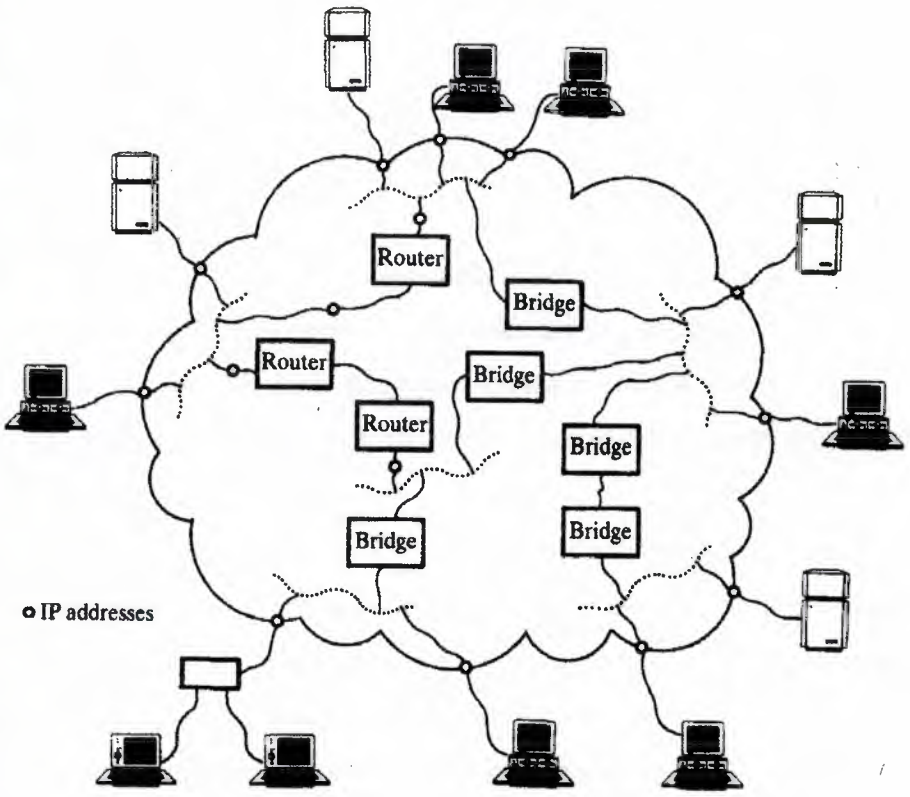

Figure 3.1 Allocation of IP addresses.

For reasons that will become apparent later, it would be unusual for two ports on the same machine to have the same network number.

Table 3.1 The limitations of each class of IP address

Network class	Max. number of network numbers	Max. number of hosts on each network number
A	126	16 777 214
B	16 382	65 534
C	2 097 150	254
D	not applicable	not applicable
E	reserved for multicast systems-Reserved for IAB use	reserved for multicast systems-Reserved for IAB use

The terms net id and host id are in common usage for network number and host number.

3.3.2. Classes of address

IP addresses are divided into five address classes (A to E). It is important to understand that these divisions were conceived to ease the management of addresses by the Internet Architecture Board (IAB); they have less immediate significance for the network manager.

The first three classes, A, B, C, are available for normal allocation and for host-to-host communications. Class D and class E addresses are not for general use; class D is reserved for special use by IAB designated protocols and class E is reserved for future use. One use of class D addresses is by routers, so network managers may encounter them while monitoring router protocols.

There are no practical distinctions in the way in which computer systems with class A, class B or class C addresses use those addresses. With some noted exceptions, which are discussed further below, any address from any class is equivalent for communications purposes. In principle, any host can communicate equally well using an address from any of these three classes. The distinguishing feature of each class of address is the number of network numbers and the number of host connections which each network number can support. The limitations of each class are shown in Table 3.1.

Most implementations give network managers full control over the IP addresses they allocate and configure into the equipment. It has been known for an implementation not to accept anything other than a class C address. There are situations where this could be unsatisfactory.

3.4. Network numbers and host numbers

The choice of network numbers is most important.

• Network connections with the same network number should communicate directly on the same physical LAN.

• Network connections with different network numbers do not communicate directly; they must use the services of a router. This router either is directly connected between the two LANs to which the hosts connect or it in turn knows of a router which can relay the message towards its ultimate destination (Figure 3.3).

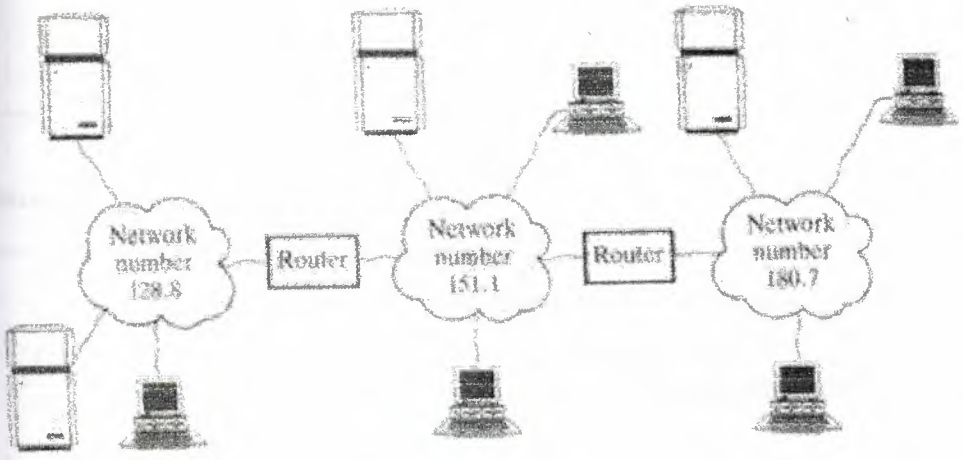

Figure 3.3 Choosing network numbers

These factors determine the basic rules for choosing network numbers and host numbers. They may alternatively be stated as:

• Computers or workstations separated only by bridges or repeaters will have the same network number.

• Computers or workstations separated by routers have different network numbers.

For the network implementor, the most important factor is that the choice of IP address for a particular host is affected by the presence of routers. The positioning of routers is determined by the exact geography of the underlying physical networks which connect computers, floors, buildings, sites and countries into a complete internetwork. It is also heavily influenced by organizational needs, traffic flows and traffic volume. The layout will change as the network grows', and matures. If a bridge is replaced by a router, the network number on \ one side of the new router must change.

When combined with the limitations of Table 3.1, the maximum i number of network connections that are likely to communicate within one j network (or organization) is therefore a key to determining which address class should be selected.

There is a mechanism for dividing one network number into subnetworks. In practice, hosts on either side of a router must be in different subnetworks rather than within different network numbers. However, if they do have different network numbers, they must communicate via one or more routers.

Table 3.2 IP address representation

Dotted decimal	Dotted hexadecimal	C-style hexadecimal	Binary
44.123.110.224	2C.7B.6E.E0	0x2c7b6ee0	00101100011110110110111011100000
129.6.48.100	81.06.30.64	0x81063064	10000001000001100011000001100100
128.240.1.109	80.F0.01.6D	0x80f0016d	10000000111100000000000101101101
192.33.33.109	C0.21.21.6D	0xc021216d	11000000001000010010000101101101

A TCP/IP addressing scheme evolves as the needs of users are better understood during the design process. Adding a TCP/IP addressing scheme to a LAN must take into account the existing layout of that LAN.

One factor is key. Network planners must attempt to design an addressing scheme that can remain substantially unchanged in a changing environment.

3.4.1. Writing down the address

The convention is that the address is written down, or described to software, in dotted decimal notation. Each eight bits of the address (each octet) is converted to a decimal number in the range 0 to 255, and separated by a dot (.).

While the US standards initially specified 'dotted decimal notation' some computer staff are more at home with hexadecimal notation. Dotted hexadecimal (and UNIX or C-style hexadecimal) and octal notations are sometimes used and will be accepted by some (though by no means all) implementations. Occasionally it is useful, if somewhat long-winded, to represent these numbers in binary. Some valid addresses are shown in Table 3.2. Leading zeros in each octet, which have no significance, should not be included. 128.1.0.9 is valid and more usual than 128.001.000.009. However, 128.1.9 or 128...9 are not valid. (Some systems interpret numbers with leading zeros as 'octal' numbers.)

3.5. The IAB and network number registration

Since addresses in an internetwork must be unique, there has to be a registry which ensures that such a policy can be enforced.

TCP/IP was developed for use on the Internet. This collection of networks is managed by many different authorities, with the IAB as overall design authority. In this environment of devolved management, the network number has two functions. First, since each IP address on the whole set of interconnected networks must be unique, the network number is used to identify one addressing authority (often called an autonomous system) responsible for issuing unique connection ids in that part of the address space. The second function is to support routing.

The IAB ensures that IP addresses on the Internet are unique, by registering and issuing network numbers to organizations that have reasons to connect to the Internet. Such organizations do not choose their own network number it is issued to them. They agree to undertake

Those who have followed the text carefully may have noticed that an organization appears to require use of many network numbers if routers are to operate correctly. The IAB will issue at most a few different network numbers to one organization, which will not, on its own, apparently satisfy the requirements for a large router network. The alternative is subnetting of a single network certain network management responsibilities when they request an official IAB address allocation. They must then issue host ids so that they are unique within their own network number address space. If they fail to issue unique host numbers, only the services of that organization will be affected.

The IAB has devolved the clerical task of registering and issuing IP network numbers to the Internet Registry, part of the Internet Assigned Numbers Authority (IANA). The Internet Registry is operated by the Department of Defense Network Network Information Center (DDN NIC), now located in Chantilly, Virginia. The NIC has devolved address management to RIPE (Reseaux IP Europeens) in Europe, to APNIC (Asia-Pacific Network Information Center) for the Pacific region, and to Internet service providers.

3.5.1. To register or not to register

For any commercial organization, applying to a registration authority for a registered network number ensures that your IP addresses are unique in the whole world (of those who have similarly registered). But as Table 3.1 showed, the IP address space, other than class C addresses, is limited with only 16408 'large' networks. Class C gives a further 2.1 million small networks. As worldwide interest in TCP/IP has grown, the

address space is already under some pressure with much of the class B address space already allocated and all the class A address space issued or reserved. There has to be good reason for allocating one or more registered class A or class B addresses from this finite resource.

In early 1995, the total number of hosts accessible from the Internet through the name service was 4.852 million. As not all registered addresses are connected to the Internet community, this is a very conservative number. Growth is already onto the vertical portion of an exponential growth curve.

When should an organization apply for an IP address registration? At least one Request For Comment (RFC) recommends that all users of TCP/IP should register. We endorse that policy but the following gives guidance on which organizations must register, in decreasing order of importance:

- Any organization that connects to the worldwide Internet has no option; they must request a registered network number.
- Any organization that wishes to communicate with a parent organization which is itself connected to the Internet is advised to register. Accidental duplicate addresses advertised to the true Internet do not make the advertising organization popular with the Network Operations Center.
- Organizations that have subsidiaries or that acquire subsidiaries which connect to the Internet should consider registering.
- Large multinationals that wish to use TCP/IP for worldwide communications but which perhaps have limited control or knowledge of the activities of individual operating companies should consider registering.

For those who do register there are additional issues to resolve:

- Larger organizations, with many thousands of employees worldwide, will prefer a class B or class A address. All class A addresses are allocated or reserved; recent allocations have been to governments outside the USA. Class B addresses with 65 534 connections are difficult to obtain.
- As we analyse the use of these IP addresses in TCP/IP, you will discover that the address space is consumed more quickly than might be expected; in practical networks, it often is not possible to use the available space efficiently. Large organizations will then require more than one class B address to satisfy their requirements.
- Most organizations that register will be forced to use one or more class C

addresses. Unfortunately, class C addresses cannot have more than 254 connections on one bridged or switched LAN. This is unacceptable for those organizations that wish to use large port switching LANs within one building.

3.5.2. Advantages and disadvantages of registering

The key advantage of registering is ensuring that as you use and extend TCP/IP to new applications, which may include communicating with external organizations using TCP/IP, your address and naming conventions are protected. If you are not registered and an address clash occurs in the future then the onus must be on the nonregistered organization to make changes.

Those wishing to pursue registration further should contact the Network Information Center, an Internet service provider or, in Europe, the IP Information Centre (RIPE).

3.5.3. How to choose your own network number

If you decide that registration is unnecessary, then do consider carefully which address you choose.

The maximum likely size of the network is the most important factor which will influence the choice of IP address class. Private self-contained networks may use any or many different network numbers, though there are good reasons to limit the range of network numbers used.

It is important to keep IP addresses unique in any internetwork of intercommunicating TCP/IP machines. As networks grow and develop and as TCP/IP is used to communicate with other external companies (for example for Electronic Data Interchange, EDI), this will become more difficult if the network number is not registered and a large number of network numbers are in use. An address is more likely to be unique if the network number is registered. If your organization does not fall into that group which should register, then to reduce the likelihood of an address duplication with another organization in the future you should:

- Avoid class A addresses, for they are all allocated to the largest networks to which everyone connects.
- Do not copy examples given in manufacturers' handbooks, textbooks or sales literature.
- Avoid well-known addresses published in TCP/IP RFCs and other

documentation.

• Do use one, or a limited number, of class B addresses which are registered to organizations you never want to communicate with.

• Do use a limited number of class C addresses chosen in a similar way.

• Do use one or more of the network addresses reserved for private networks by RFC 1597.

If at some point you wish to communicate with organizations with which your addresses clash, or with the Internet, then all is not lost. A 'gateway' that performs address translation between the two networks will map addresses from one network to the other in a static or dynamic fashion.

3.6. Autonomous systems

An autonomous system (AS) is a collection of LANs and WANs and the interconnecting routers which are under the control of one management authority for configuration, addressing, naming and routing decisions. Decisions made within such an authority should not be visible to the rest of the interconnected Internet. Ideally, one commercial organization should form one autonomous system. The IAB also registers AS numbers that uniquely identify an autonomous system.

The concept of autonomous systems is extended in the Internet to the Autonomous Confederation, a group of autonomous systems with a common interest which requires them to establish direct links to each other rather than route through the core Internet. The term autonomous confederation has a less clear value to commercial organizations unless they cannot achieve a single centralized control authority. This term seems to be falling into disuse with changes in Internet policy.

Autonomous system numbers

As well as registering a network number, commercial organizations should now register an AS number. The AS number is used by routers with the newer OSPF and BGP routing protocols to identify which AS is the source of routing information. This forms part of a low-level security facility.

Some organizations have registered more than one AS number. AS numbers can only distinguish 65535 different organizations so, in due course, this could become another scarce resource, with the rapid growth in interest in internetworking with TCP/IP.

3.7. Private network addresses

In March 1994, the IAB recognized the need for IP addresses for private networks, that is, for those networks not connected to the Internet. In an advisory document known as RFC 1597 (Address Allocation for Private Internets), they reserved three address ranges for use in private internets. They are:

10.0.0.0	- 10.255.255.255	1 Class A network
172.16.0.0	- 172.31.255.255	16 Class B networks
192.168.0.0	- 192.168.255.255	256 Class C networks

This RFC is an informational document only - it is not an accepted IAB standard.

Where companies that use these addresses become Internet connected, they must obtain one or more registered addresses and use address translation routers so that a registered address is seen on the Internet. Such techniques are now well proven and have considerable security advantages for commercial organizations where it may be undesirable that every workstation can be seen from the Internet.

RFC 1597 was not universally applauded; in July 1994, RFC 1627, [Network 10 Considered Harmful (Some Practices Shouldn't be Codified)] was issued. The authors of RFC 1627 argue that the deliberate duplication of addresses proposed in RFC 1597 is restrictive, reduces network reliability and security and will be costly in the long term. RFC 1627 argues that IP addresses can become embedded in software, and that changing IP addresses throughout an organization is costly; by adopting the address ranges of RFC 1597, an organization inevitably has an address clash in due course and must change addresses. It also argues that it is impossible to predict which machines must be visible from the Internet and hence the basis of using registered or unregistered addresses is unsatisfactory. It argues that the IAB should continue to promote only the use of registered addresses, even where those systems are not at first Internet connected.

3.8. Configuring the IP address

Every computer, workstation, personal computer, network router and network management station which 'speaks IP' must have a different IP address for each of its connections. The IP address is entered through the normal configuration management processes. IP addresses are usually represented and entered in dotted decimal notation; occasionally dotted hexadecimal notation is accepted. The address is usually stored in

nonvolatile memory, either RAM or on a disk file when available.

Diskless workstations often have no nonvolatile storage. They obtain their IP address across the network at start-up time using BOOTP or RARP.

Visiting each device to change the IP address is time-consuming, particularly in dispersed networks. BOOTP and DHCP provide mechanisms for centralizing the distribution of IP addresses and other useful TCP/IP parameters, even for machines which do have storage. The key to distributing the correct IP address and other information via BOOTP/ DHCP is the MAC address of the interface card. It is then essential to record accurately the MAC address used by each machine on the network that uses the BOOTP service. This is excellent policy for fault diagnosis and security in any case, though it is often neglected.

BOOTP uses a static table to map MAC address to IP address. As its name suggests, DHCP can dynamically allocate IP addresses from an unused pool of IP addresses to new requesting workstations.

In an operational network, IP addresses become recorded in distributed tables and in software in machines. While this is bad practice, it is unavoidable. Frequent wholesale changes to IP addresses must not Occur, so preplanning is essential. Such changes are labour-intensive because of the widespread distribution of addresses; the change introduces errors and has unpredictable consequences for the integrity of operational applications and for the time taken to restore any-to-any communications.

3.9. Reserved IP addresses

Certain IP network numbers and host numbers are reserved for the use of particular aspects of TCP/IP communication. If you configure a connection with a reserved IP address, the faults caused are likely to be obscure, apparently intermittent and difficult to isolate. Not all host software checks (or is even able to check under all circumstances) that the IP address it is requested to use is valid.

The following addresses are reserved:

- Network number 127.X.X.X, where X.X.X is any set of numbers. This is used for a local software loopback test.
- A network number of all Os is classless and means 'my current network which I do not know the number of - sometimes referred to as 'this network'.
- A network number of all Is with a particular class.

• Host number 0 is reserved to refer to a particular network number. For example, 192.0.1.1 is the first class C network address which could be allocated.

• Host number 'all Is' is reserved to broadcast to all hosts on a specific network; it can only be used as a destination address.

• The full address 0.0.0.0 is reserved. It is used in two ways: as a source address when the host does not know its genuine address, for example during bootstrapping of a diskless workstation, and by routers in a list of addresses to advertise the default route, the route to all networks which are not explicitly listed. (It is not a source or destination address, merely an entry in a table.)

• The full address 255.255.255.255 is reserved as a destination address to mean 'broadcast to all hosts on my network'. 0.0.0.0 as a destination is an obsolete form of255.255.255.255.

Further addresses become reserved when subnetworking is introduced. These are more difficult to recognize.

-IP addresses with routers and dial-up devices

Some older router implementations require that a point-to-point wide-area PTT or dial-up link between two LANs should be identified with its own unique IP network number with only two connections attached. In the larger network this is wasteful of the limited resource of IP network numbers. Check with your router manufacturer that the most modern techniques, which conserve network and subnetwork numbers, are used in their implementation.

3.10. Common mistakes in choosing IP addresses

Common mistakes in choosing and managing IP addresses centre around the possibility of address duplication or using the wrong network number for the LAN to which the port is attached. This arises through the following:

AB standards say that any address beginning with 127. must not be transmitted across; the physical interface and appear on the cable. Not all software implementations check and, reject IP addresses beginning with 127 and some will transmit IP datagrams with that source or destination address. This has been known to cause networks to fail.

• Not controlling the issuing of addresses (and subaddresses) centrally, or at a sufficiently high point in the organizational tree.

• Using the example in the software manufacturer's literature.

• Giving a supplier the responsibility of choosing and configuring an IP addressing scheme without understanding the overall needs of the organization.

• Copying the addresses from another part of the organization.

• Duplicating the address from the machine next door.

• Choosing an address from a nearby machine and modifying it slightly.

• Not fully understanding the concept of network and host number.

• Misunderstanding that workstations on the same LAN must use the same network number.

• Misunderstanding that stations either side of a router must have different network numbers.

• Using a reserved network number or host number.

• Failing to anticipate how and why IP addresses may change at a later date and avoiding that change.

3.11. The organizational structure and the IP address

Allocating TCP/IP addresses in larger companies raises political and organizational issues. If no scheme for address allocation has been developed and agreed, the first people to introduce TCP/IP may find themselves attempting to take addressing decisions on behalf of the whole organization, but without the necessary authority and influence to carry their decisions through.

If you are in that position, you either take up the challenge of agreeing the address structure for the whole organization or you take the pragmatic approach to make a local decision and trust that it will not cause problems with address clashes or major updates later. Be assured that it will. In a changing environment, this pragmatic decision will eventually rebound in increased catchback costs, for, even if address clashes do not occur, the address will change as routers are introduced in the growing network.

The second organizational issue of IP addressing is that IP addresses are entered into the software of applications computers. Network routers also use IP addresses to communicate among themselves and with host computers. In large networks, network routers and application processors may be controlled by different managements. The addressing authority must issue IP addresses, but they may not be welcome if they attempt to configure them in the larger mini and mainframe application processors. The central authority must be able to enforce IP address configuration and even

reconfiguration decisions equally on the owners and operators of host processors and network routers. Fortunately, once addresses have been allocated they should not change frequently, if the network has been well designed. But over the life of even a well-designed network, it is to be expected that some address reconfiguration will be necessary because of activities in another department.

3.11.1. Moves and changes

It is a feature of life in many organizations that staff move office location regularly. If their PC or workstation at the new location is in a different network number or subnetwork number, the IP address must change. This may have a consequential impact on other levels of TCP/IP, the domain name service, mail, and on other parameters of the machine's software (default router, BOOTP server, domain name server).

3.11.2. Record keeping

If TCP/IP management is to be successful, there must be records of each IP address and where, how, and by whom it is used. Some features of TCP/IP require a record of the Ethernet or Token Ring MAC address associated with the IP address. BOOTP and DHCP supply the IP address for a specific MAC address; domain name service supplies an IP address given the host domain name, or, given an IP address, will supply the host domain name.

To keep these records up to date requires proper clerical procedures, training of clerical staff and the correct vetting of entries by software, so that the task becomes routine. The aim of system managers must be to reduce the amount of skilled time consumed by such clerical activities.

3.11.3. Subnetting

The network manager has an immediate need: to solve the issue of net-1 work numbering on either side of a router when only one network num-1 her can be used. The solution, introduced in 1985, is called subnetwork addressing and it is now standard on all TCP/IP equipment.

3.11.4. IP Version 6 addresses

The IPv4 addressing scheme appears to give access to over 2100 000 networks with a total of 3720 million hosts. This seems a large number even given the number of PCs in use. Not all of them are yet using TCP/IP. But because of the strict divisions between network numbers and host numbers, the address space is now under considerable pressure. For example, an organization may have a class B address for 10 000 hosts and the other 55 000 host addresses are effectively lost to the Internet community.

The advent of the personal computer running TCP/IP, with each requiring an individual IP address, coupled with the worldwide success of the IPS, has revealed the limitations of IPv4 addressing. For example, in the mid-1990s, allocations of registered addresses to large organizations have been of multiple class C networks because registered class B network addresses are now a scarce resource.

Starting in 1990, the IETF studied methods of overcoming these limitations. By late 1994, it had agreed the 'IP Next Generation' (known as IPng) for the long-term solution to addressing and other technical limitations of IPv4. This version is known as IP Version 6 (IPv6).

IPv6 uses 128-bit addresses. A design aim of IPng was that it should support at least 1000 million networks each of many addressable devices, an aim which is met in excess by 128-bit addressing. The concept of network number and host number remains for routing and for addressing of individual machines. However, there is much more structure in the way network numbers are used, leading to hierarchical routing and a significant improvement in the efficiency of the worldwide communications system that the IPS has become.

We consider three features of IPv6 to be particularly important:

(1) Workstations will be self-configuring, determining their own network number and unique host number. This reduces the costs of installation, moves and changes.

(2) Individual IPv6 systems can be installed, or IPv4 upgraded to IPv6, while continuing to interoperate with IPv4 systems. IPv6 will therefore support a phased implementation while coexisting with IPv4 systems.

(3) A single IPv6 network number can support an unlimited number of hosts. (The concept of Class has disappeared.) Large networks without routers are essential if the maximum benefit is to be gained from LAN switching hubs and Asynchronous Transfer Mode (ATM) transmission systems.

The 128-bit IPv6 address can contain an IPv4 32-bit address. By adding a unique registered address modifier to an unregistered IPv4 address, that unregistered address can be used on the Internet without other modification.

Other features allow for automatic configuration of mobile and portable systems which may be temporarily connected anywhere in the internetwork.

As requirements for these IPv6 systems emerge, a network manager must plan IPv6 addresses and upgrade routers in the network to support the two Network Layer Protocols simultaneously - IPv6 and IPv4. Routers can relay information between older systems which understand only IPv4 and newer systems which operate to either IPv6 or IPv4. In time, though it may be a long time, IPv6 is likely to replace IPv4 as the single network layer protocol.

CHAPTER FOUR: SUBNETWORKS AND SUPERNETWORKS

With the increase in the use of information technology, and hence in the number of attached computers in a typical corporate network, the strict class A, B and C divisions of the IP addressing scheme, as described so far, are too limiting. The solution introduced in the mid-1980s was subnetwork addressing or 'subnets'. In this chapter, we discuss what subnetwork addressing is and how it can be used to give the TCP/IP network planner much more control over the allocation and sizing of network numbers and subnetwork numbers.

The purpose of subnetting and the subnetwork mask is to allow network managers to allocate different subnetwork numbers to each port of a router using only one registered network number. If selecting an IP addressing scheme is the first major decision the TCP/IP implementor must make, the second parallel decision is how to subdivide an IP network number so that organizations can make efficient use of the available address space. It is often important to devolve management of IP addresses to local departments to give them more control of their own network implementation.

To optimize the allocation of a limited address space the IAB will issue an organization with the smallest number of class B or C network numbers for the total required hosts. But network managers must be able to control the allocation not only of host numbers but also of network numbers if the standard TCP/IP mechanisms for using routers are to operate. These require that two stations on the same bridged LAN share the same network number; on either side of a router stations must have different

network numbers. How can this be achieved if an organization has only one or a few network numbers available? The solution to this problem is subnetworks or subnetworking through the use of the subnetwork mask.

4.1. Subnetworking and the subnetwork mask

Subnetworking or subnetting increases the network manager's control over the address space and provides a mechanism for using routers when only one or a small number of full network numbers is available. The mechanism for creating subnetworks is the subnetwork mask, another 32-bit number which is configured at the same time as the IP address.

Subnetworking divides the normal range of host ids (16 million for class A, 65 534 for class B and 254 for class C) into a number of subnetworks and a reduced number of hosts on each subnetwork. The product of these two numbers cannot exceed the original number of host ids. So, subnetting of class C addresses leads to a small number of networks with a very few host addresses on each. Some implementations may not allow subnetting of class C addresses, though it is sometimes used in research departments with high-performance RISC-based workstations where a few devices can saturate one LAN cable.

When should we use subnetting?

The simple answer is: 'Always plan for it even if you don't use it initially'. If an IP addressing scheme is planned without subnetwork addressing, then it is likely to be very labour-intensive to introduce it later, as a large percentage of addresses will change. Planning ahead is key. Most network implementors should plan their address space on the basis of subnetworks rather than multiple network numbers.

Small networks that will not use routers need not use subnet-' working. Such networks are never likely to grow beyond a few physical LANs and up to 200 stations. Also, if you do not need to use a registered address and do not mind how many IP network numbers you use, then there is less impetus to use subnetwork addressing. However, careful use of subnetwork addressing, rather than multiple network numbers, can simplify future network growth and change, so it has advantages.

4.2. The subnetwork mask

All TCP/IP equipment which operates the network layer IP protocol, and hence

as an IP address, must be capable of using subnetting and the subnetwork mask. The ubnetwork mask is the 'number' under the network manager's control which determines ow the subnetworks are structured. When two network numbers are compared to etermine if a router must be used, they are compared after being 'filtered' by the ubnetwork mask.

Inappropriate choice of the subnetwork mask can lead to great confusion when ne tries to explain how routing will operate. The following description is a simple tarting point.

The subnetwork mask is a 32-bit number that allows a network manager to choose he number of subnets and hosts on each subnetwork. There are still certain strict limits mposed by the basic structure of class A, B and C addresses. There cannot be a greater umber of connections than the basic address gives; in practice the total number of vailable connections is less due to inefficiencies in the way subnets can be allocated. What you gain is flexibility and control. By defining your own subnetworks you have some freedom to place routers for your own convenience without having to use different network numbers.

The rule, shown in Table 4.1, is simple:

• If there is a 1 in a bit position in the subnetwork mask, that bit forms part of the subnetwork number in this address space.

• If there is a 0 in a bit position in the subnetwork mask, that bit is part of the host (or connection) id.

Table 4.1 The subnetwork mask

	Binary	**Dotted decimal**
IP address	10000001.10000010.01001111.01010101	129.130.79.85
Subnetmask	11111111.11111111.11111000.00000000	256.265.248.0
Network id	10000001.10000010.01001000.00000000	129.130.72.0
Host id	00000000.00000000.00000111.01010101	0.0.7.85

Table 4.2 The default subnet masks

Address class	Default mask dotted decimal	Default mask dotted hex
Class A	255.0.0.0	FF.0.0.0
Class B	255.255.0.0	FF.FF.0.0
Class C	255.255.255.0	FF.FF.FF.0

Even if you do not choose to manage subnetwork addressing on your network, all equipment will still use the technique because it will generate default subnetwork masks that select the normal class A (8-bit), class B (16-bit) or class C (24-bit) network numbers. Once a piece of equipment is configured with its IP address it examines the 'class bits' to set the correct default mask automatically.

Subnetwork masks are represented in the same format as normal IP addresses: dotted decimal or hexadecimal. The three default subnet masks are shown in Table 4.2. The network manager takes control of the address space by changing Os to Is in the mask, which successively halves the number of hosts and doubles the number of subnets available. It is strongly recommended that these Is are added from the left of the address and that no intermediate 0 gaps remain. This is sometimes referred to as a mask of contiguous Is. Some software enforces this recommendation.

Subnetworks are used by routers. You may be able to work out how your routers will respond to a mask which is not contiguous, for example:

11111111.11111111.00011111.00000000

But have some thought for those who follow behind you, however. If you make a great success of this network, you will be promoted, but someone less able may be coming along behind! For the long-term viability of the network and the sanity of those who follow, do keep life as simple as possible!

The complete range of class B subnet options is shown in Appendix D. This shows the number of subnet bits, the resulting mask and the first and last usable host addresses on each subnetwork, together with the subnetwork broadcast address and the address that describes 'this subnetwork'.

Because the subnet bits are added from the left of a byte (the most significant bit) but are then represented in dotted decimal notation, the third octet has an unfamiliar

pattern to it. These mystical numbers are formed as shown in Table 4.3. Only these eight different numbers should be used in a subnet mask. If a number in the mask is not 255. the numbers to the left of it must be 255.

4.2.1. Reserved subnet numbers

Subnetting is further complicated because a subnet of all Is and a subnet of all Os cannot be used. All Is is a broadcast to all subnets in this overall network number. A host id of 'all Is' remains the broadcast to all hosts, either on a specific subnetwork or on all subnetworks in this network number.

Table 4.3 How a mask octet is calculated

Value	How 3rd octet is calculated	Binary 'mask'
0	0	00000000
128	128	10000000
192	128+64	11000000
224	128+64+32	11100000
240	128+64+32+16	11110000
248	128+64+32+16+8	11111000
252	128+64+32+16+8+4	11111100
254	128+64+32+16+8+4+2	11111110
255	128+64+32+16+8+4+2+1	11111111

This makes the range of useful addresses more restricted than might have been thought. Because 'all Is' is reserved, the first useful class B subnets are not 128.1.0.0 and 128.1.128.0 with a subnet mask of 255.255.128.0 or FF.FF.80.00 as might be expected. These addresses cannot be used, as 128-l.O.x refers to 'host x on this subnetwork' and 128.1.255.255 is 'broadcast to all subnetworks in the network 128.1.0.0'. 128.1.255.255 could also be 'broadcast to all hosts on the subnet 128.1.128.0'. The two conditions cannot be distinguished and therefore this subnet range is not used.

The first usable mask is the next mask of 255.255.192.0 or FF.FF.CO.O, giving four subnet values, but with only two usable subnetwork numbers of 128.1.64.0 and

128.1.128.0. While 128.1.192.0 at first seems usable, the address 128.1.255.255 again means 'broadcast to all hosts on all subnets in network number 128.1.0.0'. There is no separate way of defining a broadcast to all hosts on the single subnet 128.1.192.0. Any address 128.LO.xxx refers to host xxx on 'this subnet' not on the specific subnet 128.1.0.0, so that subnet is also unusable. This is another reason why the IP address space is not, in practice, as large as might be thought.

4.2.2. Keeping it simple - one value of subnet mask

When represented in dotted decimal, it can be difficult to recognize the broadcast to a single subnet. Address management and routing is much simpler if the same subnet mask is configured on all IP equipment that has the same network number. If more than one network number is available, different masks can be used with each number.

Before using different subnet masks at connections with the same network number, there are a number of implications for routing that must be fully understood for the particular manufacturer's equipment you have chosen to use. The use of different masks is not fully defined by today's standards.

4.2.3. Choosing a subnet mask

Moving the subnet mask boundary exchanges host numbers for subnet numbers. Hosts on different subnetworks cannot communicate directly; they must have an intervening router. So changing the subnet mask may equally be thought of as exchanging bridges for routers. Before choosing the subnet boundary, you must know:

- What is the largest subnetwork that you can successfully manage when it is interconnected with bridges alone?
- What are the cost, performance and management implications of using routers rather than bridges?

The first question implies that, as a network grows, bridges will be replaced by routers at strategic points and hence either addresses will change or the subnet mask will change on all equipment on either side of the new router.

The answers to these questions determine the upper and lower bounds to the maximum size of a single subnetwork. These decisions are important for, in an internetwork with a single subnet mask, all subnetworks will have the same maximum size. For initial simplicity, we just recommended that a single network number should use one subnet mask throughout the internetwork. In normal corporate networks, it is

probable that there will be a few large subnetworks and a larger number of small networks. To optimize the use of the address space around the larger number of smaller networks, the larger networks may have to be subdivided into small subnetworks and interconnected by routers rather than bridges. As the price and performance of routers continues to improve, the performance and cost disadvantages of routers decrease, though the relative performance of bridges may be expected to improve in the same way. It may be appropriate to use more than one network number with different subnet masks to cater for different size and traffic requirements.

So, the power of subnetting is that it allows the network design authority a degree of flexibility in organizing the address space and in developing a router network to suit the organization. The issue is that this flexibility is not as great as managers of the larger corporate network would desire; they are restricted unless they can use multiple network numbers or different length subnet masks in different parts of the networks.

4.2.4. Subnetting a class B and class C address

Most larger commercial networks use one or more class B addresses. (The principles of subnetting apply also to class C addresses; only the scale is different.)

The limits of a single class B address space are from 2 subnetworks of 16382 hosts to 16382 subnetworks of 2 hosts. It is likely that most managements will choose one of the following subdivisions of a class B network:

- 254 subnets of 254 hosts (quite small subnets)
- 126 subnets of 510 hosts
- 62 subnets of 1022 hosts
- 30 subnets of 2046 hosts (large subnets)

These figures produce an acceptable balance between the use of bridges and routers. It has been found that it becomes increasingly difficult to manage a large network consisting only of bridges. A network based only on routers is manageable but can be expensive. Subnet addressing with a mask of 255.255.255.0, giving 254 subnets of 254 hosts, is easier to explain and manage.

There are only two practical options for class C subnets: 14 nets of 14 hosts with a mask of 255.255.255.240 and 6 nets of 30 hosts with a mask of255.255.255.224.

4.3. Configuring subnet masks

Subnet masks are entered into each device at the same time that its IP address is configured and are stored along with that address. In a large network, making any change to the subnet mask may involve visiting the equipment and manually changing the mask. Obviously, this is to be avoided, particularly in large, geographically spread networks. The boot protocol, BOOTP, allows the subnet mask and the IP address to be obtained from a central server

If you do not provide a subnet mask, the installation software generates the default mask at the time you configure the IP address, by examining the class bits of the address you enter.

TCP/IP standards do not specify how particular parameters should be entered into machines, for they do not specify the user interface. Different TCP/IP implementations accept the subnet mask in different ways. These are:

- The full mask in dotted decimal (or hexadecimal) format,
- The number of bits or 'extension' bits in the mask.

A command of the form:

netmask = 255.255.224.0

allows to enter the full subnet mask in dotted decimal format; such an entry may allow any dotted decimal number to be input, but remember it is incorrect to do other than extend the default subnet mask with contiguous Is. For this reason an alternative (and preferred) representation is the 'number of Is' in the mask, using a command like

subnetbits = 19

This example also produces a mask of 255.255.224.0 since the mask consists of 8+8+3 ones. Yet another implementation expects the user to enter the extra number of bits added to the default mask. In such a case, for a class B address:

bits = 3

specifies the mask 255.255.224.0, since the default mask of 255.255.0.0 already has 16 ones.

4.4. Difficulties with a single mask

A single subnet mask and network number cannot satisfactorily span all the requirements of the larger corporation unless a class A address is used and the risk of an address clash is discounted. With a class B address, it is possible to have spare address

capacity in some smaller subnetworks and to have exhausted address capacity in the large networks, but be unable to transfer the unused addresses. Without the flexibility of variable subnet masks, the network manager must either simply live with inefficient allocation of the address space or use more than one network number, possibly with a different subnet mask for each number. Using more than one network number has an impact on the operation of routers.

Subnetworking with a single mask introduces a second level of hierarchy into IP address management. If available the use of variable subnet masks would allow a full hierarchical structure, doubling the number of subnets and halving the nodes for each bit added to the mask. The first level of division in IP addressing is the network number, the second level is the site or location and the third level is the area or subnetwork within that site. Each area has its contingent of attached hosts. A class A address gives the network planner 24 bits to divide into sites, areas and hosts with full flexibility as to where the boundaries should be, at least when the decision is made. Once deployed in equipment, the flexibility is reduced if not eliminated.

The technique applies equally to registered class A addresses but they are unobtainable.

This technique can be used to overcome the problem of changing a bridge for a router, provided the addresses have been chosen correctly in the first instance. Using a router in the presence of subnets begins to bring local geography into play. If all addresses in one building are chosen from one of the subnet address ranges (see next section) and all those in another building are chosen from a second and correctly adjacent range then, merely by changing the subnet mask for the network, each area can be part of one or of two subnetworks, according to choice.

Consider the example of Figure 4.1. Two sites are initially connected by remote bridges using a 2 Mbps digital circuit. Two buildings on each site are connected by fibre optic links from the same bridge. All have the same network number, 191.250.0.0. But the distribution of addresses in the two sites and two buildings is such that by changing the subnet mask, the two sites and buildings within sites can become different subnetworks. At stage 1 the mask is the default of 255.255.0.0 and there is only one network, as required for a fully bridged LAN. In stage 2, by changing the mask to 255.255.240.0 no IP addresses need be altered when routers are introduced between the sites. A second change of the subnetwork mask to 255.255.252.0 places the buildings in different subnetworks (stage 3).

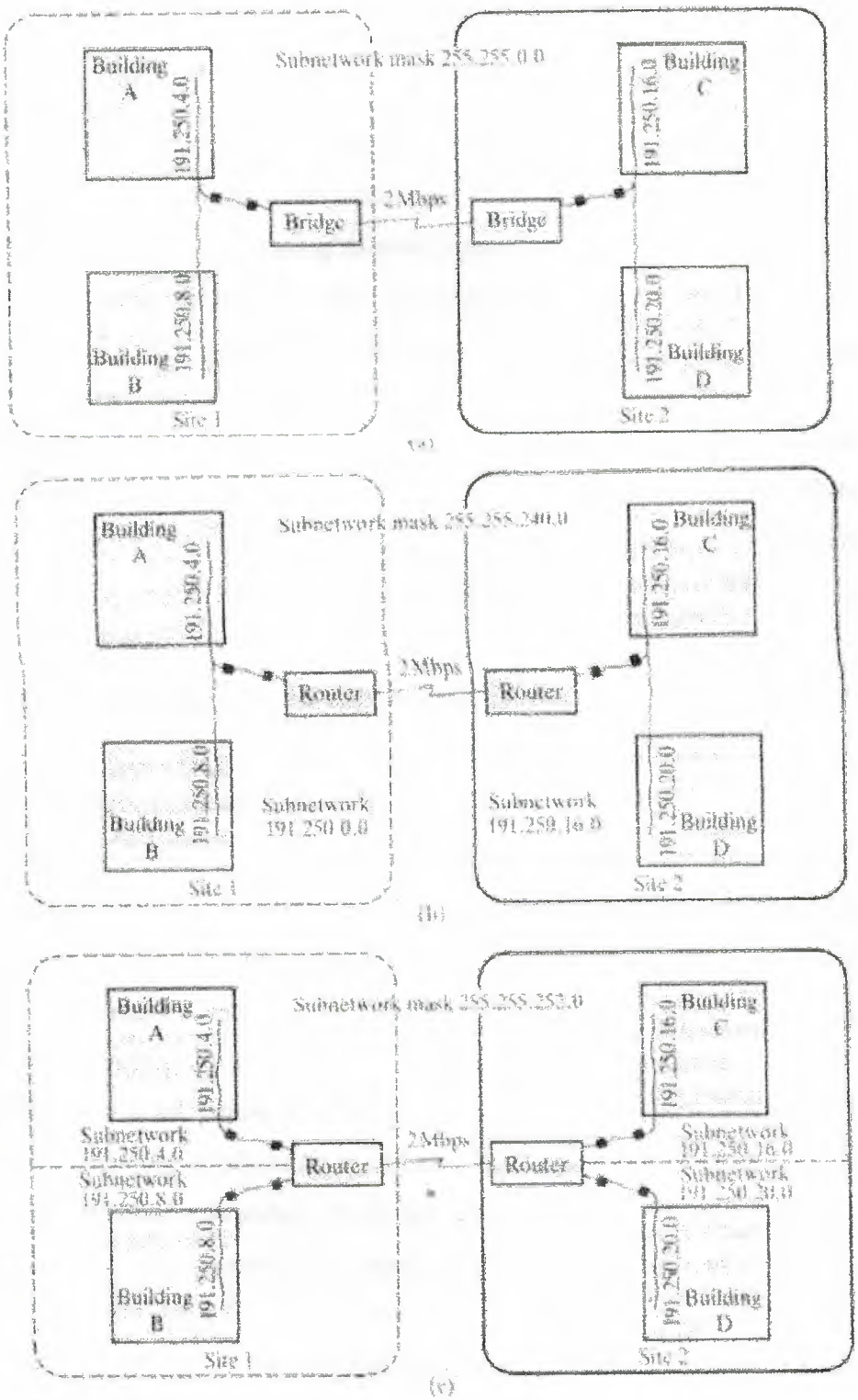

Figure 4.1 Planning routers using Subnet addressing (a) All one network (b) Two Subnetworks (c) Four Subnetworks

Grouping addresses together in this manner also assists with fault finding. When a fault is discovered, the IP address of the faulty equipment immediately identifies the

location of that equipment. If a number of devices in the same area are found to be faulty at the same time, it gives a good indication of where to begin looking for the fault.

Should traffic become unmanageable in a bridged LAN, the changes required to incorporate routing are simplified by subnetting. This is only possible because the upgrade is predicted and the addressing scheme is preplanned. The alternative is labour-intensive, with major changes to the IP addressing scheme required possibly simultaneously at different parts of the network, and having to take place outside normal working hours to avoid disruption.

The practice is rather different, however. By default each IP address and its corresponding subnet mask is configured at each attached workstation. Unless addresses and subnet masks can be managed and changed centrally, it is just as labour-intensive to change one bit of a subnet mask on each machine as it is to change a complete IP address. There is increasing justification for managing and distributing IP addresses and subnet masks centrally through the BOOTP or DHCP.

4.5. A class B hierarchical network

The network number in an IP address does not contain any concept of a geographic location, only of a controlling authority. The IP addressing scheme is not hierarchical. But since the subnet portion of the address is totally under the control of the autonomous system manager, he or she can introduce a geographic component into the address if there is value in so doing.

The following is an example of what can be achieved within a single class B address. IP addresses can be allocated to different geographic sites, in blocks which are on a high-numbered binary boundary. Buildings within a site can be allocated on binary boundaries lower down the hierarchy, according to the number of host addresses they require.

For example, suppose Integralis Ltd has been allocated the class B addresses 191.250.0.0. They wish to cover up to 16 sites with a maximum of 4000 devices on a site. There are up to four large buildings on each site. By using a subnet mask of six extra bits, that is, a mask of 255.255.252.0, there are 64 networks (of which 62 are usable) each with 1022 hosts. The result is shown in Table 4.5. The progression of addresses seems perhaps strange, because the boundaries are determined by binary bit

patterns which are then expressed in decimal. It is important that those who configure TCP/IP equipment are fully familiar with the addressing scheme and the subnet addressing scheme. They should be able to spot instantly any peculiarities in a chosen address.

Such a scheme is wasteful of address space where not all sites are of equal size. But it introduces a true, hierarchical address space where certain bits designate a site, certain bits designate a building and other bits can specify a floor. Some further flexibility can be achieved with the correct combination of bridges and routers.

Note also the scheme is based on geography which may not exactly correspond with an organizational structure. Individual groups in the organization must accept that they are allocated a specific address by a central authority. With care it is possible to allocate a block of addresses, but meeting the needs of groups of widely different sizes may mean that strict geographic boundaries are infringed. Each time such guidelines are infringed it will complicate future changes.

Using a subnet mask of 255.255.255.0 for a class B network simplifies management at the expense of more routers. The third octet will indicate the site or building number directly with no confusion to boundaries.

4.6. Using different subnet masks

Using a single subnet mask can lead to exhaustion of part of the address space, while other parts may have spare capacity which cannot be transferred. In RFC 1219, P. Tsuchiya of Bellcore has suggested a technique which reduces these problems.

For this method to be successful, it requires an initial judgement of the maximum number of subnets and the maximum size of subnetworks. It also requires the use of variable length subnet masks and of routers that understand, exchange and interpret variable length subnet masks. To date, the use of variable length subnet masks is still under discussion by the Internet Engineering Task Force (IETF); not all routers can be guaranteed to handle these in the same way.

The principles of the scheme are fairly simple, but the details could be difficult to implement in a real network with devolved management, unless there is close cooperation and a very good technical understanding of subnet masking.

Table 4.5 A hierarchical class B network with a mask of 255.255.252.0.

Site	Building	Network number	First host	Last host	Comment
		191.250.0.0.	191.250.1.1.	191.250.3.254	Range not usable seen as "this subnetwork"
1	1		191.250.4.1	191.250.7.254	
	2		191.250.8.1	191.250.11.254	
	3		191.250.12.1	191.250.15.254	
2	1	191.250.16.0	191.250.16.1	191.250.19.254	
	2		191.250.20.1	191.250.23.254	
	3		191.250.24.1	191.250.27.254	
	4		191.250.28.1	191.250.31.254	
	"	"	"	"	
	"	"	"	"	
15	1	191.250.224.0	191.250.224.1	191.250.227.254	
	2		191.250.228.1	191.250.231.254	
	3		191.250.232.1	191.250.235.254	
	4		191.250.236.1	191.250.239.254	
16	1		191.250.240.1	191.250.243.254	
	2		191.250.244.1	191.250.247.254	
	3		191.250.248.1	191.250.251.254	
	4		191.250.252.1	191.250.254.254	
					Range not usable seen as "broadcast to all stations this subnet"

The principle is that subnet numbers are allocated from the top of the mask boundary working downwards, while host ids are allocated from the bottom up. Some space is allocated between subnet bits and host bits for growth of either the number of subnetworks or the number of hosts. As the network develops this growth space can be

allocated to hosts or to additional subnetworks as required. The procedure can become complex at the point an additional growth bit is allocated to either host or subnet; it requires a full understanding of the implications and careful coordination across different subnet addressing authorities in an organization.

Variable length subnet masks only work with routers that can exchange masks as part of their routing tables. It may be impossible to broadcast throughout a number of subnets reliably if a variable length subnet mask is used.

4.7. Supernetworks, bridging and switching

The standardized way to communicate between hosts which have IP addresses in different network numbers is via a router. Where an organization has been allocated a number of registered class C addresses, but has more than 254 hosts, there is apparently no option but to use routers to communicate between these different class C networks. Newer technologies such as high-speed switching hubs and 'cut-through' hubs operate at the medium access control layer (like bridges) and therefore cannot be used to the greatest effect if routers are present too.

Supernetworking is a technique which can overcome this limitation, if it is available in host and router software. Class C networks can then be aggregated together to form a smaller number of larger (bridged) networks, reducing the number of routers and making the optimum use of bridges and switches.

Supernetworking is achieved by using bits which belong to the network address as host bits. This is realized by altering the default class C subnet mask such that some of the bits which relate to the network address are set to 0 instead of Is. This technique is not yet standardized; not all host software will allow a mask with fewer '1' bits than the default for a given class to be configured.

When the Network Information Center issues multiple class C addresses in one allocation to an organization with a requirement for more than 254 hosts, it will issue the addresses in multiples of 2, 4, 8, 16 or even 32 network numbers; those network numbers will be chosen so that they are on a 'binary boundary' - that is, some bits in the network address remain constant for that organization and that organization alone. For example, one organization requiring 1000 addresses might have four class C networks, 222.231.32.0, 222.231.33.0, 222.231.34.0, 222.231.35.0; another organization requiring 1000 addresses could have 222.231.36.0, 222.231.37.0, 222.231.38.0, 222.231.39.0. In

this case, 32 and 36 in the third octet are both multiples of 4, the number of networks allocated.

CHAPTER FIVE: ROUTING

Routing, the feature within the TCP/IP protocol suite which allows the 'best path between two communicating systems to be chosen for a given application, depends on the correct choice of equipment and routing software. In this chapter we examine the choices available and the decisions to be taken.

The topics considered are:

- The need for routers
- What routing is
- Routers and IP
- Routing advantages
- Routers and the IP address
- Routing tables
- Choosing a routing protocol
- RIP, OSPF, integrated IS-IS, BGP, IGRP, Hello, EGP and GGP
- Configuring routers

5.1. The need for routers and their management

When designing and managing the IP internetwork, the router network is important to the overall stability and performance of any system based on TCP/IP. Particularly for larger networks of more than 250 devices, routers provide the network manager with more control over traffic flows and the containment of faults than does a network composed solely of bridges.

Routers (or gateways as they were originally called in TCP/IP literature) existed in TCP/IP networks before the simpler and lower cost bridge was introduced to LAN technology in the mid-1980s. In early commercial use of TCP/IP, UNIX systems were used as routers. UNIX systems are still delivered with routing software and any multihomed host can become a router. Indeed, if the default configuration of some UNIX implementations is not altered, any multihomed host will become a router.

Successful routing is complex and processor-intensive. Current recommendations are that routers should be computers reserved for that purpose and not shared as

application servers. Not only does this leave computing power for delivering applications, but it clearly also separates the two different functions of communications management and applications management, a division which, if not enforced, can lead to conflicts of interest.

Routers are programmed to interpret Internet protocols. They take an active role in TCP/IP networks and must be managed as part of the IP addressing scheme.

What is routing?

Routing is the process by which two communicating stations 'find' and use the optimum (best) path across an internetwork of any complexity. The process has several components:

- Determining what paths are available.
- Selecting the 'best' path for a particular purpose.
- Using those paths to reach other systems.
- Adjusting the datagram formats to fit the underlying technology.

In routing, these decisions are made using the network number of the network layer address.

5.2. Routers and IP

The devices which perform routing based on IP addresses are (IP) routers. Historically, routers in TCP/IP have been called (IP) gateways, but since, in OSI, the terms 'router' and 'gateway' describe two different functions, we shall use the more modern term 'router' throughout, except where quoting an RFC standard which uses the word 'gateway'.

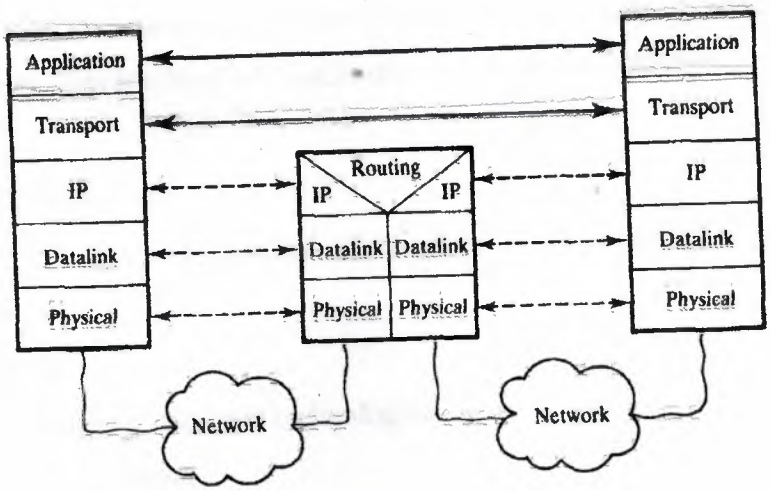

Figure 5.1 Routing with IP.

TCP/IP literature has begun to use 'router' rather than 'gateway', but routing protocols are still referred to as gateway protocols in existing terms such as GGP, IGP, IGRP, EGP and even in the recent BGP.

True network layer routing is performed in the IP layer of TCP/IP and uses the network number or subnetwork portion of the IP address to make routing decisions. The router relays IP datagrams between the IP layers of end systems (Figure 5.1).

Features of IP also allow routing decisions to be made on type of service, precedence, security and predefined routing options specified by the communicating partners. These predefined fixed routing options are called loose and strict source routing.

Routers must support fragmentation. This is the ability to subdivide received information into smaller units where this is required to match the underlying network technology. Fragmentation would occur when a router relayed an 8000-octet datagram from a IGMbps Token Ring to a 4 Mbps Token Ring which can only support 4472-octet frames. If each of the fragments was subsequently relayed to an ISO 8802.3 network with SNAP encapsulation it might be further fragmented into two 1492-octet frames and one shorter frame. Each fragment must have its own IP header, and after fragmentation the parts are treated as independent datagrams which can follow different paths. While fragmentation is normally invisible to the application programmer and application user, it can produce some bizarre effects for the network manager trying to optimize the use of limited resources. In general, repeated fragmentation in a network is undesirable because of the overhead of fragmentation in routers, the extra headers and the processing power for reassembly. Where options are available they should be set to avoid fragmentation. This is a level of functionality not available in bridges.

5.3. Routing advantages

Why use IP routing rather than the much simpler and less costly bridges? The key points are:

- Better choice of routes,
- Matching different link-level technologies,
- Resilience and control,
- Error reporting.

While these are features of the IP standards, not all end systems and commercial routers support them.

Modern routers are designed so that the routing complexity is contained within a relatively small number of powerful multiprocessor machines specifically designed for the task. Hosts use simpler methods to discover and use their nearest and best router for the communications task in hand. This leaves all the processing power in the host and workstation end systems for the task that they should really be doing, processing the users' data, not expending effort trying to be a good communications processor. Applications processes should be performed in machines chosen for that specific task; likewise network relaying and communications should be devolved to specialist processors designed for that intensive process.

Better choice of route

Routers, using the best modem routing protocols, can have multiple parallel operational links, can find any available alternative route and will share traffic across those paths according to criteria established by the network manager and by software in the communicating hosts.

Matching different link-level technologies

Long-distance (wide area) point-to-point links and the variety of ISO/IEEE LAN standards have very different performance and data transmission characteristics. It is not possible to disguise the detail of the underlying technology from a bridge. A router is specifically designed to take account of the differences as it is made aware of maximum transmission units (related to frame data size) and can convert with fragmentation.

Resilience and control

Routers form an integral part of network layer protocols and network layer addressing. End systems (hosts) can choose one specific router to relay information. Routers will not relay MAC layer broadcast frames. They act as a barrier between network areas and prevent the propagation of certain types of faults (broadcast storms) from one area to another. Routers only act upon IP datagrams that are sent specifically to them.

Error reporting

Routers and hosts must use a protocol called the Internet Control Message Protocol (ICMP) to report and record error conditions and to try to control network congestion.

Routers and the link layer

While routers are network layer devices, they must encapsulate their network layer information in link layer frames. All manufacturers must support DIX Ethernet and SNAP encapsulation on ISO networks.

5.4. Routers and the IP address

Routers and end systems use the network number or subnetwork number to decide how to relay a particular IP packet (IP datagram). There is a very close correspondence between the network layout and structure (due to the geography of the organization) and the (subnetwork) address plan. Any change in the location of a host or workstation computer may mean that its IP address must change. (Its domain name may remain the same.)

A connection on a router is an IP connection on the network number or subnetwork number to which it attaches. Every router port can be the source or destination of IP datagrams. This is how a router knows which networks it is directly attached to. The router must be configured with IP addresses and subnet masks for each of its network connections. This is normally done through the hardware terminal port on the router chassis.

5.4.1. Routers with point-to-point wide area circuits

A router can be the source and destination of IP datagrams, so each of its connections must have an IP address with the appropriate network number. Where two routers connect two LANs over a point-to-point PTT-provided circuit, some router implementations require that this single intermediate circuit has its own unique network or subnetwork number. This is unsatisfactory. In a large network it will consume considerable numbers of network or subnetwork ids; more recent router implementations do not require point-to-point circuits to be explicitly labelled.

It may be appropriate to use a different network number and subnet mask for these point-to-point links. The subnet mask can be 255.255.255.252, which leaves two usable addresses for each end of the link; the addresses would progress ending in 1 and 2, 5 and 6, 9 and 10, 13 and 14,17 and 18.... Addresses ending in 0 and 3, 4 and 7, 8 and ll... cannot be used as they are the 'this network' and 'broadcast' identifiers for the

subnet.

When two routers from different manufacturers are connected together over point-to-point leased telephone company (PTT) circuits, there is no guarantee that they will work together unless both manufacturers have implemented the identical link-level protocols on the circuit. Even where two manufacturers claim to use the same protocol, the two routers may not interoperate when joined by a point-to-point link. If interoperability is important at this level, check with the manufacturers that it has been tested.

The accepted wide area protocol for routers is PPP. PPP is a low-level protocol that can be used by routers on point-to-point circuits; it is not a 'router protocol' or 'routing protocol.

5.5. Routing tables

Every router contains a routing table of the network numbers (or subnetwork numbers) that it knows about. The table records which router connection(s) can be used to reach a particular network and some indication of the performance or cost of using that connection to reach that network number. TCP/IP routing is based on a knowledge of only the 'next hop' in the path to the destination network number and host. To use an analogy from the road network, routers do not keep a 'road map' of the complete network, rather they have a set of'signposts' with 'distances' or 'speed limits'. At each 'crossroads' (the next router), the next signpost is consulted to find the direction of the shortest or fastest path there. Like the road network, occasionally something turns a road sign around; traffic then takes a less than optimum route, or may even go round in circles.

In the simplest and older routing protocols the measure of performance of a particular link is the number of hops (routers) that have to be traversed to get to the destination network. Such a simple measure can lead to the use of totally anomalous routes.

One routing option is static routing. This may be appropriate where the network is small or because network managers wish to enforce a fixed routing policy for security or operational reasons. Static routing tables are generated manually by the network manager and loaded into the router either through the control port or across network links. Such a procedure is error-prone and cannot give the best response to network and wide area link failures. It becomes increasingly difficult to manage as the numbers of

networks and routers increases, and only a limited number of routes and alternative routes can be devised and managed successfully by hand.

It is unusual in a closed corporate TCP/IP environment to implement static table-based routing. It is more usual to allow the routers to operate a routing protocol or interior gateway protocol (IGP) which will update routing tables dynamically. Routers exchange information about network numbers they are aware of and of the performance of the routes to those networks, on a regular basis, using the routing protocol. These exchanges provide routing table updates as conditions change. So "routing protocols" are used not to perform routing but to exchange information about available routes and, in some cases, about their performance. The skill in designing routing protocols is to ensure that changing conditions and routes are tracked quickly throughout the internet without producing temporary routing anomalies or using up all available (wide area) capacity with routers exchanging routing information rather than useful user data.

5.6. Classless inter-domain routing (CIDR)

By default, a routing table will contain an entry for every network number or subnetwork number that exists. In an isolated corporate network these tables will, at most, be a few hundred entries. In the global Internet, there are potentially hundreds of thousands of operational networks, which would require an impossibly large table.

Such problems are solved in other network technologies by hierarchical routing: part of the address space is used to route to a continent, another part for routing to a country, yet another part for an area or organization within the country and finally to a site or building. This reduces the size of the table but can waste address space. Such a scheme is used in the addressing plan (numbering plan) of the international telephone network - country code, area code, exchange code, subscriber's line.

IP addressing and routing was not initially designed with hierarchical routing in mind. As the global Internet grew, the routing tables in the core of the internetwork also grew (to about 13 000 entries) and were reaching unworkable proportions. The number of entries has been reduced and the rate of growth brought almost to a standstill, by using Classless Inter-Domain Routing (CIDR).

In CIDR, groups of previously unused class C addresses were pre-allocated to three addressing authorities, one in the USA (NIC), one in Europe (RIPE) and one in Australasia-Pacific (APNIC). By using a single subnet mask which forms a large

supernet for these class C addresses, routers can hold a single route entry for the main trunks between these three continental areas, or between Internet service providers or into a single corporate network designated by multiple class C addresses. Since the concept of a class C address becomes meaningless on these routes between 'Domains', the technique is called 'Classless Inter-Domain Routing' or CIDR.

The routing table at a particular router can also be minimized by configuring a 'default route' and 'default router'. The default route is towards a router (the default router) which is assumed to know the routes to every network that is not explicity listed in this router's table. There must of course be a router that knows the routes to everywhere. This technique is particularly important in corporate networks which connect to the Internet - the router which provides access to the Internet is the ultimate default router for that organization.

5.7. Choosing a routing protocol

The first function of a routing protocol is to provide information from one router to another about the network numbers and the subnetwork numbers that are known to it, combined with some measure of performance such as distance, throughput, transit delay, error rate, and cost. These measurements are known as (routing) metrics. Today IP supports Delay, Throughput, Reliability and Cost (DTRC), where reliability indicates the desire for a low probability of datagram loss.

Few routing protocols can support all metrics. Even fewer current computer applications indicate to IP the performance environment they wish to operate in. Much of the power in IPv4 is unused and yet to be fully exploited by applications.

Routers do not usually measure and update the DTRC metrics associated with a particular link or set of links; they are configured as static values by network managers. It has been found that dynamic load-measuring schemes can become unstable with traffic oscillating from one route to another and back again in response to rerouting from highly loaded links. TCP/IP documents discourage standards developers from devising dynamic load-balancing routing schemes.

Where variable length subnet masks are used, the routers must also exchange the network mask which they use to make a routing decision for a particular subnetwork. This itself depends on the choice of the routing protocol designed for the purpose. Variable length subnetwork masks are important for hierarchical and efficient

addressing in TCP/IP networks.

A routing protocol does not directly perform routing. It updates routing tables in routers so that they may perform routing. In the jargon of TCP/IP, a routing protocol 'advertises reachability information'. However, the performance information that is recorded in the routing table is dependent on the routing protocol for update; the basis of the router's choice of a particular route is completely determined by the ability of the routing protocol to convey that performance information.

Those implementing a TCP/IP-based internetwork have a wide choice of proven and newer routing protocols. Normally an Autonomous System (AS) manager will attempt to standardize on one routing protocol for use within the AS. This is referred to as the Interior Gateway Protocol (IGP) for that AS. So an IGP is a concept, not a single technical specification. In practice, many network managers find they must support several routing protocols simultaneously for reasons of flexibility and equipment compatibility.

Routing protocols are chosen on a number of criteria:

- Availability from suppliers;
- Speed to adapt and find alternate routes round network failures;
- Conveying best route information according to service needs or performance;
- Support for active alternate and parallel routes to a destination network;
- Speed to resolve routing anomalies such as routing loops and 'black holes';
- Load imposed on networks and internetwork links by the routing protocol itself;
- Load imposed on host systems by the presence of the routing protocol;
- Load imposed on the routing processors by calculating the 'best routes';
- Scalability - change in routing performance and routing traffic with the size of the network;
- Security;
- Support for 'policy-based routing';
- Legal requirements.

Some routing decisions that would be legal between organizations in one country may be against PTT and government policy in another. Restrictions are diminishing in many countries, but equally the regulations are often very different from the substantially deregulated nature of communications in the USA and the UK. The available protocols are not necessarily sophisticated enough for this type of routing decision (nor would governments allow control of routing to be within the authority of

corporate network managers). Network managers working under such restrictions should be aware that inappropriate interconnection and configuration of their networks to their trading partners' networks could mean that they unwittingly act as a carrier for transit traffic between third parties.

The oldest routing protocols only convey the distance between networks as the number of hops - the number of routers that the information passes through to get to the destination from that router. This is a very limited and unsatisfactory measurement, for it takes no account of the performance and changing load of a route.

The routing protocols available for TCP/IP routers and end systems are:

Routing Information Protocol (RIP)

- Open Shortest Path First (OSPF)
- Border Gateway Protocol (BGP)
- Interior Gateway Routing Protocol (IGRP)
- Hello
- Exterior Gateway Protocol (EGP)
- Gateway to Gateway Protocol (GGP)

The protocol in most common use is RIP, because of its widespread availability. It has some limitations for internets with more than a few networks. OSPF is a recent full IAB standard. It has a number of advantages over RIP for the larger network.

5.8. Configuring routers

Many modem routers are easily configured from a management terminal through simple menu-driven commands. The key parameters to configure for IP routing are, for each connection:

- IP address,
- Subnet mask,
- Routing protocols to be enabled,
- Maximum transmission unit of the interface,
- 'Cost' of using that interface.

CHAPTER SİX: TCP/IP UPPER LAYERS, TRANSPORT AND APPLICATION SERVICES

The TCP/IP standards leave each manufacturer considerable flexibility in how the

standard upper layer communications services are presented to end users. The first products to the market provide basic functionality. As the market matures and feeds back ideas for improvements to developers, subsequent products will build more sophisticated user services onto the same basic standards. The size of the TCP/IP market expanded considerably in the early 1990s. The basic facilities have been overtaken by more highly integrated solutions, many based on graphical and Windows interfaces.

6.1. The transport layers

The transport layers of a protocol are the first layers in most communications systems that operate in the host end systems only and are not used by the intermediate network equipment. The transport layers allow multiple applications or users to share the same path between the same two machines; they provide a consistent interface to the applications and a mechanism for them to specify their communications requirements. As shown in Figure 6.1, the transport layer operates end to end whether the communication path is across a single network (interconnected with bridges) or across multiple networks interconnected with routers.

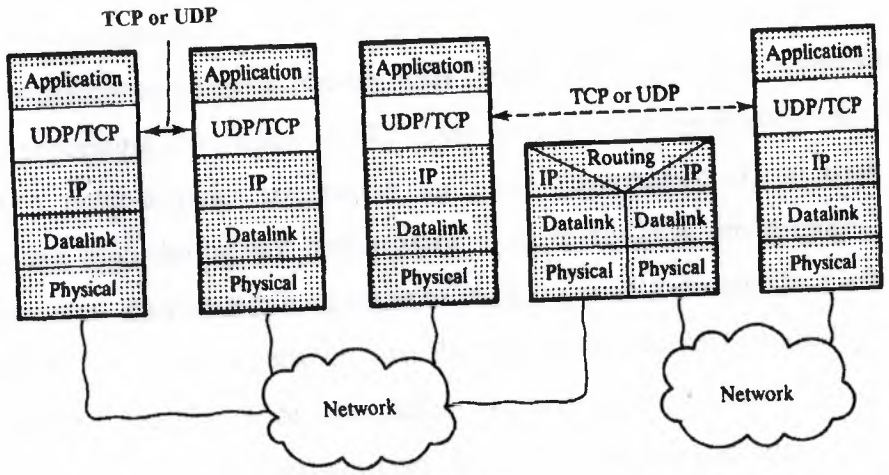

Figure 6.1 TCP/IP transport layers.

Today, most TCP/IP applications use one of two transport protocols, the User Datagram Protocol (UDP) and the Transmission Control Protocol (TCP). For applications which must operate over long distance circuits, TCP is the protocol of

choice, as it can be 'tuned' to provide better performance. UDP has no alterable parameters.

UDP provides an unreliable service; there is no guarantee of message delivery. If the application requires absolute reliability, it must perform its own error checking and error correction. UDP is ideal for distribution of information which changes every few seconds - status reports for stock market screens or airport information systems for example. If one transmission is lost there will be another more up-to-date value transmitted in a few seconds. Unreliable datagram services can be 'broadcast' to many different destinations simultaneously, since the receiving systems can just listen, they do not need to reply.

On the other hand, TCP provides a reliable 'point-to-point' service and has facilities which improve performance over slower, wide area (PTT long-distance) circuits. TCP would be the choice for operating in a wide area environment. It was designed for point-to-point applications that must have reliability - file transfer, mail transfer and terminal access systems.

The transport protocol used by a particular application is usually fixed (at compile time) by the application developer. For the network operator, there is little opportunity to choose between UDP and TCP for a particular product. Further, because UDP involves less processing overhead and is perfectly adequate in a purely local LAN, many application programmers, who may not fully consider the implications for wide area telecommunications systems, choose UDP rather than the extra complexity (and hence larger processing overhead) of TCP. Alternatively, the application may be too restricted by a point-to-point connection for each conversation As a result, these applications have high throughput performance in a strictly local environment, but, as bridges, routers and slow-speed, long-distance circuits are introduced, the performance degrades quickly. If the path between two devices which use UDP includes multiple bridges, routers or multiple wide area links, the performance will be poor. System planners may have to plan the location of network resources to avoid wide area links for certain applications that use UDP. Even though they may be able to justify and afford high-speed, multimegabit connections, the 'slow' speed of light and the delays in buffering of information will reduce the throughput of UDP.

6.2. Winsock

Winsock (Windows Sockets) is an important Application Programming Interface (API) developed for and promoted by most TCP/IP software manufacturers. It allows application programs running under Microsoft Windows to access the TCP and UDP layers of a TCP/IP system. It has been instrumental in opening up the market for TCP/IP applications running on PCs with Windows. Any application written to the Winsock standard may be used with any TCP/IP communications software which provides a Winsock interface. Most Windows-based software uses the Winsock standard. The success of many of the Internet programs is in part due to the availability of Winsock.

In operation, the Windows application searches for a set of communications software routines in a file called WINSOCK.DLL. While the interface that this file provides to the application program is always the same, the way Winsock connects to a particular TCP/IP stack is different. In the unusual circumstances that there is more than one' version of TCP/IP on a PC, it is important that an application finds the correct WINSOCK.DLL for the TCP/IP it is to use. This means that there should only be one WINSOCK.DLL accessible via the path statement or alternatively that the path to the required Winsock must be earlier in the path statement than any other copy of WINSOCK.DLL.

6.3. Open network computing or the network file system

The Network File System (NFS), as it is still commonly known, or Open Network Computing (ONC), as its creator Sun Microsystems Inc. now calls it, was developed for high performance UNIX workstations and later released to the 'public domain' as three RFCs which describe the major protocol components. Recognizing the importance of the IBM PC, Sun also developed NFS software for the IBM PC, the PC-NFS client. This allows PCs to use UNIX workstations as remote file and print servers.

NFS client and server software is now available for most computer systems. NFS provides user services familiar to any user of a proprietary resource-sharing LAN such as NetWare, LAN Manager or Banyan VINES. Files, directories and peripherals on a remote NFS server are mapped to 'virtual' drives, directories or peripherals on the local system. For most purposes, these virtual drives and printers are indistinguishable from, and are used identically to, local resources. This is illustrated in Figure 6.2.

On a UNIX system, each remote directory appears as an addition" al directory off the main local directory structure. For a PC, with PC-NFS, the remote directories appear as additional disk drives.

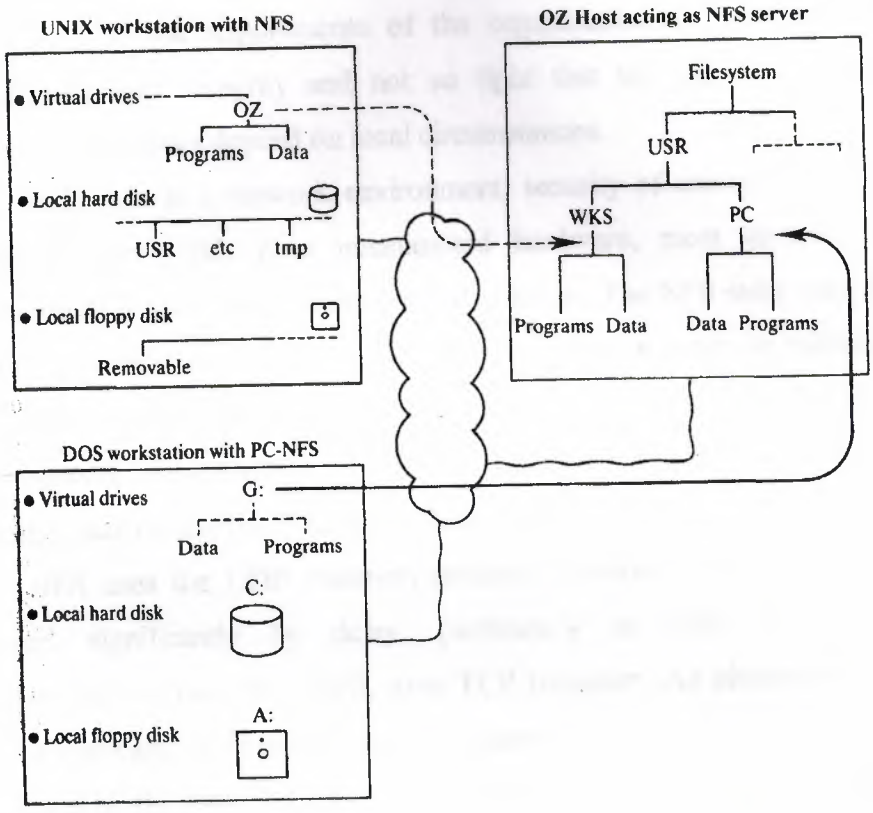

Figure 6.2 NFS representation of resources.

From the perspective of the commercial information systems manager, this is ideal, for like other TCP/IP applications described below, most business users "require no new training to use NFS; they merely have access to more information with bigger disks and better printers.

But NFS can go beyond simple resource sharing. It provides the mechanism for true distributed computing, where processing power, not just data, is shared among networked machines.

6.4. NFS management

The major management issues of NFS can be hidden from the users.

These issues centre on the security and mapping of user identifiers and access rights between different operating systems. NFS works in conjunction with the Network

Information Services (NIS), a distributed database system for NFS security which provides for centralized management of a highly distributed system. Technically, the tools have been provided by Sun; the administration of a large user population still requires careful organization, with the correct clerical procedures in place to ensure that security meets the requirements of the organization. It must not be so lax as to compromise data integrity and not so tight that no one can get work done. The organizational issues depend on local circumstances.

As always in a network environment, security of access from a PC represents a problem. Unless PCs have nonstandard hardware, most security barriers can be overcome by a proficient and determined attacker. The NFS suite uses an authentication program for PCs, the PC-NFS Daemon, running on a server to validate user access to the filing system. This provides a basic level of security. Some operational environments require additional security features. 'Secure NFS' addresses these issues by using Data Encryption Standard (DES) authentication techniques.

NFS uses the UDP transport protocol. Remote disk data transfer rates can be reduced significantly by delay, particularly in wide area links. There are implementations providing NFS, over TCP transport. An alternative solution is to buy more servers and place them close to the users.

Where the main NFS activity is resource-sharing, poor configuration options in the workstations can have a dramatic effect on network traffic. The use of print spoolers, temporary files, backup files, PC batch files and the format of the 'path' statement can all influence the traffic a workstation generates on the network. To make NFS completely invisible to end users, each workstation should have a script or batch file which executes the NFS 'mount' commands to make the remote resources locally available. The details of the command line can be totally hidden.

In the past, memory requirements for PC network software have left too little of the available 640k to run some of the more demanding PC applications. Some implementations of PC-NFS require in excess of 130 kbytes to load a LAN card driver, the TCP/IP core software and the NFS client application. By reducing the number of file systems which may be mounted and by tuning other parameters, this figure can be ;reduced by a few kilobytes. More advanced processors and more sophisticated operating systems than DOS are eliminating these concerns.

6.5. The X Window System

The X Window System was developed by Massachusetts Institute of Technology (MIT). It is a method of controlling an advanced graphical 'windowed' interface. From the perspective of TCP/IP, the X Window System is a message protocol between an X server and an X client (Figure 6.3).

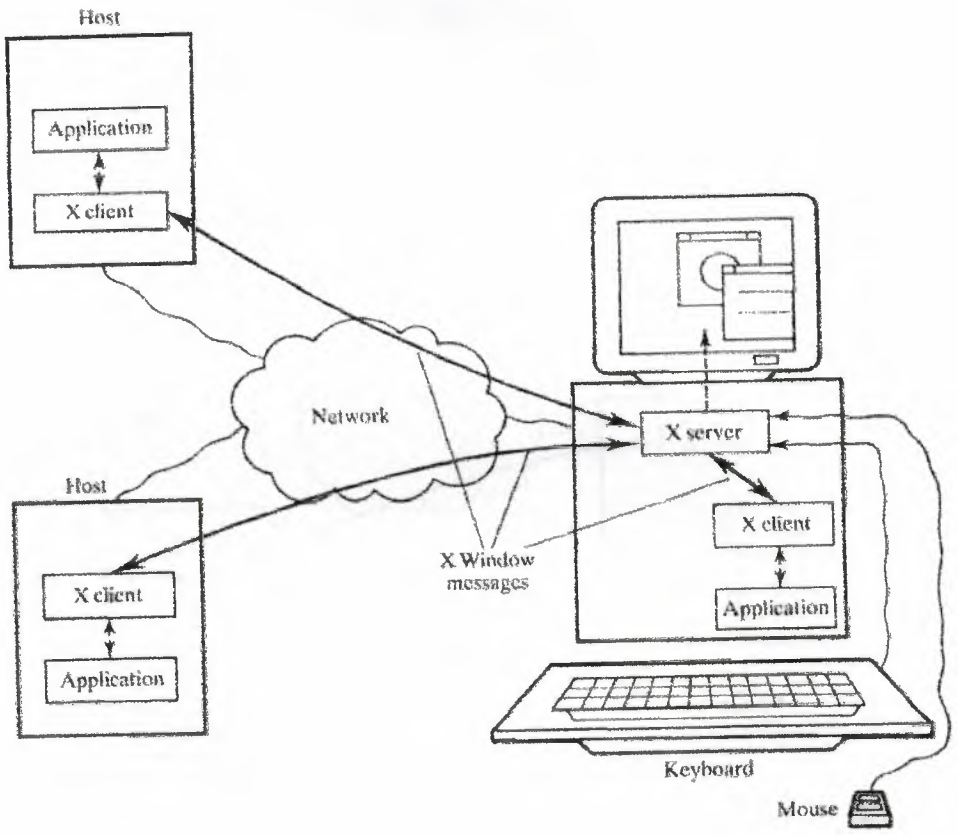

Figure 6.3 X Window protocol.

This protocol is described in RFC 1013. Copyright remains with MIT, though permission is given to distribute the RFC document as long as the copyright is acknowledged. Other aspects of the X Window System, although not published as RFCs, are described in standards available from the X Consortium at MIT.

With the X Window System, the boundary between user interface, which is not normally denned, and the communications protocol may seem to been breached, but in fact it is intact. RFC 1013 only describes the protocol between server and client. The style of the display is determined by other standards, typically the OSF/Motif display

standard promoted by the Open Software Foundation Inc. (which includes DEC, Hewlett-Packard and Microsoft), Open Look from AT&T or Open Windows from Sun Microsystems Inc.

Unlike every other reference to client and server in TCP/IP, with X Window the server is normally at the user's workstation and the client, which generates the new drawing instructions, is at the application host.

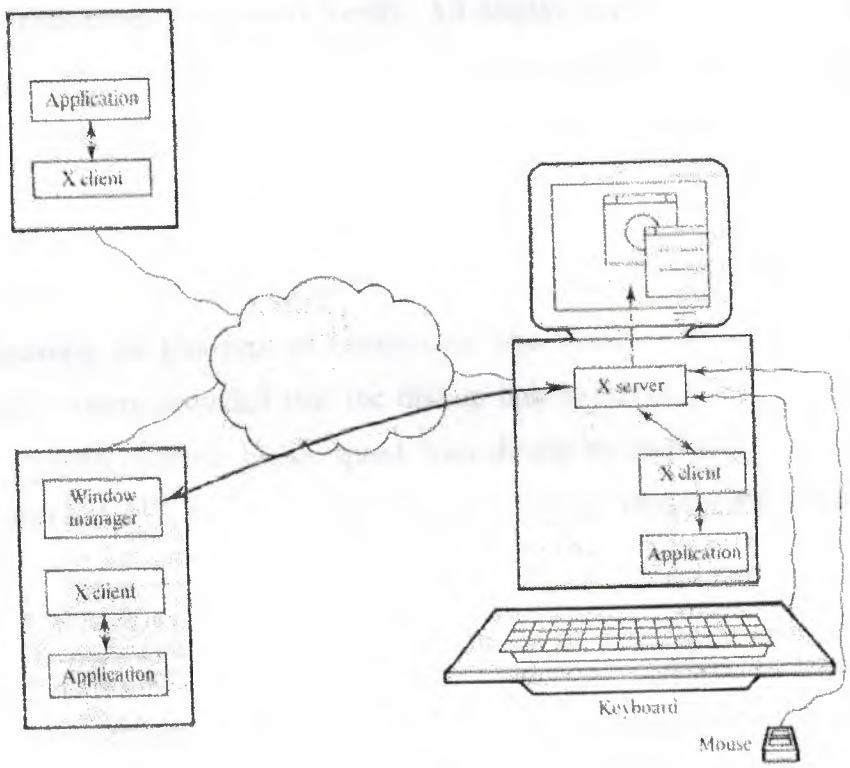

Figure 6.4 The X Window manager

The X server operates the display terminal, drawing graphics objects and text in response to messages from the X client. The server must also report user actions such as keystrokes and mouse, movements to any X clients that will be affected by them.

Since a window system may display output from many different applications and hosts simultaneously, each display should have a window manager, a special X client that supervises the construction of all the graphics objects on the screen (Figure 6.4). It is the window manager that implements the window style, or 'look and feel' as it has been called, of the display standard (OSF/Motif, Open Look or some other standard).

More practically, it is the window manager that adds and controls the scroll bars, title line, move buttons, sizing, scaling and overlaying of windows in response to user actions. Any graphical interface with the modifications to provide the correct software interface to the X server can act as a window manager for an X Window server. Microsoft Windows has been adapted for this role.

6.6. The X terminal

The X terminal is an X Window display station that implements the X server. It runs no user applications (X clients) locally. All display requests are received on the network connection. Extensions to the X user interface can provide for colour, image support and Display PostScript among others.

Some X terminals have been adapted to operate over dial-up modem links (using either the SLIP or PPP protocols described in Chapter 10). Since modem links are limited to 14400 bps or 28800 bps before compression, many suppliers offer some form of data compression for this type of connection. The result is a usable, if somewhat sluggish, display system provided that the dial-up link is carrying data for a single X terminal user. Where possible, higher speed lines should be used for X terminals. The increasing availability worldwide of ISDN 64 kbps 'dial-up' circuits will alleviate these restrictions.

Managing the X Window System

Graphics applications, particularly when bit-mapped graphics is involved, are demanding both of processing power and of communications capacity. The communications requirements will increase if the X server and window manager are not on the same workstation. The earliest X Window terminals operated with a remote, host-based window manager; the standards specifically provided for it. In this case, every user action, from a key press to a pointer (mouse) movement generates network traffic, with a large movement of the pointer potentially generating a stream of X protocol messages.

The X Window protocol uses TCP reliable connections between a server and its clients. Each TCP data segment sent may be individually acknowledged, almost doubling the expected traffic. If the protocol is confined to a LAN segment reserved for the purpose, this traffic is unlikely to be an issue. Where X systems cross bridges or routers between LANs or more particularly cross wide area lines, the traffic generated

by particular actions should be measured for every X implementation being considered. The network capacity and layout should then be planned carefully to carry the expected traffic. Some of the more recent X terminal implementations use a local window manager, which removes a high proportion of the traffic from the network.

Tuning TCP may not improve the performance of an interactive protocol like X Window as much as a bulk transfer protocol like FTP.

6.7. Telnet

Telnet is the virtual terminal protocol of TCP/IP. It operates over the TCP error-corrected transport layer. It provides the terminal service that was mentioned in the very first pages of this book - total terminal interconnectivity and interoperability.

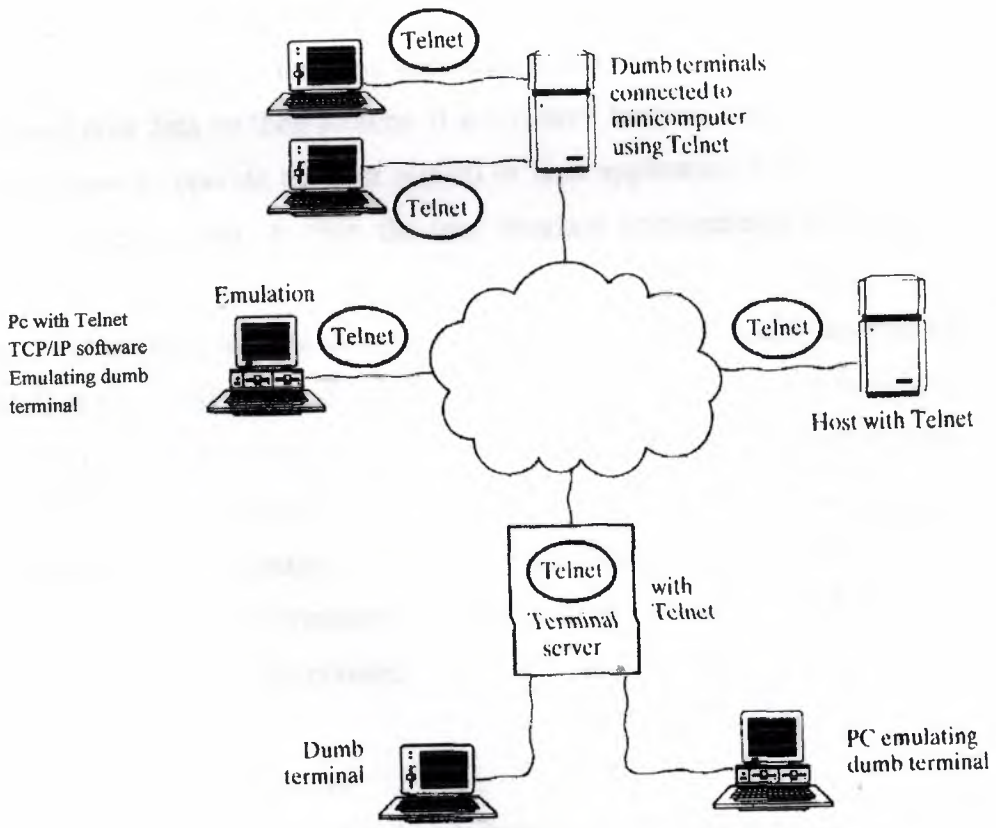

Figure 6.5 Telnet support for terminals.

Telnet gives terminal users the ability to log-on to many different 'Telnet hosts' from a single terminal on their desks. As shown in Figure 6.5, there are four ways of using Telnet. The terminal could be a simple 'dumb' terminal connected to a TCP/IP

terminal server with a standard communications interface (V.24/V.28 or EIA232). Equally, that simple terminal could be a workstation or PC running a terminal emulator and connected to the same terminal server in the same way. A minicomputer can run Telnet software which provides the terminal server capability from within the minicomputer. Alternatively, a PC or other workstation could run a version of TCP/IP and a terminal emulator and connect directly to a LAN.

The Telnet service was one of the first to be provided in a standard way by the TCP/IP architecture. It delivers a similar service to that provided by the switching statistical multiplexers of the early to mid-1980s which provided many companies with their first example of flexible access to information.

6.7.1. The Telnet user interface

Telnet allows terminal or workstation users to gain access to a host system and to display host data on their screens. It is a remote terminal service; any terminal user must know how to operate the host system or host application from a terminal of that type. No attempt is made to map the user interface environments from one system to the other.

The Telnet service is most commonly used with ASCII asynchronous terminals such as the ANSI standard terminal or the DEC VT series. Most manufacturers of other terminal types have registered them with the Internet Assigned Numbers Authority (IANA) so that they can be identified to the Telnet protocol. That does not mean that a particular implementation of workstation Telnet will necessarily match the host software. Any-to-any protocol conversion is not necessarily a feature; systems which match must be carefully chosen.

6.7.2. The terminal server

The advantage of the Telnet terminal server is that it provides a low (capital) cost connection. But the interface often operates in asynchronous character-by-character mode; the terminal is strictly limited in its capabilities with little possibility of upgrading the service in the future.

Asynchronous interfaces can involve large management overheads. Cabling for the terminals may require only three wires, but for some terminal types and uses, it could be more. A structured cabling system with centrally located terminal servers can reduce management costs. At each interface, a number of parameters must be configured -speed, parity, flow control method and type of terminal (to ensure the correct control sequences are used). Unless a single standard for these parameters can be enforced throughout the organization, any apparent cost savings can be quickly swallowed up in increased management costs. This is particularly so in the larger installation, where the unit costs of supporting large numbers of terminal servers may not reduce with increasing size. Where terminal servers can be remotely configured from a central network management system, support costs may be more easily contained.

6.7.3. Configuring Telnet

One of the most difficult aspects of Telnet is ensuring that character and keyboard mappings are standardized across all systems. In a large organization, host configuration will be the responsibility of many different groups; the responsibility for terminal servers may also be devolved. But the workstation or terminal servers and hosts must be matched, particularly in their use of the backspace, delete and 'arrow' keys. A consistent policy is required across the organization. One of the responsibilities of technical management in an organization where many different systems run Telnet will be to standardize key mappings.

If every nationality used the same (US) keyboard with PCs, then there might be some possibility of consistency, but, when national keyboards and currency symbols are involved, that consistency may elude even the most vigorous attempts to introduce it. PC implementations of TCP/IP are famous for not being configured correctly. Trying to operate a remote system when keys marked '@' or '#' or' I' or '£' produce a different character on the screen can be quite demoralizing even for an experienced technical user. This situation is improving as US software developers recognize the international features of the IBM PC and include support for different national keyboards.

Many implementations of Telnet for workstations and PCs support IBM 3270 terminal operation intended for accessing IBM hosts (with an appropriate implementation

of TCP/IP Telnet on the mainframe). Such a connection eliminates the need for IBM 317x cluster controllers. The support is likely to be limited to text only, but could include an emulation of 3278/9 colour terminals on PC colour screens with extended sizes of up to 32 x 80, 43 x 80 or 27 x 132 characters (model 3, 4 and 5). Where Telnet is used like this across different platforms, it may be impossible to map every key on one keyboard to a logical keystroke or combination on the other. At least one manufacturer includes the ASCII to EBCDIC translation table (PC to IBM character set translation) so that any modifications can be made by user management. As always, once a decision is made to change these defaults, continued local support will be essential throughout the life of the system. For keyboard mapping, the manufacturer's manuals should be kept to hand. 'Quick reference' keyboard templates permanently attached to each terminal may save a great deal of user frustration.

Most Telnet implementations allow the workstation user to open multiple connections simultaneously to different destination hosts. Terminal users may also 'escape' to the operating system (for example, a DOS shell) while still connected in Telnet. Depending on the operating system and program memory requirements, this allows other programs to be run. Some Telnet versions allow file transfer and Telnet to operate together. While this can provide a useful environment for computer development and maintenance staff, it is less useful as a general business tool, unless the mechanisms can be hidden from the commercial user by an in-house application.

Another feature of some PC Telnet implementations is the provision of a simple data transfer mechanism to which users can integrate their chosen terminal emulator package. (In a PC implementation, the software interface is often through software interrupt 14H.) The Telnet connection then replaces the PC's hardware communications port (COM1 for example). This means that any communications features of the emulator package can be used with Telnet. These could include support for nonstandard terminal types, graphics or colour support, and the use of a 'script' communications language to automate frequent communications tasks.

6.7.4. The place of Telnet in the 1990s

Even into the 1990s, a character mode terminal service can play an important role in

commercial information systems, though it is beginning to look dated.

With the correct (text) menu-driven applications on the hosts, the data users can be protected from the command line of a particular operating system and can quickly navigate around complex commercial applications. Accessing text systems from one window in an X Window, Presentation Manager or Microsoft Windows environment with the ability to move data between applications easily, further enhances this traditional information systems environment. Delivering a text answer to an enquiry across a LAN in a fraction of a second, rather than in seconds across slow-speed links, brings a new perspective to such systems. Once the data is retrieved from the central resource, it can be massaged and presented in a graphical form at the workstation, if necessary.

But many factors dictate that for an increasing number of users, more transparent and sophisticated data retrieval and presentation are likely to be required.

6.8. File transfer protocol

The function of the File Transfer Protocol (FTP) is self-explanatory; it provides a means to move files from one computer system to another. So FTP usually only has a part to play on a workstation or host which has a local filing system, normally its own hard disk. The FTP RFCs define the protocol which computers use to pass file transfer information between them, not how the user application or interface generates those commands.

As well as transferring files, FTP also provides facilities for managing files on remote systems; you can show or change the current disk directory, list the contents of that directory, delete files from the directory and rename files. While it is not part of the FTP standards, most systems also allow similar file management commands on the local system while connected to a remote computer. When FTP is running on a multitasking operating system, connections can be opened to different remote computers simultaneously; PC implementations often only allow a single remote connection.

FTP regards each file as a stream of bytes to be transferred. There is no concept of structured data or of opening a file to access individual records.

6.8.1. The FTP user interface

The user interface to FTP is not defined by the RFC standards. The normal UNIX

nplementation is character based, and is similar to entering commands at a DOS or UNIX rompt. It was designed in the early 1980s for use by computer 'professionals'. While the nterface looks similar, not all commands are implemented in the same way, or at all, in very implementation. The exact format may also be subtly different. Each command typed y the user translates into a sequence of FTP protocol commands on the network. Since ome FTP commands are mandatory and others are optional, there may be a mismatch etween workstations and host. These differences in user interface and host and orkstation implementation can be a source of user frustration.

There are a number of FTP implementations that use Microsoft Windows or an X Vindow manager to present remote file systems and the FTP operations that are available n a familiar form of file and command selection boxes. This simplifies the interface to miliar mouse actions. This type of interface is provided by NetManage Chameleon as nown in Figure 6.6.

Like Telnet, FTP is not an application that should be delivered to the untrained or asual user whose main activity is commerce rather than computing. Some FTP versions llow a sequence of FTP commands to be taken from a local file. Some user tasks could be educed to a single command which would log-on to the remote machine and manipulate nd transfer files without intervention. But this command file is often not a true 'program'; cannot change its actions because of user input and it cannot recover from changing onditions or errors encountered from one run to the next. Its use must be limited to simple ecurrent tasks. An alternative but more costly option is to develop a custom application hich uses FTP as its underlying file transfer mechanism. The FTP standards still limit the ype of actions available.

8.2. Configuring FTP

There are few options with implementations of FTP. A common trap for new users is at FTP is aware of two types of files: text (or ASCII) files and binary (or image) files. If ot told otherwise, FTP assumes that it is transferring text files. When files are stored in ifferent environments (IBM, VAX, UNIX and PCs), this may corrupt any file which is not xt and where the transfer is between different types of computer. Often the file transfer ype must be changed manually once the workstation is logged-on to the remote host.

While almost any type of computer can act as a central file store for one other type of computer using FTP, file-naming conventions are a barrier to simple transfer among a range of different computer types. As an example, PCs use a name of eight characters with a three-character extension; other systems use any length of name and file version numbers that cannot be represented in the PC environment. Some implementations attempt to translate long names to DOS names in a predictable way using 'globbing'.

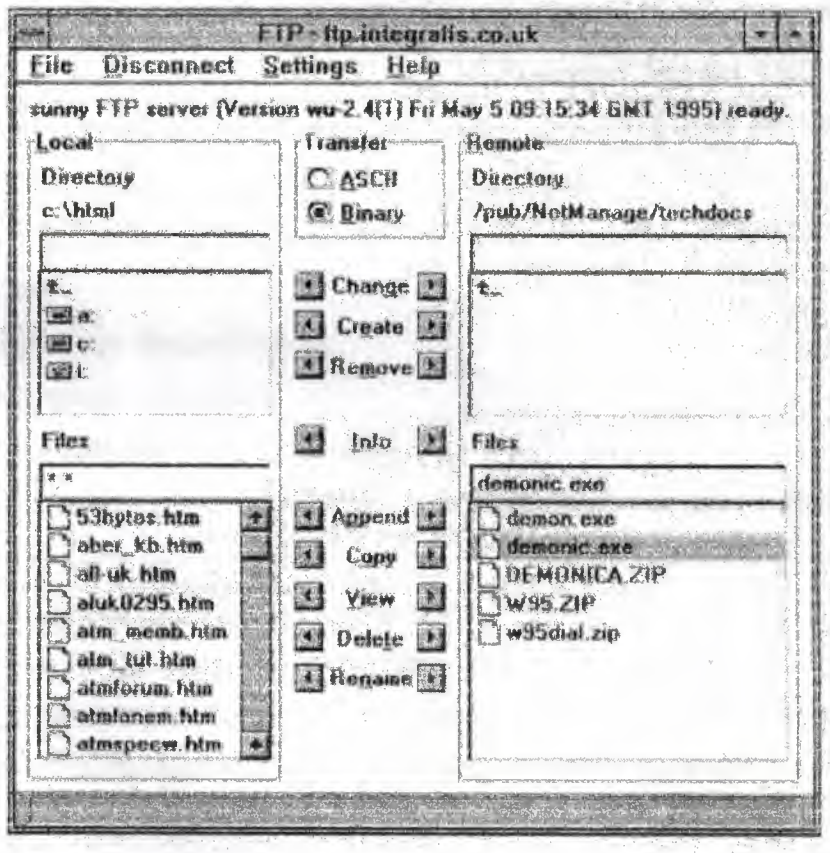

Figure 6.6 NetManage Chameleon FTP interface.

FTP uses the TCP transport layer for reliability. Optimizing TCP improves the throughput of FTP over wide area links.

6.9. Trivial file transfer protocol

The Trivial File Transfer Protocol (TFTP) is a simple program with minimal

acilities, designed to be implemented in permanent memory (PROM), so that diskless omputers may perform their initial loading of an operating system. It is often used in onjunction with the boot protocol. It is not usual for it to be used for other purposes.

TFTP lacks the security features of FTP and it is normal to disable the TFTP server n hosts which do not provide a BOOTP service.

5.10. Simple mail transfer protocol

The Simple Mail Transfer Protocol (SMTP) provides the mechanisms for TCP/IP isers to exchange mail messages. SMTP with Multipurpose Internet Mail Extension MIME) now supports binary file attachments or enclosures. Mail users are identified by heir mail address, a combination of their name or nickname and the fully qualified domain name of a particular host computer which receives and stores their mail. Some possible mail addresses are:

catherinej@in.tegv.integralis.co.uk

medavies@eclecticsys.co.uk

registrar@nic.ddn.mil

SMTP client software uses simple techniques to send text mail often directly from the source machine to a destination machine over a TCP (error-free) connection, using ASCII data streams and commands. Alternatively, the destination mail address can contain a list of machines that are to be used for forwarding mail to the ultimate destination. The sophistication in the mail protocol is that only a single copy of a mail item is sent to a group of users who are all on the same machine. The protocol between the machines can indicate a number of different user names at that machine and check that they are recognized at the destination. Where a user is not recognized, three different actions are possible:

(1) The receiving machine may refuse the mail.

(2) The receiving machine may, if it knows where that user is located, offer to forward the mail towards the user, in which case no further action is required by the sender.

(3) The receiving machine may refuse the mail and, again, knowing a possible address for that user, may suggest an alternative address.

Mail may also be sent to a distribution list retained at the receiving machine. Where implemented, the sender can request an expansion of the mail list into individual mail

ames.

For SMTP to operate, each sender must be able to convert the host and domain name nto an IP address. The network must use the domain name system, or each host must have list of destination hosts in a hosts file to map the domain name to an address. An lternative is to deliver all mail to a gateway computer which can use the domain name ervice.

The simplicity of SMTP and its ready availability mean that it is frequently used as ne gateway protocol between other proprietary mail systems, though in some instances .400 has assumed this role.

.10.1. Managing SMTP

Some of the organizational issues of SMTP arise because of the association of a erson with a particular computer. Those computers are the mail servers of SMTP and they ust always be available to receive mail. An individual workstation or PC is unlikely to ake a satisfactory mail server as its availability cannot be guaranteed. A mail server is ore likely to be a part-time task carried out by another server machine.

Building on SMTP protocols, some manufacturers have introduced the mail relay or ail store, which receives and stores mail on behalf of workstation users until they access ne store using an SMTP client on their workstation. A workstation can be a satisfactory enerator of SMTP mail, but it is a poor and inattentive recipient.

Post Office Protocol (POP2 or POPS) is an extension to the SMTP mail protocols. It lows a mail user at a machine which is only occasionally used to interact with a mail orage system which is permanently available. This mail store will act as the permanently vailable destination for that user's mail. It may also act as an intermediate gateway for utgoing mail for that user.

Staff who travel must be able to access the mail system they are registered on, for that where their mailbox resides. Alternatively, a method of mail forwarding can be devised, ut such features are beyond SMTP itself and are implementation dependent.

If mail users change location permanently, their mail addresses may change. They ill have to advise all their correspondents of that change. TCP/IP standards do not include directory service to relate people to mail addresses (or for that matter to relate application

ervers to network addresses). Worldwide research into directory services has concentrated on perfecting and implementing the ITU-T X.500 recommendations which are accepted as part of OSI. Indeed, the Internet is said to be the biggest trial of X.500 directory services in the world. Several RFCs describe how X.500 is to be used on the Internet, and the adaption required to operate over TCP/IP protocols.

Another potential problem is that the name part of a mail address can in some implementations be case sensitive. The remainder of the domain name is not.

6.11. Internet applications - World Wide Web and News

The Internet is that worldwide collection of computer systems and physical networks which uses the Internet Protocol (IP) as its common communications standard. As the Internet has moved from research, more towards academic and commercial usage, a number of specific applications for information searching and retrieval have become widely used.

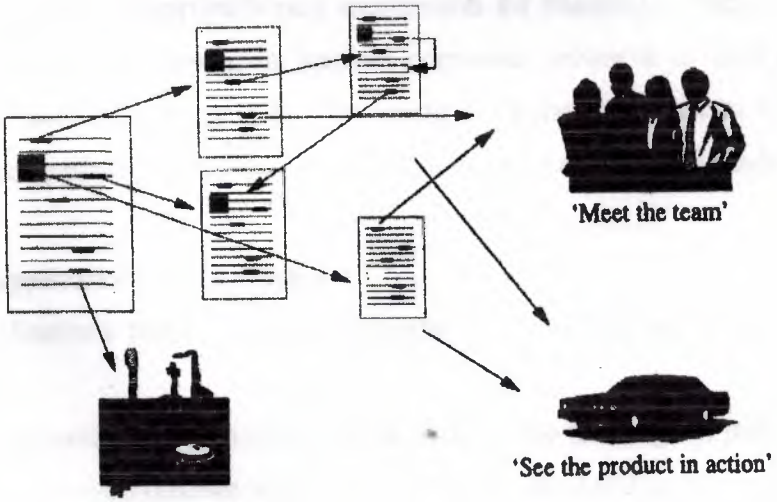

Figure 6.7 Hypertext links related topics - text, pictures, movies or sound.

These take advantage of the interconnectivity of computers provided by the Internet to present one 'integrated' information source.

A number of early systems, such as Archie and Gopher, are still in use. But the more

idely used systems now are the World Wide Web (WWW or the Web) and News rvices. These applications are again based on a client/server architecture. With WWW, e client software, referred to as browser software, deals with the presentation of formation to the user, and the server software manages the distribution of information to e browsers and the execution of special scripts (simple programs) designed to enhance the nctionality of the basic text and graphics system.

The World Wide Web service is so called because of the logical connectivity it hieves between its servers; if you attempted to draw the connectivity across the Internet these servers you would end up with a vast number of crossing lines which would semble a spider's web spanning the globe. The web is an extremely powerful system for e distribution of information. Today's web browsers can present text, graphics, audio, imation and movies to the user. Through the use of underlined hypertext links, a user es a linked set of related 'pages' or topics, each expanding on the information presented reviously (Figure 6.7). The basic idea will be familiar to those who have used Help within Microsoft Windows.

The Web provides a completely new mechanism for business applications and the istribution of information relevant to internal corporate networks as well as external nternets. In effect web technology is fast becoming a standard mechanism for accessing nformation and a completely new marketing, advertising and distribution medium.

.11.1. Internet application configuration

Internet applications require little configuration at the client end beyond choosing imple preferences.

For the web browsers, configurable options relate to the first page of information that vill be retrieved on start-up (known as the default Home Page) and to storing the location f favourite places to go. These are all simple options to set. There are more sophisticated ptions for specialist situations such as interpreting sound and animation files.

For the news system users will 'subscribe' to the newsgroups they wish to be involved n. 'Subscribing' to a newsgroup involves no additional monetary outlay, merely a decision o retrieve messages from that group and perhaps a commitment of time (and sanity) to eading them.

Subscribing is done simply by point and click operations through the graphical user terface.

ews server configuration

Configuration of a News server can be very complex. The server needs to know hich machines in the Internet it retrieves articles from (the news feeds) and which ones it nds articles to ('posts' to). The newsgroups advertised by the server must also be nfigured. A server can be set only to acquire certain groups or it can be configured only send certain groups, depending on where control is desired and bandwidth restrictions. ere are also the issues of removing old messages in the group -there are a number of tions on how this can be managed by the server - and the creation of new newsgroups. It also possible to password protect groups, which requires registration of users and their sswords. To configure a News server requires knowledge of the specific News server ing used and a clear understanding or policy on how the server is to be run.

b server configuration

The configuration of a web server is an art, requiring not only technical knowledge also artistic and graphic design skills to present information in the best possible way. ML defines the setup of web pages and the links between items on those pages and er pages in the Internet.

CONCLUSION

When Berkeley Software Distribution released Berkeley UNIX 4.2BSD, a comprehensive set of 'ready-made' communications protocols called TCP/IP became much more widely available and well-known than it had been before.

In the late 1980s, TCP/IP received a further boost to its fortunes, when Sun Microsystems published the specification for Open Network Computing (ONC), often called the Network File System (NFS). NFS adds important functions to TCP/IP and is now very widely available and regarded as an integral part of the TCP/IP protocol suite. It is particularly valuable for the commercial implementor because of the simple user interfaces that it provides.

An architecture describes three facets of communications in an abstract way which is independent of particular hardware or technology. The three aspects are; Data exchange (intercommunications), Data interpretation (interoperation), System management.

Like the OSI reference model, communications architectures are described in layers, each layer providing its own functions but using the functions of the layer below. This layering decouples the functions of one layer from another so that layered architectures are flexible; their designers can respond to changes in technology and in application software without a major upheaval for existing users.

We have introduced Transmission Control Protocol/Internet Protocol (TCP/IP) to readers in a clearly exposition by introducing and establishing the network foundation, explaining the planning and managing IP addresses, Subnetworks and Supernetworks and Routing.

REFERENCES

Reference to Books

[1] Oppliger R., *Internet and Intranet Security*, Artech House, Inc., New York NY,1998

[2] Tanenbaum, A.S., *Computer Networks*, Prentice-Hall Inc., United States of America USA, 1996

Reference to Web

[1] *"http://www.cisco.com/warp/public/535/4.html"*

[2] *"http://www.webopedia.com/TERM/T/TCP_IP.html"*

[3] *"http://www.yale.edu/pclt/COMM/TCPIP.htm"*